



Problems for ‘standard’ dispositionalist accounts of semantic content

Ásgeir Berg¹

Received: 25 March 2024 / Accepted: 9 December 2024
© The Author(s) 2025

Abstract

A popular view in metasemantics is the view that a speaker’s dispositions regarding the use of a symbol determine the meaning of that symbol for the speaker. Kripke (Wittgenstein on rules and private language, Harvard University Press, Cambridge, 1982) arguments against simple versions of semantic dispositionalism have inspired ever new versions. A recent account in the literature, due to Warren (Noûs 54(2):257–289, 2020) offers a sophisticated version of semantic dispositionalism whereby certain conditions are imposed on speaker’s dispositions to count as meaning determining—conditions we can refer to as ‘standard’. In this paper, I argue that there are a number of cases that suggest that even under such conditions, a speaker’s meaning and their dispositions can come apart, and that this suggests that dispositional accounts of semantic content presuppose that there are semantic norms, independent of a speaker’s dispositions, and thus do not explain semantic content. I conclude that, since these problems generalise to non-standard dispositionalist accounts, the dispositionalist’s prospects are dim.

Keywords Semantic dispositionalism · Rule-following paradox · Semantic content · Saul Kripke · Kripkenstein · Jared Warren

Naïve dispositionalism about semantic content is the view that a speaker’s dispositions regarding the use of a symbol determine the meaning of that symbol for the speaker: if S is disposed to say of red things, and only of red things, that they are ‘red’, for example, then S means red by ‘red’. Likewise, if S is disposed to reply with the sum of x and y when calculating ‘ $x + y$ ’, then S means *addition* by ‘+’. Such naïve dispositionalism was decisively refuted by Kripke (1982). but since the pub-

✉ Ásgeir Berg
asgeirberg@hi.is

¹ Institute of Philosophy, University of Iceland, 102 Reykjavík, Iceland

lication of Kripke's book, more sophisticated versions have been in almost constant development.

One such account, which has received quite a bit of attention in the recent literature, is due to Warren (2020).¹ According to Warren, we can solve all the problems of dispositionalism by imposing on an agent *S*'s dispositions a series of conditions that define what we can call 'standard' conditions.

In this paper I will examine a number of cases that Warren's account struggles with and argue that (a) these cases suggest that a speaker's meaning can come apart from their dispositions, and (b) that this in turn suggests that such 'standard' dispositional accounts of semantic content presuppose that there are semantic norms, independent of a speaker's dispositions—and since semantic dispositionalism is meant to be a *reductionist* account of semantic content, whereby semantic content is nothing above and beyond the speaker's dispositions, such accounts do therefore not explain what they purport to explain.

I will conclude that these problems generalise to dispositionalist accounts in general, standard or not, and hence, that the prospects for a successful dispositionalist account of semantic content are poor indeed.

1 Guardo's privilege argument

In a recent paper, Guardo (2022) has identified a problem for dispositionalist accounts of meaning that is of a slightly different kind than those identified by Kripke. Guardo refers to this as 'the privilege argument'. Despite being distinct from Kripke's arguments, it does have its roots, I believe, in his discussion of idealised conditions and *ceteris paribus*-clauses (Kripke 1982, p. 28), as well as some remarks by Paul Boghossian (Boghossian 1989, §23) and Hattiangadi (2007, pp. 106–8 and p. 117).

For the exposition of the rest of the paper, it is helpful to look at this challenge and how one might respond to it—as I will refer to these different ways later. The problem is relatively simple: It is plausible to think that *S* might give different answers when tired, than when well-rested, and so on for all kinds of mental, physical or environmental states *S* might be in. *S* might be disposed to say, for example, that $68 + 57 = 125$ in one state, but not in another. If dispositions determine meaning, we need to explain why one set of dispositions, which we can call *C*, is supposed to be meaning-determining for *S*, and not some other set of dispositions, *C**. This needs to be done in a way that is neither question-begging nor *ad hoc*. The dispositionalist thus needs to explain why my dispositions when I'm well-rested and sober are meaning-determining, and not the ones I have when I'm tired and drunk. What privileges the former over the latter?

1.1 Three ways to respond

In his paper, Guardo identifies three different ways a dispositionalist may respond to this problem. The most naïve way is simply for the dispositionalist to claim that

¹ See e.g. Lane (2022), Miller and Sultanesco (2022), Berg (2022), Guardo (2022) and Guardo (2023).

the dispositions associated with the set of conditions C are meaning-determining because those are the conditions under which S gets things right: The conditions in C determine meaning because those are the conditions under which S adds and doesn't quadd.

