Frontloading and Fregean Sense: Reply to Neta, Schroeter, and Stanley

David J. Chalmers

I would like to thank Ram Neta, Laura Schroeter, and Jason Stanley for their generous and probing comments on *Constructing the World*. [1] All three of them home in on key issues in the argumentation of the book. Neta and Schroeter focus on the argument for A Priori Scrutability, and especially the key frontloading argument for that thesis. Stanley focuses on whether the semantic values that are defined in terms of scrutability can truly serve as Fregean senses. All of them raise serious challenges that need to be addressed.

1 Frontloading Principles (Neta and Schroeter)

In chapter 3 of *CTW*, I argue for a Conditional Scrutability thesis: for all ordinary truths M, subjects are in a position to know (from the armchair, given ideal reasoning) that if $PQTI$, then M. (Here $PQTI$ specifies complete physical, phenomenal, that's-all, and indexical information.) In chapter 4 I argue from here to an A Priori Scrutability thesis, holding that subjects are in a position to know this a priori.

Neta and Schroeter both reject the step from Conditional Scrutability to A Priori Scrutability. If the former but not the latter is true, then subjects are in a position to know that if $PQTI$, then M, but they are not in a position to know this a priori. That is, knowledge of the conditional will have ineliminable empirical justification. Schroeter makes a direct case for this empirical justification, while Neta (and also Schroeter) focuses on my argument against it.

Both of them focus on the argument from frontloading, so it is useful to outline that argument here. The frontloading argument says that 'If $PQTI$, then M' is justified by empirical evidence E, then 'If $PQTI\&E$, then M' is justified independently of E. The argument from knowability in

chapter 3 made the case that any claim that is justified by empirical evidence can be justified by a restricted class of core evidence concerning primary and secondary qualities and phenomenal properties. Given that this core evidence is itself a priori scrutable from $PQTI$, it follows that 'If $PQTI$, then M' can be justified independently of empirical evidence, so that M is a priori scrutable from $PQTI$.

The key to this argument is a frontloading principle along the lines of the following:

> *Frontloading Principle*: If one is justified in believing M based on empirical evidence E, one can be justified in believing 'If E, then M' with justification independent of E.

I give two arguments for frontloading theses a formal and an informal argument. The formal argument pivots on the following principle"

> *Diachronic Frontloading*: If $cr^*(M) = \phi$ at t_2, and one acquires total relevant evidence e between t_1 and t_2, and $cr^*(M|E)$ is defined at t_1, then $cr^*(M|E) = \phi$ at t_1.

Here 'cr^*' represents a rational credence, the credence (or range of credences) that a subject should rationally have at a given time. The Diachronic Frontloading principle is a straightforward conseqence of the orthodox Bayesian principle of conditionalization. Diachronic Frontloading is very much in the spirit of the informal frontloading principle, since it says that if a certain unconditional credence in M is rational upon acquiring E, the corresponding conditional credence in M given E was rational even before acquiring E. This temporal priority claim strongly suggests the epistemic priority claim that the conditional credence in M given E is justified independently of E at t_1. This epistemic priority claim in turn suggests that even at t_2, after E has been acquired, the same conditional credence in M given E is justified independently of E at t_2. This last claim, called Synchronic Frontloading in *CTW*, is extremely close to the informal frontloading principle.

One can also argue informally for the frontloading principle as follows. Given that E justifies M, then one could in principle (i) suspend judgment concerning E, (ii) suppose (for the purposes of conditional reasoning) that E, (iii) conclude (under this supposition) that M, with justification provided by E's support for M, and (iv) discharge the supposition, yielding a justified conditional belief in M given E. This conditional belief is justified even though one has suspended judgment concerning E, so that E played no non-suppositional role in its support. So the conditional belief in M given E is justified independently of E.

Here is a device that does not appear in *CTW* but that I have found useful in making a case for frontloading principles. There are two rooms, the Knowledge Room and the Supposition Room.

We take twin subjects in the same state and put one in each room. In the Knowledge Room, we convey the knowledge that E, perhaps through testimony. From there the subject engages in some reasoning and concludes that M.

In the Supposition Room, we ask subjects to suppose that E. We tell them that E may be true, but we do not confirm this. Where the subject in the Knowledge Room reasons from knowledge of E to knowledge of C, the subject in the Supposition Room reasons in a parallel way from the supposition that E to the conclusion (conditional on that supposition) that C. This subject then discharges the supposition to conclude that if E, then C.

The key claim is that if the reasoning in the Knowledge Room is justified, then the corresponding reasoning in the Supposition Room is justified. That is, if the judgment that C is justified in the Knowledge Room, the judgment that if E, then C is justified in the Supposition Room. On first sight, this claim is intuitively obvious. The two subjects have engaged in near-parallel reasoning, and if the reasoning is good in one case, it is also good in the other case.

Perhaps there are some varieties of epistemic evaluation (perhaps including reliabilist evaluation) on which the two sort of reasoning can be evaluated differently. But it is hard to deny that there are also varieties of epistemic evaluation on which they must be evaluated the same. Following the methodology of the book (see especially the discussion of reliabilism on pp. 66-9), it is these varieties that are most relevant to my claims.

Neta and Schroeter put forward epistemological views that they take to be incompatible with frontloading principles, and that they take to give reason to reject those principles. I think that the plausibility of frontloading principles (deriving from arguments such as the above) itself gives reason to reject views incompatible with them. In fact, I think that there are views in the vicinity of Neta's and Schroeter's that are compatible with frontloading principles. These views lead to no objection to my arguments. There are other views in the vicinity that are incompatible with frontloading principles. I will argue that we have good reason to reject these views.

2 Reply to Neta

Neta raises two major objections to frontloading principles. One is an objection to the formal Bayesian version, while the other is an objection to the informal version.

Neta's objection to the formal version starts with the observation that conditionalization is not always rational. If one's prior credences are irrational, then may be rationally better not to conditionalize. He also notes that there are plausibly constraints of which of our "ultimate priors"—prior

3

probabilities we should have prior to any evidence—are rational. These constraints hold whether or not beings without any evidence could appreciate them.

So far, this poses no objection to the frontloading principles above. A Bayesian can naturally hold that any such constraints are justified a priori. However, Neta goes on to argue that there may be a posteriori constraints on the rationality of ultimate priors.

