Charity and Analyticity

Candidate Number: 1008185

Word count: 29,639

Abstract

Chapter One attempts to develop one component of Williamson (2007)'s defence of a knowledge-maximisation principle of charity. In particular, Williamson speculates that, roughly, a causal connection is a channel for *reference* to an object if and only if it is a channel for the acquisition of *knowledge* about that object. After first arguing for this claim the problem of deviant causal chains, I defend it against a counterexample due to McGlynn (2012). Finally, I relate the preceding discussion to the idea that 'belief aims at knowledge' (Williamson 2000).

Chapter Two takes up the question of why, on the hypothesis that meaning is socially determined, it would follow that there are no analytic sentences, for it seems possible in principle that linguistic practices impose conditions on membership. Williamson (2007) suggests the answer that a language which encoded analyticities would be *dysfunctional*. I attempt to develop this thought more precisely, and suggest that such practices are dysfunctional primarily for excluding good reasoners, rather than those with reasoned disagreements.

Chapter Three argues against fictionalism, understood as the view that sentences which apparently refer to numbers makes knowledgeable assertions, even though their semantic content itself is not known. In particular, I argue that fictionalism's distinction between semantic and assertive content is theoretically redundant, independently of the plausibility of mathematical scepticism itself. If this is correct, it may suggest that the logical space of responses to what Williamson (2007) calls *judgement scepticism* is more restricted than it seems.

Contents

1 Knowledge, Reference and Causal Connections 4

2 Analyticity and the Social Determination of Meaning 38

3 Fictionalism as Judgement Scepticism 69

1 Knowledge, Reference and Causal Connections

1.1 Introduction

There is a crude line between semantics and metasemantics.[1] Roughly, on the one side are facts about what the constituents of thought and language mean, while on the other side are facts about why the constituents of thought and language have the meanings that they actually have. For instance, it is a semantic fact that 'and' means *and*, but a metasemantic fact that English has a term that means *and* in part because it is extremely useful for public languages to have terms that express conjunction. Why 'and' in particular means *and* is a further question. On one familiar story, 'and' means *and* in part because the correct assignment of reference to languages is constrained by a truth-maximisation principle of charity: if all is equal, then an assignment of reference to a language is correct only if it maximises true belief in that language. Given that an assignment of reference to 'and' on which 'and' means *and* presumably outperforms most alternative assignments of reference to 'and' on the score of truth-maximisation, the truth-maximisation principle of charity defeasibly supports an assignment of reference to 'and' on which 'and' means *and*.[2]

[1] I use 'metasemantics' in the sense of Kaplan (1989a).

[2] For further development, see Lewis (1983) and (1984) and Williams (2007). Given that, in light of Putnam's paradox and Kripkenstein's rule-following paradox, a truth-maximisation principle of charity drastically underdetermines reference, Lewis notoriously supplements his charity principle with an eligibility constraint: if all is equal, then the objects of reference should be relatively natural. For detailed discussion of naturalness, see Dorr & Hawthorne (2013).

It is another semantic fact that 'water' means *water*. Another application of the truth-maximisation principle of charity suggests that 'water' means *water* in part because an assignment of reference to 'water' on which 'water' means *water* presumably outperforms most alternative assignments of reference to 'and' on the score of truth-maximisation. But another familiar story, according to which the meaning of terms is constrained by speakers' causal connections to their environment, seems at least as illuminating.[3] Notoriously, one can imagine another world, qualitatively just like this one, which contains a water-like substance, twater, with a quite different (but perhaps just as natural) molecular composition. Perhaps with some finessing, it seems possible that, before it was discovered that water was constituted by H_2O molecules, an assignment of reference to 'water' on which 'water' meant *twater* would have maximised true belief to just the same extent as one on which 'water' meant *water*. Plausibly, the factor that explains why 'water' has nevertheless always meant *water*, not *twater*, is that our causal environment contains water, not twater.[4]

In recent work, Timothy Williamson (2007) has defended a *knowledge*-maximisation principle of charity: if all is equal, then an assignment of reference to a language is correct only if it maximises *knowledge* in that language. According to Williamson, a picture of reference-determination grounded in the knowledge-

[3] So-called 'causal' theories of reference are typically claimed to originate in Putnam (1975) and Kripke (1980). Incidentally, Kripke himself rejects the name 'causal theory of reference' (preferring 'historical web picture'), noting that he does not generally take causal connections to be required for reference, given reference to mathematical objects (Kripke 2019).