This is, as Guardo points out quite obviously circular (Guardo 2022, p. 872–3): S means addition by '+', the naïve dispositionalist seems to be saying, because S 's dispositions when they add are meaning-determining and they are meaning-determining because when S has those dispositions, they are adding. This is, while bad enough, not the ultimate reason why a naïve approach fails in my view, or rather, not the most useful lens through which to view it.

In the literature on the rule-following paradox, there is widespread agreement that meaning has correctness conditions—that meanings establish norms. By this, I simply mean that if S means something by '+', then certain uses or applications of '+' by S are correct or incorrect. For example, if S means *addition* by '+', then S 's assertion that " $68 + 57 = 125$ " is correct, while " $68 + 57 = 5$ " would be incorrect. If, however, S meant *quaddition* by '+', then " $68 + 57 = 5$ " would be correct, and " $68 + 57 = 125$ " incorrect.²

As this example shows, what counts as correct and incorrect in a given case is determined by what S means: If S means *addition* then one utterance is correct, and if they mean something else, something else is correct. One might then think that talk of semantic norms is superfluous, since what S means seems to fully account for what is correct and incorrect. However, speaking of correctness and incorrectness is useful, as I will emphasise later, because what S means and what S actually does can come apart: It is possible that S means *addition*, but nevertheless asserts that " $57 + 68 = 5$ "—in which case, S has made a *mistake*. If we only speak of what S means, and never of what is correct and incorrect, we have no way of making this distinction.

We can now see that when the naïve dispositionalist appeals to the purported fact that certain dispositions give the correct answers, and that others do not, they have thus taken a semantic norm for granted, i.e. their account does not explain what constitutes such a norm. But if what S means determines what is correct and incorrect in a given case, the dispositions are what needs to play this role—i.e. determining what is correct and incorrect in a given case. For comparison, contrast this with community accounts of semantic content, flawed as they may be: S is correct in their response if and only S agrees with others. Here, a clear criterion of correctness is at least given. The naïve dispositionalist offers nothing on this score, and cannot, because they specify the meaning-determining dispositions in terms of what is semantically correct, and that is circular, since what is semantically correct depends on what the meaning-determining dispositions are in the first place.

²Verheggen has called this the 'trivial' way in which meaning can be said to be normative, namely that when we use or apply a word, some ways are correct and other incorrect (Verheggen, 2011) It is trivial because it is so closely connected with the semantic content itself. Anandi Hattiangadi has also made a distinction between the normative proper and the merely norm-relative (Hattiangadi, 2007). On this view, meanings imply a standard of correct and incorrect use, but otherwise have no prescriptive powers. In this paper, I do not assume that meaning is normative in the former sense. For further discussion, see also Glüer and Pagin (1999), Hattiangadi (2006) and Glüer and Wikforss (2009).

Another way one might try to respond to the challenge is by characterising C as dispositions in ‘ideal’ conditions. Here, C are those conditions under which S would have the right dispositions, e.g. the dispositions S would have if they wouldn’t get tired, drunk, distracted, would live long enough to hear the problem and had their brain stuffed with extra matter to deal with the complex case in front of them, as well as their environment being optimal for their perceptual faculties, e.g. well-lit, quiet and not filled with a dangerous gas, etc.

Kripke does not argue against ideal conditions in the context of the privilege argument, since he does not identify it explicitly, but his arguments against *ceteris paribus*-clauses in general do apply here: If we appeal to what S would do under idealised conditions, then our account will only work if that idealisation already includes a specification that S will respond under those conditions according to what S meant. After all, S would only respond with the sum, if S meant *addition* by ‘+’, and if S meant *quaddition*, then S would respond with the quum. If the dispositionalist were to insist that things would come out right because S really meant addition, and not quaddition, they would be begging the question against the sceptic.

The real problem, I believe, is however again the same as before. Here is Kripke:

Such a counterfactual conditional [i.e. about what S would do] is true enough, but it is of no help against the sceptic. It presupposes a *prior notion of my having an intention* to mean one function rather than another by ‘+’. It is in virtue of a fact of this kind about me that the conditional is true. But of course the sceptic is challenging the existence of just such a fact; his challenge must be met by specifying its nature. (Kripke 1982, p.28. Emphasis mine.)

For Kripke, the problem with the dispositionalist appeal to idealised conditions is that it implicitly makes use of a notion that in turn requires semantic content, the speaker’s intention: If there is a problem with what the speaker meant, there will likewise be a problem with what they intended.