His key case involves two conditionalizers, Harry and Sally, who have different ultimate priors regarding sequences of coin flips. Harry regards all sequences of results as equally likely, while Sally regards relatively uniform results as more likely. Given a sequence of 12 heads, Sally will conditionalize to expect heads in the future, whereas Harry will not. So Sally seems to be in a position to learn from evidence more efficiently than Harry. This tends to suggest that Sally's priors are more rational than Harry's.

Neta suggests that Harry's priors are irrational *because* they do not allow efficient learning. Furthermore, he suggests that if priors do not allow efficient learning *within a human lifetime*, they are also irrational. But the length of a human lifetime cannot be known a priori, so for such a prior it may be an empirical fact that it does not allow efficient learning with a human lifetime. Neta suggests that if so, it is an empirical fact that this prior is irrational.

There is a lot to say about this argument. I think that the most important problem in the argument is at the final step. The fact that it is a posteriori that Harry's priors do not allow human efficient learning does not entail that the priors are excluded a priori. The premise of this inference suggests at best that there is an a posteriori justification for excluding these priors. But the existence of an a posteriori justification does not entail the nonexistence of an a priori justification. Many a priori truths have both a priori and empirical justifications.

Furthermore, there is good reason to think that just this structure obtains in Harry's case. It is plausible that the fact that Harry's priors will not allow efficient learning in his lifetime *indicates* that those priors are irrational (thereby providing an a posteriori justification for excluding them). But it is also plausible that this fact is not what *makes* those priors irrational. There is plausibly another, deeper reason for this irrationality that does not involve facts about human lifetimes. Neta has given us no reason to think that this deeper reason is not a priori. For example, given the widely held view that there are a priori considerations favoring simplicity and uniformity, Sally's priors will be favored a priori over Harry's.

I think this analysis provides sufficient reason to reject Neta's argument. Another issue with the argument is worth noting, though. Even if it is an empirical fact that Harry's prior credences are irrational, this does not entail that those credences are justified a posteriori. Prior credences

are by definition prior to empirical evidence, so they cannot be empirically justified. (This applies to both doxastic and propositional justification, unless one holds that an agent's credences can be propositionally justified by evidence the agent doesn't possess). So insofar as prior credences are justified at all, they must be justified a priori.

Here Neta has two choices. He could allow that prior credences are justified independently of empirical evidence, while holding that empirical factors are part of what make them so justified. This would be analogous to a case in which empirical facts about one's brain make it the case that one is a priori justified in believing that 2+2=4, or perhaps the case of the "lucky mechanism" discussed in chapter 8 in which environmental facts play a role in the existence of an a priori justification. Strictly speaking, we would still have a case of a priori justification, and no bar to A Priori Scrutability. At best we would have a case in which A Priori Scrutability has surprising grounds.

Second, Neta could suggest that the agent's ultimate priors are not justified at all (a priori or a posteriori), but that later credences based on them by conditionalization are justified a posteriori. On this view, conditionalization on empirical evidence E can take one from non-rational credences $cr(M|E)$ at t_1 to rational credences $cr(M)$ at t_2. A version of this view could lead to the rejection of Synchronic Frontloading. Nothing about the structure of Neta's case gives much reason to support this view, however. Furthermore, one can apply the usual argument from supposition. If reasoning of this sort is justified in the Knowledge Room, why isn't it also justified in the Supposition Room, suggesting that the credence $cr(M|E)$ is rational even independent of E?

Neta's objection to the informal frontloading principle concerns the hypothesis (W): 'I am not dreaming now'. He says that I can know W based on total empirical evidence E, but I am not justified a priori in believing that if E, then W. He suggests that we are at best a priori justified in believing 'If I am having a normal body of empirical evidence, then I am not dreaming' (G). But getting from G to 'If E then W' require the further knowledge (call it N) that E is a normal body of evidence, and that can be known only empirically.

A minor point: Neta says that closure under known logical entailment tells us that if W is justified then 'If E, then W' is justified. This seems right if the conditional is a material conditional. But the relevant case here is one of conditional belief, which is not belief in a material conditional, so that the closure principle does not seem to apply here. But given that both E and W are justified, it is plausible that 'If E then W' is justified all the same.

More importantly, it is not clear why justification for believing that if E, then W depends on justification for N. On the face of it, the reasoning from N is questionable, since N is consistent

with the hypothesis that one has always been dreaming (so that dreaming is normal). That is to to say that the bridging premise *G* is itself questionable. Perhaps Neta is thinking that *N* tends to undermine the hypothesis that one is *now* dreaming, although one has not normally been dreaming. Perhaps there is more of a case here, though it depends on what to say about evidence in common between dreaming and non-dreaming scenarios. Still, it is not at all clear that this reasoning from *G* and *N* is *required* for justification of 'If *E*, then *W*' in either case, or even that it is the best source of justification. On the face of it, there are other sources of justification, including abductive sources, that do not rest on *G* or *N*.

Furthermore, the suppositional argument for frontloading seems strong in this case. If *E* justifies *W*, presumably one could suspend judgment about *E*, suppose *E*, conclude *W*, and then discharge to yield a justified belief that if *E*, then *W* with justification independent of *E*. The same goes for if *E*, then *N*. All this tends to suggest that if *W* and *N* are justified, these conditionals can be justified independently of justification for *E*. Neta has to say that because the goodness of the reasoning depends on knowledge of *E*, the merely suppositional version of the reasoning is not justified. To say this is at least to bite a large bullet.

Neta suggests that frontloading principles get their intuitive support from cases of deductive reasoning and that the support does not generalize to cases of non-deductive support. He says that in nondeductive cases, our ability to infer *M* from *E* can depend on our justification for believing *E*. He suggests that Goodman's green/grue case is such a case. Our total evidence supports favoring the green hypothesis over the grue hypothesis, but this very support relation depends on the evidence. If one contemplated the cases prior to the evidence, we could not see why the green hypothesis is better than the grue hypothesis.