[4] For more careful development of a 'mirror universe'-style objection to the truth-maximisation principle of charity, see Hawthorne (2007).

maximisation principle of charity promises to explain both the appeal of the truth-maximisation principle and the appeal of causal theories of reference. First, since knowledge entails true belief, that the assignment of reference is constrained by a knowledge-maximisation principle of charity explains why the assignment of reference also tends to maximise true belief. Second, Williamson suggests that the knowledge-maximisation principle of charity explains why causal connections are often channels for reference. Namely, causal connections are often channels for *reference* because they are often channels for the acquisition of *knowledge*. Thus, Williamson speculates that the following qualified biconditional holds:

> **Hypothesis:** Roughly: a causal connection to an object (property, relation, …) is a channel for reference to it if and only if it is a channel for the acquisition of knowledge about the object (property, relation, …) (2007: 264; my label).

This chapter primarily aims to develop and defend **Hypothesis**. It proceeds as follows. In §1.2, I offer some general reasons that one might expect **Hypothesis** to hold, given a knowledge-first picture of mind. In particular, Williamson briefly notes that causal theories of reference are vulnerable to the problem of deviant causal chains; I attempt to bring out the severity of this problem if one denies that causal connections are often channels for reference in virtue of often being channels for knowledge. In §1.3, I argue that a counterexample to an unqualified version of **Hypothesis**, due to Aidan McGlynn (2012), fails, partly in light the

considerations raised in the preceding section. Indeed, even if it is conceded that McGlynn's case does falsify the claim that causal connections are channels for reference *only* if they are channels for knowledge, a defender of **Hypothesis** itself can arguably repurpose the case as evidence that our grasp of which causal connections are channels for reference is derivative on our grasp of which causal connections are channels for knowledge. In §1.4, I situate the preceding discussion with respect to Williamson's well-known slogan 'belief aims at knowledge'; and briefly and tentatively discuss the implications of the idea that belief aims at knowledge for the knowledge-maximisation principle itself.

1.2 Knowledge and Deviant Causal Chains

In this and the following section, I primarily discuss **Hypothesis** on its own terms, independently of its role in Williamson's broader defence of the knowledge-maximisation principle of charity. Before proceeding to a defence of the principle, it will be helpful to briefly clarify how it will be understood. First, it will be assumed that a causal connection between a subject and an object is a *channel* for the subject to refer to that object if and only if the subject is in a position to refer to the object in virtue of the relevant causal connection; given that a channel may be open without any message being sent through that channel, the subject may not in fact come to refer to the object in virtue of the relevant causal connection. Second, it will be assumed that a causal connection between a subject and an object is a channel for the subject to acquire knowledge about that object if and only if the subject is in a position to know that the object is something; that

is, the subject need only be in a position to know that the object exists, not necessarily what kind of object it is.[5] Third, the left-to-right direction of **Hypothesis**, according to which a causal connection to an object is a channel for reference to that object *if* it is a channel for the acquisition of knowledge about it, will be taken as uncontroversial; if one is in a position to have knowledge about something, then one is in a position to refer to it. Thus, at issue is the plausibility of the following principle, perhaps taken in some suitably qualified form:

(1) A subject S is in a position to refer to an object O in virtue of some causal connection that obtains between S and O only if S is in a position to know that O exists in virtue of that causal connection.

(1) should be read as universally quantified and nomologically necessitated (the same qualifications apply to analogous principles below). Note that, at least in principle, (1) neither seems to imply nor be implied by (2):

(2) A subject S refers to an object O in virtue of some causal connection that obtains between S and O only if S knows that O exists in virtue of that causal connection.[6]

[5] Of course, some theorists might maintain that, at least for some entities, one is only in a position to refer to those entities if one knows some special fact about those entities—in effect, one must perceive them under the right guise. However, this is consistent with the claim that knowing that an object exists is sufficient for referring to it; it just implies that one must know *more* than that it exists in order to know that it exists. Chapter Two, especially §2.5, argues that the fact that demanding acquaintance conditions are a fortiori demanding conditions on having any knowledge at all is a compelling reason to reject the existence of the former.