However, just as before, we can weaken the problem and get rid of intentions: Suppose for example that community accounts of meaning are true and that S is a part of a community where there is widespread agreement that giving an answer according to *addition* is correct. In this case, we would have no problem declaring one answer that S gives correct and all the others incorrect, and thus if S meant what their fellow speakers mean by ‘+’, S would give the answer according to addition. Ideal conditions thus suffer from a similar defect as we saw before, namely that the dispositionalist has taken for granted that there is a norm of correctness and not specified how such a norm is constituted.³

Finally, we might try to focus on ‘standard’ conditions, which are, like ideal conditions in that we take them to be those conditions under which nothing interferes with S ’s cognitive function, and so on, but they not given in terms of a counterfactual, but

³ I do not mean to endorse community accounts of meaning here. I’m merely using them for illustrative purposes—to show that by assuming a norm of correctness which is independent of S ’s dispositions, we can give a coherent reply to the problem. It is precisely this possibility which is closed off to the semantic dispositionalist. Elsewhere, however, I do defend a communitarian solution to the paradox (see Berg 2022).

some kind of story that is meant to single out these conditions non-circularly, perhaps by an appeal to human biology or evolution. Taking this route seems more promising, since it is not as obviously circular as the other two.

For the rest of the paper, I will be concerned with demonstrating the problems with this kind of approach. I will take Warren's account as a representative, since having something concrete to focus on will make exposition easier. This should not lead to any loss of generality, even in cases where I criticise the specifics of Warren's account, since, as Guardo puts in his paper, "the current state of the debate on dispositionalism can practically be seen through the lens of Warren's discussion" (Guardo, 2022). Some of what I have to say, however, will be applicable to dispositionalist accounts of semantic content more broadly, and not just those that end up appealing to standard conditions.

2 Warren's 'standard' dispositionalist account

In this section, I will give a quick overview of Warren's account and how he responds to the problem of error, one of Kripke's main complaints against dispositionalist accounts of semantic content.

2.1 The problem of error

To be brief, the problem of error is as follows: Any account of meaning must make room for the possibility that *S* might make mistakes; to explain why *S*'s disposition to incorrectly add isn't simply a disposition to correctly quadd. Any purported mistake must not merely signal a change in meaning.⁴ To solve this problem, Warren tries to appeal to standard conditions, however in quite a complex manner, involving three different conditions, meant to "screen out" different kinds of errors.

Warren starts off by defining a situation as (1) *normal* iff no internal or external factors are interfering with *S*'s general cognitive functioning. Despite this being only the first of Warren's conditions, one might think, right here at the outset, that taking the notion of "interference" with *S*'s cognitive faculties for granted begs the question against the sceptic. For example, *S* would presumably have a different set of dispositions in the presence of oxygen, than in its absence. We wouldn't want to say that oxygen is thereby *interfering* with *S*'s cognitive faculties, in which case we seem to come dangerously close to have implicitly defined normal situations as those situations in which *S* is an adder, since we then need a prior understanding of which factors interfere, and which not.

In order to stave off charges of circularity, however, Warren appeals to the "causal-historical situations in which Ludwig's [i.e. our *S*'s] general cognitive mechanisms were evolving and developing" (Warren 2020, p. 30f) and so, the fact that *S*'s species has evolved in an oxygen-rich environment would allow us to explain why we can deny that oxygen interferes, and *mutatis mutandis* for other factors. I am not sure

⁴See e.g. Wikforss (2001), p. 208–9 for discussion.

how persuasive one should find this line of thought, however, but I will not press the point here.⁵

Next, Warren defines S 's dispositions to φ as (2) *general* iff in the overwhelming number of situations in some class of situations C , $S \varphi$ s.⁶ The point of this definition, however, is to account for the fact that it is unlikely that S would give the very same answer in every situation where they are asked about a given sum— S might be tired, sick, etc, and so sometimes give the sum and sometimes another answer. For this reason, Warren weakens the requirement and only requires that S gives the same answer in the *overwhelming* number of situations of a given class C —and not in all such situations, as one might first assume.

This is promising, but there is a problem. Is it really so clear what counts as a class of situations C ? Warren seems to assume, for example, that every time S gives a reply to a question of the form ' $n+m$ ', S is giving an answer to the same kind of problem, and so that these form a class. But as the familiar notion of synonymy shows, this is not so clear at all. It could be, for example, that sometimes when S is using the symbol 'bank', S is referring to a financial institution, and sometimes the land along a river or a lake. The sceptic could then ask what fact about S makes it the case that when S is using the symbol '+' today it belongs to the same class of situations as when S is used the symbol '+' yesterday?

It would take some care to distinguish these cases without referring to their meaning, and it is unclear to me how Warren could respond, since his formulation of the condition refers to the *form* of question being asked, but as the sceptical problem makes clear, we could divide S 's use of '+' up in such a way that S means quaddition by '+' in some cases where we'd want the verdict to be addition. To see this, recall that Warren's purpose with this condition is to allow us to identify mistakes: if S only occasionally answers according to quaddition when adding $68 + 57$, then we can refer to this condition to explain why those answers should not show that S meant quaddition—after all, in the overwhelming number of cases when computing this sum, S answers according to addition, and hence, Warren wants to say, S was making a mistake.