As before, it is not clear why suppositional reasoning does not apply here. What if our contemplator merely supposes that the total evidence in question obtains? On the face of it, even using that evidence in a suppositional role lead them to favor the green hypothesis (at least if their psychology is like ours), and on the face of it that reasoning will be justified (at least if ours is). And on the face of it, they could then discharge, yielding belief in the conditional. Neta must reject one of these steps—perhaps the claim that the reasoning is justified, or perhaps the discharging of the conditional—but intuition seems to remain clearly against him here. All this is to say that the intuitions seem strong in the inductive cases and not merely the deductive cases.

Even if one accepts Neta's points about induction versus deduction, it is worth noting that the key uses of frontloading principles in *CTW* all concern cases that are more akin to deduction than induction in important respects. Whereas Neta raises worries about cases concerning non-

extremal credences and knowledge that falls short of certainty, the key cases in *CTW* all involve extremal credences and knowledge that amounts to certainty. So even if there are worries about frontloading principles that apply distinctively in non-extremal or inductive cases, these will not undermine the applications of these principles in arguing for A Priori Scrutability.[2]

3   Reply to Schroeter

Schroeter offers two models of scrutability on which she suggests that A Priori Scrutability will fail: a token-reflexive model and a hermeneutic model. More precisely, she argues that Conditional Scrutability holds on these models, but that Conclusive Generalized A Priori Scrutability (which strengthens the scrutability claim to require certainty, apriority, and generality across all epistemic possibilities) fails.

Schroeter starts by outlining a "point-and-shoot" model that she takes to be friendly to the strengthened scrutability thesis. On this model, one can scry what counts as X in any scenario, independent of any empirical presuppositions about the actual world. I do not quite accept this model. I think there are often empirical presuppositions to acts of thinking such that if those presuppositions fail, relevant concepts will fail to refer. An example: for response-dependent concept, which picks out the cause of relevant mental responses, the existence of those responses may serve as an empirical presupposition without which the concept will not refer. It follows

---

[2]In correspondence, Neta says that he holds there is a kind of knowledge based on empirical evidence that does not start from prior credences (and perhaps that empirical knowledge is a necessary conditions for having credences at all). On this view, my argument from diachronic frontloading may be set aside on the grounds that the rational credence $cr^*(M|E)$ is not always defined at t_1 in the relevant cases. Still, I think that even if one accepts this view, the intuitive argument from synchronic frontloading principles remains strong. At time t_2, the subject could suspend judgment in the empirical evidence E and instead assume it as a supposition. It is not clear why they would not then be justified in (conditionally) concluding M in this version of the Supposition Room as they are in (unconditionally) concluding In correspondence, Neta says that he holds there is a kind of knowledge based on empirical evidence that does not start from prior credences (and perhaps that empirical knowledge is a necessary conditions for having credences at all). On this view, my argument from diachronic frontloading may be set aside on the grounds that the rational credence $cr^*(M|E)$ is not always defined at t_1 in the relevant cases. Still, I think that even if one accepts this view, the intuitive argument from synchronic frontloading principles remains strong. At time t_2, the subject could suspend judgment in the empirical evidence E and instead assume it as a supposition. It is not clear why they would not then be justified in (conditionally) concluding M in this version of the Supposition Room as they are in (unconditionally) concluding M in the Knowledge Room. All this tends to suggest that even if E plays an essential enabling role in having the relevant conditional credences, it does not play an essential justifying role.

that in scenarios where the presupposition is false, the concept will not refer. This leads to a qualification of Schroeter's Generality thesis, but not a harmful one. We already know that there are actual cases involving false presuppositions in which concepts lack referents and thoughts involving those concepts are indeterminate or false. We need only generalize this point across epistemic space: in scenarios where the presuppositions are false, a concept's intension will lack a referent and a thought's intension will be indeterminate or false.

On Schroeter's token-reflexive model, the processing of scrying the referent of a term 'X' in light of empirical information gives a central role to one's linguistic or mental tokens associated with 'X'. For example, to discover what counts as water in a world that contais both H_2O and XYZ, one traces the substance that bears an appropriate connection to one's actual beliefs about water. Schroeter suggests that on this model one will have to imagine those actual beliefs in any hypothetical scenario in order to identify the referent.

Now, I am somewhat skeptical that token-reflexivity plays as much of a role in the scrutability of reference (that is, in determining what counts as water, and so on) as Schroeter suggests. I think it plays a role in some hard cases, such as the Two Tubes cases involving demonstratives discussed in *CTW*, and perhaps related cases in which one is near-symmetrically related to H_2O and XYZ. In these cases a relation to a token may in effect break a tie between potential referents that are otherwise equally apt. It may also be important to trace relations to other people's tokens in cases of semantic deference. But it is far from clear that this sort of reasoning is needed in all cases. Even for natural-kind terms such as 'water', there are plausibly many scenarios for which token-reflexive reasoning is not required for scrutability. Still, if Schroeter succeeds in raising a problem for my scrutability theses even in a few cases, that will be a problem for my framework, so I will focus on these cases in what follows.[3]

_____

[3]In her important paper "Considering Empty Worlds as Actual" (2005), Schroeter argues that this role for token-reflexivity extends from hard cases involving 'water' to all cases, and uses this claim to argue that my intensions will be undefined in scenarios that do not contain relevant tokens. She suggests that this role is required in order for an astronaut visiting Twin Earth to count water as H_2O rather than XYZ, for example. But in cases like these, there are many paths to scrutability that do not go through tokens: for example, it suffices to identify water with the dominant watery stuff in one's historical environment. Schroeter raises the case of an amnesic astronaut who acquires a new term 'water$_2$' for XYZ and suggests that token-reflexivity is required to distinguish the referent of this term from that of the original term 'water$_1$'. This is not obvious: for example, water$_2$ (unlike water$_1$) may be identifiable as the dominant watery stuff in one's recent environment. In a sufficiently symmetrical scenario, a role for tokens may be essential. But the fact that tokens play an essential role in some scenarios does not entail that they play an essential role in all scenarios. For example, 'water' (in the mouth of a nondeferential user) plausibly picks out H_2O in a scenario in which H_2O is the only watery stuff in the historical vicinity of the center, even if words such as 'water' have never been used

Schroeter suggests that the token-reflexive model is inconsistent with Generality and Apriority. The problem for Generality arises because scrying 'water'-claims in a scenario presupposes that a relevant set of 'water'-thoughts exist at the center of that scenario. We have already seem that this sort of qualification to Generality is unproblematic. At worst it leads to intensions that are null at some scenarios, which we have good reason to expect in any case.