[6] For present purposes, we harmlessly assume that, if one knows anything about an object O, then one knows that O exists. Technically, this conditional may fail to hold, given that one has not yet inferred that O exists from whatever it is that one does know about O. Much of the discussion in

A counterexample to (1) would take the form of a situation in which one is in a position to refer to an object but not in a position to have knowledge concerning it; given that one might not actually refer to the object, such a case need not constitute a counterexample to (1). Conversely, a counterexample to (2) might take the form of a situation in which one forms a false belief about an object without forming any true beliefs about that object; given that one might be in a position to know that the object exists in virtue of one's causal connection to that object, such a case need not constitute a counterexample to (1). Incidentally, as we shall see in §1.3, it is possible in principle that a case is a counterexample to both (1) and (2) without being a counterexample to the knowledge-maximisation principle of charity itself, on the assumption that an assignment of reference can be knowledge-maximising without being *uniquely* knowledge-maximising. For then, if a subject fails to know on *any* candidate assignment of reference, every candidate assignment of reference is knowledge-maximising, so the knowledge-maximisation principle is trivially satisfied.

In the following section, I turn to a putative counterexample, due to McGlynn (2012), to the knowledge-maximisation principle of charity—although, in light of the preceding point, I shall suggest that it is more naturally construed as a counterexample to (1). For now, however, the important question is why one might believe **Hypothesis**, besides on the contentious grounds suggested by Williamson that the success of causal theories of reference is best explained by

---

this chapter could be employed in defence of (2) instead, although it is less plausible that (2) holds in full generality than that (1) does.

the knowledge-maximisation principle of charity. In fact, Williamson himself offers what is in effect a second reason to believe **Hypothesis**, namely that causal theories of reference, like causal theories of knowledge, are faced by the problem of deviant causal chains (2007: 264). Thus, explaining which causal connections are channels for reference in terms of which causal connections are channels for the acquisition of knowledge has the effect of reducing the problem of deviant causal chains for causal theories of reference to a special instance of the problem of deviant causal chains for causal theories of knowledge. Of course, assimilating the former problem to the latter problem does not *solve* the former, since it does not solve the latter. Nevertheless, there is a clear benefit to the assimilation and a clear cost to rejecting it, especially given Williamson's own knowledge-first ideology. The rest of this section attempts to make each explicit.

The deviant causal chains problem for causal theories of reference is perhaps best introduced by pairs of cases like the following:

> **Hologram:** Jones beams a hologram of himself several light-years away to Smith, using highly reliable holographic technology. Although Smith has never before met or seen Jones, upon receiving the hologram, which is a photo-realistic three-dimensional image of Jones, Smith can immediately see that Jones is of an average build, has fair brown hair, and is wearing a raincoat.
>
> **Hallucination:** Lizzy is a scientist, who works on hallucinogenic drugs. She slips a novel drug that she is testing into Mary's drink, without Mary ever

coming to be aware of her presence. By some strange fluke, after ingesting the drug, Mary has a photo-realistic, three-dimensional hallucination, which is qualitatively indiscernible from Lizzy, from her New Balance trainers to the scar on her cheekbone.

In **Hologram**, Smith is able to refer to Jones in virtue of his causal connection to Jones, even though the causal chain connecting them is extremely complicated. In **Hallucination**, Mary is *not* able to refer to Lizzy, despite the fact that there is a causal connection between them (Mary would not have had that hallucination if not for Lizzy). In some sense, the causal chain between Mary and Lizzy is *deviant*, while the causal chain connecting Jones and Smith is not. But in what does the difference between the two cases consist? Why is the former chain deviant, but the latter not?

The *benefit* of accepting **Hypothesis** that it arguably enables a straightforward answer to these questions. Namely, the reason that Smith is able to refer to Jones in **Hologram** in virtue of the causal connection between them is that Smith is in a position to acquire *knowledge* about Jones in virtue of that causal connection. Conversely, the reason that Mary is not able to refer to Lizzy in **Hallucination** in virtue of the causal connection between them is that Mary is not in a position to acquire knowledge about Lizzy in virtue of that causal connection. Crucially, the appeal to the problem of deviant causal chains provide support for the *right-to-left* direction of **Hypothesis**—that is, some suitably qualified version of (1)—rather than merely for the uncontroversial left-to-right

direction. For suppose that the right-to-left direction of **Hypothesis** is false. Then it is possible that some causal connection is a channel for reference, even though it is not a channel for the acquisition of knowledge. But this threatens to undermine the explanation of the fact that Mary is not in a position to refer to Lizzy in **Hallucination** in terms of the fact that she is not in a position to acquire knowledge about Lizzy, given that, on this supposition, the latter fact obtaining is not sufficient for the former to obtain. In other words, given that there *are* cases in which some causal connection suffices for reference, without sufficing for the acquisition of knowledge, then why is **Hallucination** not one such case? Accepting (1) has the effect of closing off this question.