The sceptic can nevertheless *still* ask why S isn't computing according to quaddition when S gives the answer according to quaddition and addition when S answers according to addition? After all, S can sometimes use the term 'bank' to refer to a river bank and sometimes to a financial institution. If we would focus on the mere form of the expression, there could be a case where S , who only rarely speaks of the land along a river or a lake, could simply never speak of river banks, since their overwhelming use regarding that symbol refers to financial institutions. That cannot be, and so there should be a way of dividing S 's use up in such a way that S uses 'bank₁' in overwhelming number of cases to refer to a river bank and 'bank₂' in over-

⁵ Here, Warren has teleosemantic and related accounts of content in mind, e.g. Millikan (1984, 1989)s, Dretske (1994), and to some extent Dennett (1981). I don't take Warren to mean that his account is itself teleosemantic, but rather that the semantic dispositionalist can use the resources of teleosemantics to explain why some conditions are standard and some are not.

⁶ Actually 'M-general' but this is a complication we need not bother with.

whelming number of cases to refer to a financial institution (where the subscript is suppressed in actual use).

But what fact about S determines that they are not in this situation with regards to ‘+’? S ’s answer according to quaddition could thus not be ruled out as a mistake, because S ’s use of the term ‘+’ could be divided up or gerrymandered in such a way that precisely that case was an example of quaddition, and not addition, analogously to the example of ‘bank’. One difficulty that should be noted is that an answer to this problem must be able to generalise to the symbol ‘bank’, too (i.e. even if we could explain why ‘+’ can’t refer to two functions, depending on S ’s meaning, then we’d still have to explain the difference where the symbol can).

One could object and say that in cases of ambiguities, we would see two distinct sets of stable dispositions, say one for river banks and one for financial institutions, but in the cases of error, we would not see any such stable regularity; sometimes S subtracts 1 and sometimes they add 5, etc. The problem with this reply is twofold, I believe: First of all, since Warren defines the condition in terms of classes of situations in which S makes their utterances, there is still a problem with deciding which set of dispositions a given utterance belongs to, even if S had two sets of stable dispositions. Suppose for example that S utters the statement “The shop is next to the bank”. Was this a mistaken claim that the shop is next to the financial institution or an utterance of the truth that it is next to the river bank? That S has two sets of stable dispositions does not help, because S can make a mistake relative to both of them, and Warren hasn’t told us how to distinguish these two cases.

Second, there is a well-defined function that corresponds to any one-off mistake, and so it should be possible for S to refer to that particular function instead of *addition*. But if S overwhelmingly utters ‘ $2+2=4$ ’ and only once utters ‘ $2+2=9$ ’, Warren’s condition means that it is impossible for S to use the symbol ‘+’ to refer to the function that takes 2 and 2 as inputs and outputs 9. But surely S can do that?

This brings us to the last of Warren’s conditions. He defines (3) a judgement α to be *stable* iff the ratio of non- α answers approaches zero as S ’s number of independent attempts to give a judgement increases. According to these definitions, Warren then stipulates that S means ‘addition’ by ‘+’ iff S has the disposition to answer with the sum and those dispositions are general and stable over normal situations. I will refer to this as S being in standard conditions, or S ’s dispositions being standard.

Before moving on to more general problems with Warren’s account, I want to raise a rather serious problem with his stability condition, namely this: We can re-run a temporal analogue of Krikpe’s *tabair* example on S ’s successive tries.⁷ In their hypothetical attempts to calculate the sum, S makes an attempt at some time t_0 , and again at t_1 , t_2 , and so on up to t_n , for some number of tries n . Now further stipulate that as n grows, the ratio of S ’s answers that do not correspond to *addition* approaches zero. In this circumstance, Warren’s stability condition is met.

Even so, the sceptic can *still* ask, what fact about S constitutes S *adding* at each step, and not *temporally quadding*, where temporally quadding is to *quadd* at every

⁷This is the problem of explaining why S ’s means *table*: by his use of the word ‘table’ when S finds himself directly beneath the Eifel Tower instead of *tabair* where *tabair* is....My objection to Warren’s stability condition is inspired by a similar, but more general, temporal argument made by Guardo (2023).

t_n where n is prime, but *adding* otherwise? Warren cannot appeal to the fact that S 's dispositions are only stable for addition, and not temporal quaddition, because S 's answers that do not fit with temporal quaddition are also approaching zero with each successive try, since the average space between the primes approaches infinity as S goes along. Thus, the stability condition does not weed out every deviant function in the way intended.⁸