The problem for Apriority arises because scrutability will now depend on empirical knowledge of one's actual thoughts. Updating a sentence S (e.g. 'water is H_2O') in light of empirical evidence E will depend on empirical introspective knowledge of the mental states expressed by S, including perhaps the introspective knowledge *This is a thought that water is* H_2O of one's thought token *water is* H_2O. In these cases our conditional judgments *if E then S* and *if PQTI then S* are also justified by this introspective knowledge. Furthermore, this justificatory role for introspective knowledge is ineliminable, so these conditional judgments cannot be justified a priori.

An initial response it is that it is far from clear that introspective knowledge is so central in updating. On the face of it, an introspection-blind creature could do much of the world-involving reasoning that we do, could discover that water is H_2O, and so on. Still, Schroeter could suggest that when introspection is not playing this role, then perception is playing a similar role, as when one relies on knowledge of a perceived expression token, or on other perceptual demonstratives.

A deeper point is that even if the token-reflexive model is correct and introspective knowledge K plays a quite general role in scrying the truth-value of M, this does not entail that K plays an essential role in justifying scrutability conditionals. On the face of it, the frontloading argument still applies. If a subject can reason from $PQTI$ to M with background justification from K, then in principle she could suspend judgment about K and come to be justified in believing if $PQTI\&K$, then M. She need only suppose K, and then engage in the same reasoning that got them to if $PQTI$ then M from K, then discharge the supposition to yield if $PQTI\&K$, then M. This conditional belief will then be justified independently of K.

Schroeter says that this supposition is just imaginative role-playing. One is imagining a hypothetical person with knowledge K, and deriving a conclusion about the reference of their words. I respond that this is not how supposition works. In cases of supposition, we take K to be true and we reason just as if it were true. If I knew that my token *water* (this very token, picked out demonstratively while thinking a *water* thought) is causally connected to H_2O, and so on, then (on Schroeter's view) I would conclude that water is H_2O. Correspondingly, if I merely suppose that

───────────────────────

in that scenario.

9

my token *water* is causally connected to H_2O, I can likewise conclude (under the supposition) that water is H_2O. Although the antecedents in this reasoning involves the metalinguistic or metaconceptual mode, the conclusions do not. The conclusions in both case are claims using one's own concept *water*, are not claims mentioning someone else's concept.

The same applies equally to non-actual epistemically possible scenarios. For example, if I suppose that my token *water* (this very token, picked out demonstratively while thinking a *water*-thought) is causally connected to XYZ, and so on, then I can conclude (under that supposition) that water is XYZ. As before, the conclusion uses my concept water, rather than mentioning someone else's concept. And as before, there is no need for actual introspective knowledge to do justifying work here. Any evidential role for introspection can be played by frontloading introspective claims (such as *This very concept is a water-concept*) into the scenario. To align these claims with one's actual concept *water*, all one needs from the actual world is that concept and a demonstrative concept referring to it. No empirical knowledge is required.

So a token-reflexive model is quite consistent with a A Priori Scrutability. There is even a descriptivist analog of Schroeter's token-reflexive model. On an extreme version of that view, a concept such as *water* might be analyzed as *The cause of this concept-token* or perhaps *The cause of* water-*tokens*, involving demonstrative or higher-order concepts that pick out one's token *water*-concept. The actual scenario verifies *That concept-token is caused by* H_2O and so verifies *Water is* H_2O. Alternative scenarios might verify *That concept-token is caused by XYZ* and so verify *Water is XYZ*.

A scrutability-based token-reflexive model need not be as extreme as this, but it can still give a key role to demonstrative presentations of the concepts in question. If a given scenario verifies enough claims along the lines of *This concept-token is caused by XYZ*, it will thereby verify *Water is XYZ*. As discussed in section 6.16 of *CTW*, claims involving these demonstratives may even be built into the indexical part of a scenario specification such as *PQTI*, and there may be a priori links between claims involving a concept such as *water* and claims involving the corresponding demonstrative *That concept-token*.

Schroeter will presumably reject these versions of the token-reflexive model in favor of one where knowledge of the actual concept (as opposed to mere possession of it) plays a more essential role in scrutability. Here I simply note that (i) nothing in Schroeter's description of the Twin Earth cases that she uses to motivate her model supports her version over mine, (ii) frontloading reasoning strongly supports my version over hers, and (iii) Schroeter's argument against frontloading reasoning (assimilating it to meta-level judgments about hypothetical people) does not succeed.

Schroeter's second model is a hermeneutic model on which the referent of a term such as 'water' depends on one's practical/theoretical interests in using the word. Updating one's 'water'-beliefs depends on empirical knowledge about one's actual interests historically associated with 'water'-thoughts. The same goes for scrying 'water'-thoughts from information such as PQTI. The justificatory role of knowledge of interests is ineliminable. If so, scrutability conditionals cannot be justified a priori.

I am not sure that I fully grasp of the hermeneutic model works and of the role that historical knowledge plays. I am inclined to agree that the interests associated with a concept often play an important role in judgments about reference. For a start, in cases of semantic deference, one needs to know what others refer to with a term in order to know the referent of one's own term, and knowledge of their interests in using the term may play a key role here. There may also be cases in which one defers to the usage of one's past self. But I there are plausibly cases where a term is used without deference, evem to one's past self. In these cases, it is hard to see a constitutive role for one's historical interests in using the term.