However, there is a complication in this argument for **Hypothesis**. For note that it does not by itself explain why Mary is not in a position to have knowledge about Lizzy, while Smith is in a position to have knowledge about Jones; it simply takes this difference for granted. But then, it is at least fair to wonder why the fact that Mary is not in a position to refer to Lizzy, while Smith is in a position to refer to Jones, cannot itself be taken as primitive. That is, even if it is accepted that the differences coincide, why think that the latter difference should be explained in terms of the former, rather than that neither needs be explained in terms of the other, or indeed, that the former should itself be explained in terms of the latter? Even though **Hypothesis** does not technically encode any claim about the direction of explanation, it is most interesting—and best supports the knowledge-maximisation principle—on the assumption that it reflects the idea

that causal chains which conduce to reference do so *because* they conduce to knowledge. Henceforth, this is how we shall understand it.

The simple answer is that, at least given certain background assumptions about knowledge, there is independent reason to think that reference should be explained in terms of knowledge, rather than the other way around. In particular, on Williamson's own view, knowledge is the central relation that obtains between mind and world. Correspondingly, Williamson takes knowledge to be essential to the nature of belief (2000: §1.5); to the nature of epistemic justification and of evidence (2000: §§9–10); to the practice of assertion (2000: §11); and to the justification of action (2017a). None of these proposed theoretical roles for knowledge is uncontroversial, and certainly the fact that knowledge putatively plays these other roles does not detract from the requirement to evaluate **Hypothesis** on its own terms. However, the question we are currently considering is why, even if differences between which causal chains are channels for reference *do* coincide with differences between which causal chains are channels for the acquisition of knowledge, it should be thought that the latter differences help to *explain* the former ones. And an appropriate (if defeasible) answer to that specific question is that, on the assumption that knowledge does play these various theoretical roles, then we have independent evidence that knowledge-theoretic distinctions are often explanatory.

Roughly, then, cases like **Hologram** and **Hallucination** at least suggest that there is some interesting relation between causal chains that, because they are

deviant, are not channels for reference, and causal chains that, because they are deviant, are not channels for knowledge; and if some relation *does* obtain. Further, the problem of deviant causal chains provides crucial support for the right-to-left direction of **Hypothesis**, specifically. We now turn to the costs of *rejecting* **Hypothesis** for the tractability of the problem of deviant causal chains. In particular, the opponent of **Hypothesis** might be tempted to think that, even if she denies **Hypothesis**, she can still identify non-deviant causal chains as the ones that involve, for instance, perception or awareness, or a high degree of reliability. However, there is a serious problem for this line of response—that it scarcely shows that being in a position to have knowledge in virtue of some causal connection is *not* necessary for being in a position to refer in virtue of that causal connection. Indeed, the friend of **Hypothesis** can potentially regard each of these suggestions as plausible precisely in virtue of the connection that she posits between reference and knowledge.

More specifically, it is natural for the friend of **Hypothesis** to endorse Williamson's well-known view that 'knowing is the most general stative proposition attitude such that, for all propositions *p*, necessarily if one has it to *p* then *p* is true' (Williamson 2000: 39), as an independently well-motivated claim in the philosophy of mind.[7] But then, if either (3) or (4) is true, then so is (1):

[7] For defence, see Williamson (2000: §1.4), Unger (1975: §4), and, more recently, Magidor and Hawthorne (2018). Hyman (2014) criticises Williamson's equation of the claims that (i) knowing is the most general factive stative attitude and (ii) knowing **p** is necessary for having some factive stative attitude towards **p**. However, Hyman's arguments target (i), not (ii); nothing in what follows requires the stronger claim. As far as I know,

(3) A subject S is in a position to refer to an object O in virtue of some causal connection between S and O only if S is in a position to be *aware* that O exists in virtue of that causal connection.

(4) A subject S is in a position to refer to an object O in virtue of some causal connection between S and the object O only if S is in a position to *perceive* that O exists in virtue of that causal connection.