One might think that Warren could simply fall back on his other conditions, but unfortunately, we can do the same thing for the condition of generality. The thought is this: In order for S to have a general disposition to add, S must give an answer according to addition in an overwhelming number of cases in some class of situations C . These answers can in principle be ordered temporally, just as we did before. And if that is the case, then Warren's condition of generality is also fulfilled by temporal quaddition: If S gives a reply according to addition in the overwhelming number of situations in some class of situations C , then S also gives a reply according to temporal quaddition in the overwhelming number of situations of C —after all, temporal quaddition agrees with addition in the overwhelming number of cases.⁹

3 Problems for standard dispositionalist accounts

The problems just outlined for Warren's three conditions are peculiar to his account, and not standard dispositionalist accounts in general—although I do think they demonstrate the difficulty in finding the right conditions on dispositions and may be enough to sink Warren's account. In this section, I will discuss problems that do generalise, however; problems that should be difficult for *any* standard dispositionalist account of semantic meaning to solve. These problems point to, as I hope to show, a deeper problem with semantic dispositionalism, namely that it confuses performance with correctness, just as Kripke originally claimed.

Now, suppose S sees some cows standing in a field on a dark and foggy night.¹⁰ Because S 's perceptual faculties are as they are, S is actually disposed to call the cows 'horses'—and thus get things wrong. What should the standard dispositionalist say? Here is Warren, discussing a different, but analogous case:

So when Ludwig slips up and forgets to carry the “1” in the 17th step of an addition problem, he is making a mistake in computing the sum, rather than correctly computing the quum, because he has a stable, M-general disposition to carry the “1” in the 17th step of this problem in normal situations. (Warren 2020, p. 17)

⁸One might think that this requirement is too strong. For this point, see below.

⁹There is a quibble here: The classes under consideration are infinite, and so it is not clear what it means to count the cases. Warren avoids this by giving a statistical interpretation of 'overwhelming'.

¹⁰This example is based on Boghossian's well-known horsey cow case, which in his hands is meant to demonstrate a problem for community accounts of semantic content (see Boghossian 1989).

But by stipulation, *S* is *not* in a normal situation—they are out on a dark and foggy night.

In this passage, Warren seems to say that what *S* means on such an occasion is determined by what *S would* say if *S were* in a normal situation. But this won't do, since saying that *S* means in non-normal situations what *S would* mean if *S were* in normal situations would just reduce dispositions in standard conditions to dispositions in ideal conditions, and hence, as we've already seen, make the account circular. Maybe Warren has a different way to respond than merely insisting that what *S* means in non-normal situations is determined *S*'s dispositions in normal situations, but I can't see it—locating the meaning-determining dispositions in normal situations will always have the problem of explaining what we mean in non-normal situations.

That may be wrong, but again, this is not the crux of the matter. In general, we should be able to explain mistakes that are not simply linguistic in nature. It is, after all, possible to use a word correctly, while still misapplying it. If *S* means to say that those cows over there are horses, because *S* really thought that they were horses, then *S* has not made a mistake concerning the meaning of words, but about what they saw: relative to *S*'s intention of using the word, *S* was correct.

The worry I'm pressing here is that it seems impossible to fill this gap with more conditions on dispositions, since we seem to need to explain two distinct kinds of mistakes: how *S* can intend to use a word in a specific way, on the one hand, and what makes a given application correct, on the other. For the dispositionalist, recall, *S* means *cow* iff *S* is disposed to apply 'cow' to cows, and so, the dispositionalist account does not seem to have anything to say about the former kind, even if it could explain the latter.

3.1 Empty predicates

This problem can perhaps be brought out better by imagining that unbeknownst to *S*, the extension of the predicate they are applying is in fact empty. Suppose for example that *S*'s friend has described to them a horse-like creature, with one horn and typically white, called 'unicorn'. *S* then sees some white horses in the distance and forms the belief that they are seeing some unicorns. This is incorrect, as they are horses, but *S* nevertheless *meant unicorn* by their use of the term 'unicorn'—and correctly so. After all, they thought they had seen what their friend had described and that was unicorns.

In this case, however, we cannot explain their meaning by them applying the term 'unicorn' to unicorns and only to unicorns, or by them saying of unicorns and only of unicorns that they are unicorns—after all, there are no unicorns.