There are certainly some connections between one's present interests in using a term and its referent. For example, two subjects with different interests associated with a word such as 'planet' may well judge their terms to have different extensions. But in this case (at least setting aside semantic deference) it is natural to take them to have different concepts with different extensions, for which different scrutability conditionals will hold. That is to say that in these cases, interests seem to play an enabling role rather than an evidential role: they enable possession of relevant concepts rather than providing evidence for judgments involving those concepts. But a mere enabling role poses no problem for a priori scrutability.[4]

I am not sure whether there are ordinary cases in which one's own historical interests play an evidential role in judgments about reference (setting aside cases such as 'My interests'). But insofar as there are such cases, we can once again apply frontloading reasoning. If evidence E about interests (say *I use planet-concepts for scientific purposes*) enables justified belief in S (say *Pluto is a planet*), then prima facie, supposition about interests can yield justified conditional belief that if E, then S, justified independently of E. So knowledge of one's actual interests will

---

[4]The enabling/evidential distinction is also useful in addressing Schroeter's (2008) discussion of the concept *I*. Schroeter suggests that identifying the referent of this concept depends on knowledge of underlying cognitive mechanisms and the like. On the face of it, these mechanisms clearly play a role in enabling possession of the concept, but evidentially, knowledge of them is inessential to the subject's applying the concept. Of course they might play an evidential role for an external observer interpreting someone else's use of 'I', but this is not the relevant case here.

play no essential role in justifying scrutability conditionals.

The same goes for hypothetical scenarios. One would likewise be justified in believing *If I use planet-concepts for non-scientific purposes, then Pluto is not a planet*. I think it is implausible that conditionals like this are really justified, but this is only because it is implausible that these interests play an evidential role. If they did, then the conditionals would be fine. So even an evidential role for interests does not give them an essential role in scrutability.

Schroeter also suggests that on the hermeneutic model, intensions will not be interesting Fregean senses or narrow contents. She says that on this model, tracing the referent of any word (e.g. 'water') with respect to a scenario will require tracking the historical patterns of use of that expression in a scenario, and this will in effect work by a non-topic-specific metasemantic route for connecting a term to it's referents. That is, in effect the primary intension of 'water' will be the same as the primary intension of "the referent of 'water'" (where 'water' is construed as an orthographic string), which will not be interestingly different from the primary intension of "the referent of 'plus'". The primary intensions of all terms will in effect simply reflect the primary intensions of 'referent'.

This view may be a consequence of an extreme version of the hermeneutic model, but it just brings out how implausible that version is. On this model, our primary grip on what counts as addition is via our grip on what counts as the referent of 'plus'. We will be unable to say that it is a priori that 2+2=4, because of the epistemic possibility that 'plus' is used to refer to subtraction. We will be unable to say it is a priori that $P\&Q \rightarrow P$, because of the epistemic possibility that 'and' is used to refer to disjunction. Something like this may be plausible for deferential users of these expressions, for whom the primary grip is via other people's uses, but it is extremely implausible that this model applies universally. For example, it is plausible that an expert mathematician has justification for believing certain facts about addition that could not be defeated simply by evidence about her use of 'plus' or usage of 'plus' in her community.

There is also an obvious regress worry: how do we identify the referent of 'referent'? To avoid a vicious regress, we need to assume some sort of substantive grip on what counts as reference. But once we have done that, we can also presumably assume a substantive grip on what counts as conjunction, addition, water, and so on. These substantive grips will ground routes to scrutability that are not wholly grounded in metasemantics and that yield substantive primary intensions.

Perhaps a Quinean should say that our route to identifying referents is holistic and that metasemantic considerations will play some role along with non-metasemantic considerations. My view is that there are numerous cases (including that of the mathematician above) in which metaseman-

tic considerations play no essential role. But in any case, if metasemantic considerations play only a partial role in scrutability as one factor among many, then there is no danger that intensions will be trivialized as Schroeter suggests.

4   Reply to Stanley

Stalney presses an important worry for the semantic program of *CTW*. The worry is that because my intensions are defined in terms of an idealized notion of apriority, they are too coarse-grained to serve as Fregean senses. There are many cases in which '*A* = *B*' is a priori but still cognitively significant: '67+56=123', for example. In these cases, *A* and *B* will have the same primary intensions, but by Frege's criterion of individuation, they will have different senses. So primary intensions are not Fregean senses.

I discuss this issue in *CTW* (and also in Chalmers 2006), suggesting a multipronged response. First, moving from unstructured primary intensions to structured primary intensions deals with most of the problem cases. Second, the only residual problem cases are cognitively significant a priori identities in which *A* and *B* are both simple expressions without structured intensions, but it is not obvious that there are such cases. Third, if there are such cases, one might be able to handle the issue by moving to a finer-grained space of non-ideal scenarios and nonideal intensions defined over them. Fourth, if the prior strategies do not succeed, one may have to accept that my intensions are not as fine-grained as Fregean senses, but a semantic pluralist can allow that they can still play many useful explanatory roles.

Stanley agrees with some of this, but he thinks that there are clearly residual problem cases involving simple expressions, and he thinks that the appeal to semantic pluralism is unsatisfactory. Stanley and I do not have huge disagreements about the properties of epistemic intensions, but we disagree on how valuable this makes them in semantic theorizing. This disagreement may well stem from a deeper disagreement about the value of primitive Fregean senses in semantic theorizing.

Stanley puts his point initially by saying that my theory of meaning is not properly regarded as Fregean, because it is not a plausible theory of cognitive significance. I think this is largely a verbal point. I regard my theory as broadly Fregean, in the sense that it gives a key role to epistemically individuated modes of presentation, so that coextensive terms can have different semantic values. But it is certainly not narrowly Fregean in that it holds to all of Frege's commitments about meaning. I take it that this broad use of 'Fregean', analogous to broad uses of 'Russellian',

13

'Millian', 'Cartesian', and 'Humean', is a familiar one. It is also worth noting that I am not offering a theory of cognitive significance (or even of apriority). I am offering a theory of (a sort of) meaning or content. The real worry is that those meanings are not well-enough connected to cognitive significance.

Another preliminary point: Stanley stresses that heavy idealizations will be required in order that mathematical truths be scrutable from other truths. I acknowledge that heavy idealizations are required here, though I do not see strong reasons to think that such idealizations cannot succeed. It is worth noting that if some mathematical truths are inscrutable and have to be built into the base, this will not cause obvious problems for semantic applications of the framework (although it might cause problems for some metaphysical applications). In any case, for Stanley's main worry, the heaviness of the idealizations play no essential role. His worry arises just as strongly for truths knowable a priori by quite ordinary reasoners, such as the mathematical case above.