This is merely meant to draw out the problem I was describing above, namely the possibility of correct use with incorrect application, but nevertheless points to, I believe, a general problem with dispositionalism, that it cannot explain the meaning of empty predicates. I may, after all, use the term 'unicorn' correctly and even apply it to pictures of unicorns, descriptions, etc. Nevertheless, it is simply false that I mean *unicorn* by 'unicorn' because I apply the term 'unicorn' only to unicorns, or because

I say of unicorns and only unicorns that they are ‘unicorns’—at best that would entail that I mean the same thing by any empty predicate.¹¹

The dispositionalist might try to object by saying that we shouldn’t say that I mean φ by ‘ φ ’ iff I apply ‘ φ ’ to φ s and only φ s after all, but because I am disposed to use the term φ *correctly*. Such a reply, however, would explicitly assume a standard of correctness, independent of S ’s dispositions, and since what is correct or incorrect in S ’s use of a given predicate depends on what S meant in the first place, this modification of the view would implicitly rely on what S meant in order to explain S ’s meaning—therefore presupposing a semantic norm, rather than explaining it.

3.2 Correct use versus correct application

But just as S should be able to *use* a word correctly, but nevertheless *apply* it incorrectly, S should be able to use a word incorrectly, but apply it correctly—for example, if S thinks that ‘cow’ means *horse* and mistakenly, relative to their own intention, calls some cows ‘cows’. This is a correct application, but S is nevertheless using the word incorrectly, since they think it means *horse*.¹² Here, we do not even need to posit that S is in non-normal conditions, as this could occur on a clear and sunny day. But how can the semantic dispositionalist explain this? After all, for them, if S is disposed to call cows ‘cows’, then S means *cow* by ‘cow’—by stipulation, however, S means *horse* by ‘cow’. Again, an appeal to some norm of correctness seems to be needed, other than just S ’s dispositions.

It is possible to give yet another objection, again without supposing that S is in non-standard conditions. Recall, that Warren claims that S ’s mistake in following an algorithm is explained by S ’s dispositions in standard conditions. It seems to follow, however, that there are cases where S cannot make a mistake in standard conditions. Suppose, for example, that S , like Putnam, isn’t quite clear about how elms are different from beeches.¹³ S might then consistently call both elms and beeches, ‘elms’. In this case, S would have a general disposition to say that both elms and beeches are ‘elms’. It should be quite intuitive to say that S is making *mistakes* in these cases—that they are not using the word correctly, even relative to their own meaning, since it does not follow that by having these dispositions, S *thinks* that ‘elm’ means *beech or elm*. S might well agree that elms and beeches are different, and that the word ‘elm’ only picks out one of these things, and yet have the dispositions described.

But the standard dispositionalist cannot allow that S is making mistakes here, since their account predicts that by the word ‘elm’, S means *beech or elm*. So, the standard dispositionalist must allow that S thinks that ‘elm’ and ‘beech’ picks out two different kinds of trees, but nevertheless means *beech or elm* by ‘elm’. This is maybe not a direct contradiction, but an odd consequence nonetheless.

¹¹ In a way, then, I’m making a distinction between sense and reference, and arguing that dispositionalism is impotent with regards to the former. It is unclear to me how other accounts would fare on the problem of empty predicates, but I do believe it shows that we cannot focus on application—correct *use* must come first in the order of explanation.

¹² For this distinction between correct ‘use’ and ‘application’, see Millar (2002).

¹³ This example is of course from Putnam (1975). A similar case could be constructed using Burge’s arthritis case (see Burge 1979).

Similarly, we can also imagine that *S* consistently misapplies the words ‘elm’ and ‘beech’ in such a way that there is no discernible rule in how *S* uses the words: *S* might sometimes call elms ‘elms’ and sometimes ‘beeches’ and they might sometimes call beeches ‘beeches’ and sometimes ‘elms’. In this case, there doesn’t seem to be any disposition that *S* has such that *S* calls anything ‘elm’ or ‘beech’ in an overwhelming number of situations in some relevant class of situations, as Warren’s definition of standard conditions requires. Warren’s account would predict that *S* means nothing by his utterances of ‘elm’ and ‘beech’—and that again seems wrong.

In both of these cases, *S*’s actual dispositions to use words can come apart from how *S* means or intends to use a word—i.e. *S* has a disposition to apply the word in a certain way, but use it in a different way. These kind of cases of course suggest that perhaps *S*’s dispositions would be socially conditioned, and so there would be a way for *S*’s dispositions to defer to the meaning of ‘elm’ and ‘beech’ that are prevalent in *S*’s speech community (perhaps by *S* having a higher-order disposition to override his faulty ones),¹⁴ The problem with this sort of reply is not that it is necessarily false, but that it takes something outside of the agent to provide a criterion of correctness for their use of symbols—and that, I would think, throws out the semantic dispositionalism and leaves the problem of how to explain the constitution of such a social standard.

3.3 The argument from normativity

In the cases just considered, dispositionalist accounts struggle to account for mistakes that result from *S*’s usage deviating from their intention: either *S* means to say that the cows are horses, even though they are not, or *S* means to say that some cows are ‘cows’, but thinks they are talking about horses, because they thought ‘cow’ means *horse*. Likewise, *S* might be confused about elms and beeches, without thereby having a non-standard meaning for those terms or meaning nothing at all. This led to the thought that the semantic dispositionalist is missing a norm of correctness which is independent of *S*’s dispositions.

This objection to dispositionalist accounts of semantic content can perhaps better be seen in relation to the following excerpt from Warren’s paper:

I do not claim that errors are *impossible* in normal situations. I admit that some rogue interfering factors may have slipped through the net. It might seem strange to say this, since a class of situations where error is impossible seemed the obvious goal, but that seems too much to hope. In fact, Paul Boghossian has offered detailed arguments purporting to establish that there is no way to finitely and non-question beggingly single out a class of situations in which all errors are impossible. (Warren 2020, p. 15)

As we can see, Warren takes the problem of error to be a problem of *eliminating* errors.

¹⁴ In an endnote 34 of his book, *Shadows of Syntax* Warren seems to allow for something like this.

As the cases we've discussed so far should show, the problem is rather to explain how we can *identify* errors—explain why an error is an error, and not a correct reply, relative to a different concept. If that is the correct way of viewing the problem, Warren's admission that 'rogue interfering factors may have slipped through the net' is tantamount to admitting that dispositionalist accounts can *never* solve the problem of error; to answer the sceptic we need a fact about *S* that explains why *S* is adding, but making a mistake, and not correctly quadding, and thus, any error that slips by one of Warren's screens would suffice for the sceptic to claim that there is no such fact.

The point can perhaps be better seen through the following scenario: Suppose all of Warren's conditions are fulfilled and that *S* has been asked if those cows over there are horses or not. How would we go about determining if *S*'s reply is metalinguistically correct? That is to say, if *S* means *horse* by 'cow' and thinks those are cows, he would be, in this sense, wrong to say that the animals are 'cows'. Likewise, they are metalinguistically correct if they mean *cow* by 'horse' and say that they are 'horses'. The question is: what determines if *S*'s utterance is correct, relative to what *S* means by their words?

Warren's answer is that *S*'s dispositions in standard conditions determine what *S* means, and since *S* is in standard conditions, what *S* is disposed to do determines what *S* means. He has admitted, however, that a standard dispositionalist can never give a set of conditions that rule out every mistake. But relative to what was the mistake a mistake, then? As we've seen, what counts as a mistake depends on what *S* meant, and so, it can't be *S*'s dispositions that set this standard, because they are what determine *S*'s meaning, and hence *S* cannot make a mistake if *S* has those very dispositions—they are, again, supposed to be constitutive of *S*'s meaning. By allowing that no set of conditions on dispositions can identify every mistake, Warren admits that there is a standard of correctness, independent of *S*'s dispositions, but the very notion of a semantic mistake depends on what *S* means, and so, on a dispositionalist account of meaning, what dispositions *S* has. But if dispositions cannot set this standard, it follows that dispositionalist accounts of meaning are false.

The other option, that *S* simply cannot make mistakes in standard conditions also seems hard to swallow—and the reasons for this are of course the very ones that move Warren to admit that no condition could ever rule out every mistake. One such reason, mentioned by Boghossian, is that all kinds of background beliefs might influence *S*'s belief about what they are seeing (Boghossian 1989, p. 539–540). *S* might for example see a horse in the distance and think that it is a cow *because S thinks that this is how cows look like*. In this case, *S* would presumably have a disposition that fulfills Warren's conditions, and so, mean *horse or cow* by 'cow'.

One way of responding might be to require that *S*'s dispositions only count if they are such that *S* would not change their mind if they received more information.¹⁵ But this seems problematic, because if *S*'s dispositions entail that they meant *horse or cow* by 'cow', and so, that *S* was *correct* in their judgement, then there would be

¹⁵There are accounts in the literature that make a similar move, e.g. Johnson and Nado (2014, 2016, 2017). For Johnson and Nado, *S* means *red* by 'red' iff *S* would be disposed to apply 'red' to red things, if *S* had all the relevant information. Relevant information, given Boghossian's point about dispositions and belief, seems to presuppose semantic content. For further criticism and discussion, see Andow (2016) and Nyquist (2020).

something irrational about *S* later changing their mind about the case, for example, by learning how cows actually look like. After all, they would be learning how *cows* look like, not how *cows or horses* look like. *S* should therefore only accept the correction, if they meant *cow* by ‘cow’ all along, and their dispositions have already settled this, according to the dispositionalist—they meant *cow or horse*.

The prospects of finding conditions on dispositions that can rule out *every* error thus seem dim, just as Warren admits. But his response suggests that he thinks of the sceptical problem as one where we must account for how a speaker can track pre-existing semantic facts, rather than as a challenge to explain how such facts are constituted in the first place. And this way of thinking does not seem unique to Warren, but a general assumption of dispositionalist accounts of meaning.