On the first and second prongs of response above: Stanley acknowledges that the move to structured intensions helps with some cases, but he suggests that there are residual cases involving simple expressions that raise the same problems. His key case involves Schrödinger's mathematical proof of the equivalence of matrix mechanics' and wave mechanics, which could be put as a proof that two primitive operators are identical: '$\phi = \psi$', say. In these cases he says we cannot appeal to structural complexity.

There is an obvious response in this case. It is true that 'ϕ' and 'ψ' are both simple expressions, but this does not mean that their meanings lack structural complexity. After all, both expressions are defined via an explicit mathematical definition (without such a definition there would be no mathematical proof of their equivalence). The definition can naturally be seen as giving the meaning of these expressions, and it certainly involves complex structure. So (as with the analogous case discussed on p. 249 of *CTW*, where 'Num' is defined to be 77×33) it is easy to associate these expressions with distinct structured intensions. To be sure, on some semantic approaches simple expressions are always associated with unstructured intensions, so this suggestion requires going beyond those approaches. But this is not a radical step. Given the plausible claim that simple expressions are cognitively equivalent to their definitions, it is reasonable enough for a Fregean to associate them with the same semantic value.[5]

---

[5] As Stanley notes, Frege holds in the *Grundlagen* (1884) that there are "fruitful definitions" that "extend our knowledge" (although they are analytic), suggesting that they are cognitively significant. I do not know whether the current case would count as one of these. In any case, it is certainly not biting a large bullet for a broad Fregean to hold that definitions are cognitively insignificant. Frege himself appears to change his mind later, saying in 1903 "No definition

A compelling counterexample to the structured intension approach will require a cognitively significant a priori identity '$A = B$' in which both sides are simple expressions that lack complex definitions. Mathematically provable statements are unlikely to provide examples here, since such proofs require definitions to get off the ground. Furthermore, Kripkean cases in which names are not a priori equivalent to descriptions are unlikely to help here, since Kripkean reasoning tends to suggest that identities involving these names are not a priori. Stanley offers two philosophical examples: 'knowledge is evidence' and 'justice is fairness'. But of course it is extremely controversial whether these are true identities, let alone a priori identities.

Still, perhaps there can be cases here. One source might involve informally used mathematical terms such as 'π'. Although π can be given many definitions (e.g. the ratio of a Euclidean circle to its diameter), it is arguably not cognitively equivalent to any of them in informal mathematical usage. Instead it is grounded in a large cluster of definitions and applications. One could imagine a group of mathematicians with one informal term π_1 grounded in a various geometric roles and another informal term π_2 grounded in various roles from analysis. These mathematicians could come to discover through a priori reasoning that $\pi_1 = \pi_2$. This would be a cognitively significant a priori identity, and it is arguable that neither side has a structured intension. One could argue that there is a cluster-style definition for each, but it is at least far from obvious than any such definition will be cognitively equivalent to the original expression. One might also make a case for cognitively significant a priori philosophical identities in a similar style.

Another potential source (suggested in an unpublished paper by Nick Kroll some years ago) involves a priori empty names. For example, perhaps it is a priori that Sherlock Holmes and Harry Potter do not exist. If so, it seems that both names, although cognitively distinct, will have identical primary intensions, empty at all scenarios. One could suggest in response that the names pick out distinct abstract objects and have distinct intensions, but there is at least a prima facie case here.

These examples can be resisted. Nevertheless, I do not see a compelling principled reason why there could not be a priori cognitively significant identities without structure. As Stanley notes, a Russellian approach on which all expressions are equivalent to structures consisting of super-rigid expressions (for which reference in grounded in Russellian acquaintance that precludes Frege cases) might provide a principled reason. But I am inclined to reject this thesis for reasons discussed in chapter 1 of *CTW*: for many expressions, there may be no a priori equivalent complex expressions in these more basic terms. I remain sympathetic with the idea that there is some sort of

---

extends our knowledge. It is only a means of collecting a content into a brief word or sign, thereby making it easier for us to handle".

priority of Russellian acquaintance in an account of intentionality, and it is not impossible that we could use this to analyze Fregean senses in the problem cases in some related way: for example, by inferential roles with respect to basic acquaintance concepts. But this would require going beyond the framework of structured intensions.

In the absence of a principled reason to deny it, I am inclined to think that there may well be cognitively significant apriori identities of the problematic sort. If so, it would be nice to see clearer examples. Certainly, it seems that such cases are not especially common, and one could argue that the cases are somewhat marginal. I return shortly to the question of whether this marginality reduces the significance of the problem.

Even if there were no counterexamples here, Stanley is right that the framework would lead to a non-uniform treatment of cognitively significant identities. Significant scientific identities will be analyzed in terms of scenarios where they are false, while significant mathematical identities will involve differences in structure without falsifying scenarios. Intuitively there is more in common between these domains than that. On the face of it, in both mathematical and scientific discoveries one rules out scenarios that were previously epistemically possible.

This thought can be used to motivate the third line of response to the problem (which Stanley does not discuss): introducing fine-grained epistemically possible scenarios. Here the idea is that there will be scenarios at which cognitively significant mathematical theorems (such as Fermat's last theorem) are false, even though those theorems can be known a priori. A natural suggestion is to rework the construction of scenarios with cognitive insignificance (or perhaps analyticity) playing the role of apriority. Then in the cases of significant a priori identities 'A = B', these identities will be false at some fine-grained scenarios, and A and B will have distinct fine-grained intensions. If this approach could be made to work, fine-grained intensions (structured or unstructured) could serve as Fregean senses and we would have a more uniform treatment of scientific and mathematical cases.

Unfortunately, no convincing model of fine-grained intensions with the right properties has yet been devised. Jens Christan Bjerring (2013) has shown that there are principled problems for any such approach. Either all logical truths will be true at every scenario (even though many are cognitively significant), or some logical near-trivialities will be false at some scenarios (even though they are arguably cognitively insignificant). Either way it will be hard to get a tight match between cognitive significance and fine-grained intensions. Perhaps there is room to maneuver here, for example exempting logic as a special case and holding all logical truths to be cognitively insignificant, or holding that logical trivialities are significant. And one could still apply the appa-

ratus to other a priori domains: for example, cognitively significant connections between natural and normative truths can be handled by Gibbard-style fine-grained factual-normative scenarios. But it is not easy to make this sort of model work in full generality.