My complaint against semantic dispositionalism thus boils down to the claim that such a view does not explain how semantic norms get constituted, but rather takes them for granted. This is an argument from normativity, even if it is not Kripke’s (or rather, not obviously so). I do not claim that the dispositionalist is wrong because meanings imply that *S ought* to do something, but rather because meanings imply that some things are correct and some things are incorrect—establishing *norms*, and that semantic dispositionalists cannot, by the very nature of their account, explain how such a norm could be constituted independently of the dispositions of the speaker, as the examples above suggest they need to be able to do.

4 Conclusion

In this paper, I’ve offered a criticism of standard dispositionalist accounts of meaning, focusing on a recent account by Jared Warren. I’ve argued, through a number of examples, that a speaker’s intention of using a word and their actual disposition of how to use it can come apart. These cases suggest that some standard of correctness for the use and application of words, independent of a speaker’s dispositions, is necessary in order to account for semantic content.

I’ve further diagnosed this problem as dispositionalists misconceiving the problem, of one where a speaker must track pre-existing semantic facts, rather than explaining how such facts are constituted in the first place. If that’s right, the prospects for a pure dispositionalist account of semantic content should be considered quite poor.

Declarations

Conflict of interest The author declares no Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view

a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Andow, J. (2016). Zebras, intransigence & semantic apocalypse: Problems for dispositional metasemantics. *Philosophia*, 44(1), 53–62.
- Berg, Å. (2022). Rules as Constitutive Practices Defined by Correlated Equilibria. *Inquiry*, 65, 1–35.
- Boghossian, P. (1989). The rule-following considerations. *Mind*, 98(392), 507–549.
- Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, 4(1), 73–122.
- Dennett, D. (1981). *The intentional stance*. MIT Press.
- Dretske, F. (1994). If You Can't Make One, You Don't Know How It Works. *Midwest studies in philosophy*, 19(1), 468–482.
- Glüer, K., & Pagin, P. (1999). Rules of meaning and practical reasoning. *Synthese*, 117(2), 207–227.
- Glüer, K., & Wikforss, Å. (2009). Against content normativity. *Mind*, 118(469), 31–70.
- Guardo, A. (2022). Yet another victim of Kripkenstein's monster: Dispositions, meaning, and privilege. *Ergo*, 8(0).
- Guardo, A. (2023). A dilemma for dispositional answers to Kripkenstein's challenge. *Minds and Machines*, 33(1), 135–152.
- Hattiangadi, A. (2006). Is meaning normative? *Mind & Language*, 21(2), 220–240.
- Hattiangadi, A. (2007). *Oughts and thoughts: Rule-following and the normativity of content*. Oxford University Press.
- Johnson, M., & Nado, J. (2014). Moderate Intuitionism: A Metasemantic Account. In A. Booth & D. Rowbottom (Eds.), *Intuitions*. Oxford University Press.
- Johnson, M., & Nado, J. (2016). Intuitions and the Theory of Reference. In J. Nado (Ed.), *Advances in experimental philosophy and philosophical methodology*. Bloomsbury.
- Johnson, M., & Nado, J. (2017). Actual versus counterfactual metasemantics: A reply to Andow. *Philosophia*, 2(45), 717–734.
- Kripke, S. (1982). *Wittgenstein on rules and private language*. Harvard University Press.
- Lane, E. (2022). Semantic dispositionalism and the rule-following paradox. *Metaphilosophy*, 53(5), 685–695.
- Millar, A. (2002). The normativity of meaning. *Royal Institute of Philosophy Supplement*, 51, 57–73.
- Miller, A., & Sultanescu, O. (2022). Rule-Following and Intentionality. In N. Edward (Ed.), *The Stanford Encyclopedia of Philosophy, Summer 2022*. Zalta. Metaphysics Research Lab: Stanford University.
- Millikan, R. (1984). *Language, thought, and other biological categories: New foundations for realism*. MIT Press.
- Millikan, R. (1989). Biosemantics. *Journal of Philosophy*, 86(6), 281–297.
- Nyquist, M. (2020). On complete information dispositionalism. *Philosophia*, 48(5), 1915–1938.
- Putnam, H. (1975). The Meaning of Meaning. *Minnesota Studies in the Philosophy of Science*, 7, 131–193.
- Verheggen, C. (2011). Semantic Normativity and Naturalism. *Logique Et Analyse*, 54(216), 553–567.
- Warren, J. (2020). Killing Kripkenstein's monster. *Noûs*, 54(2), 257–289.
- Wikforss, Å. (2001). Semantic normativity. *Philosophical Studies*, 102(2), 203–226.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.