Let us now suppose that none of these replies provide a sufficient response. In particular, let us suppose that there are cognitively significant a priori identities that cannot be handled by finding structure or by invoking fine-grained scenarios. Then we will have to allow that structured epistemic intensions carve somewhat less finely than traditional Fregean senses, in that some pairs of cognitively distinct expressions may share the same intension.

How bad is this? Stanley's view is that it is very bad: it follows that structured intensions are simply not acceptable Fregean semantic values. My view is that it is not so bad. Structured intensions are at least quite close to having the properties of Fregean senses, falling short only in some marginal cases. It is true that if one wants true Fregean senses that individuate by cognitive significance, one will need to go beyond structured intensions. But structured intensions seem at least to provide a good first or second approximation to what we need. After all, we do not seem to have any good positive account of Fregean senses. In the absence of such an account, structured intensions can do at least a lot of their explanatory work.

Furthermore, structured intensions may provide significance guidance in coming up with a better account. It may be that Fregean senses end up being somewhat more fine-grained entities that share many of the properties of these structured intensions: their structure, their relations to referents and truth-value, their evaluability at epistemic and metaphysical possibilities. In particular, if the generalized scrutability thesis is correct, it would be surprising if Fregean senses were not evaluable at epistemically possible scenarios. A natural way to develop a finer-grained account would be to associate simple expressions with finer-grained entities (inferential roles, perhaps) that lie behind and determine their epistemic intensions. Then many of the features of structured epistemic intensions will remain in a successor framework.

Much here depends on whether one thinks we already have a good account of Fregean senses. Stanley seems to think we do: an account on which senses are taken as primitive, and not constructed from anything prior to them. Stanley acknowledges that these accounts have a problem of obscurantism, but he does not think the problem is fatal. I think that the problem is fatal: most extant accounts on which senses are taken as primitive are no accounts at all. Rather, these accounts are akin to job descriptions that do not make the case that anything can do the job and that do not properly characterize the things that do the job.

I am not opposed to postulating primitive entities in general. But one has to say enough about

17

the entities one is postulating to give them clear and definite properties. In the case of primitive Fregean senses, this has usually not been done. Frege himself notoriously did little to give a clear positive characterization of senses. Likewise, no such account is given by Stanley (here or elsewhere, as far as I know).

I am not saying that the only good account of senses is a constructive account. One could in principle give an axiomatic account (e.g. Zalta 2001) laying out axioms that senses must satisfy, or an algebraic account (e.g. Bealer 1993) laying out the operations that can be performed on senses and the sorts of results that are obtained. Indeed, it would be easy enough to phrase the current account as an algebraic account (there is at least the beginning of such an account in "The Nature of Epistemic Space"), characterizing senses as entities that compose in certain ways, that can be evaluated at scenarios in certain ways (where scenarios are themselves algebraically characterized), and so on. With enough detail, an account of this sort may constitute an adequate account of senses even if (like Bealer's) it does not build in a construction. Still, the detail provided by these accounts usually makes it easy to find a construction that satisfies the account (for Bealer, this might be a structure whose elements are properties or linguistic entities; for Zalta, it might be a set of properties). So these accounts are relevantly similar to constructive accounts in providing a detailed formal account of senses. Unsurprisingly I think that my account is superior to these in some respects, but I will not get into the comparison here.

Constructions can still play two distinctive roles, combined either with axiomatic/algebraic accounts of senses or with informal characterizations of senses. First, they help demonstrate satisfiability. Just as Dedekind's construction of the real numbers shows that something could satisfy the axiomatic characterization of the reals, a construction of Fregean senses shows that something could satisfy an axiomatic characterization of senses, or an informal characterization of senses. Second, they help demonstrate determinacy. Some formal or informal accounts underdetermine many crucial formal properties of senses, whereas a specific construction that satisfies the account is determinate.

Both satisfiability and determinacy are real issues in the case of Fregean senses. Regarding the first: many opponents have argued that nothing can have various combinations of properties that senses are supposed to have. A construction can help to demonstrate that something can in fact have those properties, or something close enough. Regarding the second: most informal characterizations of senses leave them highly underdetermined. We are told that sense determines reference, that senses are individuated by cognitive significance, and that certain compositional principles hold, perhaps along with a few other informal remarks. That these remarks leave senses

18

undetermined is brought out by the existence of many different competing accounts of Fregean senses. A construction (or a sufficiently detailed axiomatic or algebraic account) fills in the holes and provides a determinate conception of senses.

Stanley approvingly quotes Burge saying that mathematical entities are explanatorily useful even without deciding their metaphysical nature. This seems correct, but it is not really to the point. The relevant issue here is not a *metaphysical* characterization of Fregean senses (in the quoted passage Burge is discussing Platonism and other views about the ontology of abstract objects) but an adequate *formal* characterization. We obviously have a well-developed formal characterization of mathematical objects, at least in axiomatic and algebraic terms, typically with associated constructions as well. If we had a formal characterization of senses along these lines, they would then have a status analogous to that of mathematical objects, and we could be relaxed about the metaphysics. But Stanley's alternative to a constructive account seems to be no formal account at all.

It is worth noting that in the mathematical case, we have well-developed pure mathematics and also well-developed applied mathematics. We understand the mathematical operations that can be performed on mathematical entities, and we also understand the ways that they can be used to characterize concrete systems (for example, the conditions under which a physical system has a certain gauge structure). A well-developed account of sense requires both components. We need both an account of the formal structure of senses and an account of just how they can be used to characterize concrete systems (for example, the conditions under which a person grasps a certain sense).

Frege's account of sense appears to be highly underdetermined on both the pure and applied dimensions. On the pure side, the formal structure of basic senses and the way they determine referents is left largely unspecified. On the applied side, Frege gives in effect a condition (in terms of cognitive significance) for intrapersonal sameness of sense, but this does not easily yield a useful condition for interpersonal sameness of sense or for a person's grasping a specific sense. By contrast, the account of senses as structured epistemic intensions provides a relatively determinate account on both the pure dimension (in terms of the mathematical structure of these intensions) and applied dimensions (by giving conditions for grasping intensions in terms of epistemic status and inferential role). Perhaps the account is suboptimal, but at least it is in the game.

Note that one can understand senses in terms of intensions while staying neutral on sense's metaphysical natures, as one does in mathematical cases. Strictly speaking one need not even identify senses with intensions. In mathematics, after we use Dedekind cuts to construct real num-

bers, one can kick away the construction in favor of an algebraic characterization. One could do the same here, saying senses are entities (perhaps even primitive entities) that satisfy a relevant algebraic characterization of structured intensions and that can be put into one-to-one correspondence with structured intensions without being identified with them. That will have the bonus of avoiding Benacerraf-style problems about choosing between different but isomorphic systems of entities that satisfy the characterization. But as in mathematics, it will often be useful to wheel out a construction, perhaps for modeling or for concreteness. And certainly a formal characterization cannot be dispensed with.

Structured epistemic intensions look better in this light. One could now make a case that they are the best account we have so far of Fregean senses. They do not have every property that senses are supposed to have, but they have many of them. Perhaps we will be able to find a better account in the future, but here structured intensions may be able to provide a guide. It is not out of the question, for exaple, that they could guide us to an algebraic account of senses for basic expressions, where these are individuated in part by their values at scenarios and in part by something else, such as inferential roles.

Just how one views the status of structured intensions in light of this will depend on one's semantic methodology. A semantic monist holds that there is one true sort of meaning, and one true meaning relation that relates expressions (or utterances, or expressions in context) with their meanings. A Fregean semantic monist may well hold that meanings are finer-grained Fregean senses, in which case structured intensions are not meanings. On this view, when my semantic framework is considered as a theory of meaning it is false. Still, this falsity is consistent with its being the best Fregean theory currently available, and one that may provide an approximation to and guidance toward a true theory. That is not a bad status to have. I take it that semantic monism is also compatible with taking structured intensions to be quasi-meanings that are not strictly speaking meanings but nevertheless may help in analyzing many semantic phenomena. That is also not a bad status to have.

I am a semantic pluralist, holding that there are many sorts of meanings (and many meaning relations) and that none of them is privileged as the one true sort of meaning. At root, there are any number of ways of associating expressions and utterances with abstract (or concrete) entities that can play some of the explanatory roles that we want meanings to play. No single entity will play all of these roles perfectly. But if an entity plays many of these roles reasonably well, it will be a sort of meaning. According to semantic pluralism, instead of asking "What is the meaning?", we ask "Which semantic values have which properties, and what can they explain?". Structured

intensions play many of the meaning roles quite well. They can already be used for many purposes. It may turn out that there are meanings (such as Fregean senses) that play the relevant role even better. But insofar as these meanings are undiscovered, they are no bar to the claim that structured intensions are the most useful meanings currently available.

Stanley raises some doubts about my appeal to semantic pluralism. He thinks that pluralism is acceptable in the case of character and content but not here. His first reason is that on my pluralist view, the objects of knowledge and the objects of a priori knowledge (structured intensions) are distinct: the former are Fregean senses and the latter are structured intensions. I would not put things this way. Assuming that there are fine-grained Fregean senses, then I would say that we can stand in epistemic relations of both sorts to both senses and intensions. It could turn out that relations to senses are more fundamental, but in the meantime analyzing relations to intensions is useful. Insofar as intensions are less fine-grained than senses, they will be less discriminating as objects both of empirical and a priori knowledge, and they will provide an imperfect semantics for attitude ascriptions of both sorts. But they will at least provide a good approximation in both cases, one that can usefully deployed to help analyze knowledge of both sorts. It may be that if we had the finer-grained account, it would supersede the intensional account, but we do not currently have that account.

Stanley also suggests that we need Fregean senses to construct or define epistemically possible scenarios, so that senses will have conceptual priority over epistemic intensions. I do not think this is obvious. I argue on pp. 239-41 of *CTW* that if the thesis of Super-Rigid Scrutability is true (as I think it is), we can define scenarios as complexes of properties (where these properties are those picked out by expressions in a super-rigid scrutability base). If so then we can understand scenarios without invoking senses. If Super-Rigid Scrutabilility is false, then to move beyond a linguistic account of scenarios we may need to construct them from senses (*CTW*, p. 242). But the input from senses will be limited: we can construct scenarios using only the senses of expressions in a scrutability base, and intensions will be well-defined from there. Perhaps this suggests some priority for senses, but it also suggests a useful explanatory role for intensions. They allow us to move from semantic values for a few primitive expressions to semantic values for all expressions. It may well be that we can understand Fregean senses for certain primitive expressions well before we understand them for all expressions, or that certain standard objections to Fregean senses (for proper names, say) do not apply to the relevant primitive expressions. In either case, the intensional framework will give significant explanatory purchase. And even before we have a final account of Fregean senses, the intensional account can shed considerable light on

their structure and properties.

References

Bealer, G. 1993. A solution to Frege's puzzle. *Philosophical Perspectives* 7:17-60.

Bjerring, J.C. 2013. Impossible worlds and logical omniscience: An impossibility result. *Synthese* 190:2505-2524.

Chalmers, D.J. 2006. Two-dimensional semantics. In (E. Lepore and B. Smith, eds) *Oxford Handbook of the Philosophy of Language*. Oxford University Press.

Chalmers, D.J. 2012. *Constructing the World*. Oxford University Press.

Frege, G. 1884. *Die Grundlagen der Arithmetik*. W. Koebner. Translated as *The Foundations of Arithmetic* (J.L. Austin). Northwestern University Press, 1978.

Frege, G. 1903. Über die Grundlagen der Geometrie. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 12: 319-24, 368-75. Translated in (B.F. McGuiness, ed.) *Collected Papers on Mathematics, Logic, and Philosophy* (Blackwell, 1984), pp.273-84.

Neta, R. 2014. Chalmers' frontloading argument for A Priori Scrutability. *Analysis*, this issue.

Schroeter, L. 2005. Considering empty worlds as actual. *Australasian Journal of Philosophy* 83:331-347.

Schroeter, L. 2014. Scrutability and epistemic updating. *Analysis*, this issue.

Stanley, J. 2014. Constructing meanings. *Analysis*, this issue.

Zalta, E.N. 2001. Fregean senses, modes of presentation, and concepts. *Philosophical Perspectives* 15: 335-59.