For one is only *aware* that something is so if one *knows* that it is so, and one only *perceives* that it is if one *knows* that it is so. Contraposing, if one does not know that something is so, then one is neither aware that it so, nor perceives that it is so. To put the point another way, given the thesis that knowledge is the most general factive stative attitude, in the sense that, for any factive stative attitude, having that attitude implies knowing, a theorist who rejects (1) is committed to the possibility that some causal connection suffices for the possibility of referring to an object without thereby sufficing for the possibility of having some factive stative attitude towards a states of affair involving that object. The following section will examine in detail one case in which this possibility arguably obtains. But, for the moment, the crucial observation is that, consequently, a theorist who rejects **Hypothesis** risks severely limiting the range of concepts to which she can appeal in addressing the problem of deviant causal chains. For instance, it is clear that one difference between **Hologram** and **Hallucination** which is relevant to the fact that Smith can refer to Jones but Mary cannot refer to Lizzy, is that Smith is in a position to see that Jones is wearing a raincoat, while Mary is not in a position to see that Lizzy has a scar. But since seeing that something is so is a

paradigmatic factive stative attitude, the friend of **Hypothesis** is in a much better position to emphasise this difference than her opponent.

Just the same point applies to the suggestion that the difference between **Hallucination** and **Hologram** consists in the fact that there obtains a *reliable* causal connection between Jones and Smith, but not between Lizzy and Mary. Again, this seems like a natural and plausible diagnosis of the difference between the two cases. But again, whether it is a diagnosis that is in tension with, let alone inconsistent with, **Hypothesis** is a very different question. For it is also the case that a reliable causal connection between a subject and an object must typically obtain for that subject to be in a position to acquire *knowledge* about the object. So, if the suggestion is that the causal connection between Mary and Lizzy is not reliable enough for Mary to have knowledge about Lizzy, while the causal connection between Jones and Smith is reliable enough for Smith to have knowledge about Jones, then the friend of **Hypothesis** can fully accept it. But if the required degree of reliability that a causal connection must have for it to be a channel for reference is also *not* to be understood in terms of knowledge, then how is it to be understood? More suggestively, we might wonder whether the degree of reliability necessary for reference can go much lower than the degree of reliability necessary for knowledge. For instance, consider the following simple thought experiment:

> **Horizon:** A boat is sailing off into the horizon, as Jones stands on the dock. At time t_0, the boat is right in front of Jones, so, if Jones were to look at t_0, he

could see it. By t_n, the boat has sailed hundreds of miles, so, if Jones were to look at t_n, he could not see it.

Between t_0 and t_n, the causal connection between Jones and the boat grows increasingly unreliable; by the end of the process, it is clear that it is far too unreliable to be a channel for either reference or for knowledge. Let t_k be the first time at which the boat has sailed far enough into the horizon for it be the case that, if Jones were to look at t_k, he would not be able to see that the boat existed (know that the boat existed, perceive that it existed, etc.). If Jones were to look at t_k, would the causal connection between him and the boat nevertheless be reliable enough to constitute a channel for reference? That is, given that Jones would not in a position to be *aware* at t_k that the boat existed, would his causal connection to the boat still suffice for him to refer to it—without the help of some definite description like 'the object in that direction, if there is one'?

If the answer to this question is *no*, then that would be very good news for **Hypothesis**, since it would be defeasible evidence that (1) held in close to full generality. But even if the answer is *yes*, that this result should concern the friend of **Hypothesis** is not clear. It would be a significant objection to **Hypothesis** if some of the causal connections that were channels for reference had no interesting similarity to any causal connections that were channels for knowledge, since then it would look quite *ad hoc* to postulate that the property of being a channel for reference could be explained in terms of the property of being a channel for the acquisition of knowledge. But that some causal connections

suffice for reference despite being marginally less reliable than any causal connection that suffices for knowledge does not establish any such dissimilarity. Rather, what it may show is merely that *referring* to something constitutes less of a success state than *knowing* that it is. After first sketching two strategies for defending (1) against a putative counterexample, the following section shows how a version of this idea can be deployed in support of **Hypothesis**.

1.3 Fake Barns and Reference

The preceding section attempted to offer general reasons that one might believe that **Hypothesis** holds, particularly from a knowledge-first perspective. In this section, I turn to a putative counterexample to **Hypothesis**, due to McGlynn (2012). In broad outline, McGlynn's case purports to show that the causal connections that suffice for reference may obtain even though any beliefs formed on the basis of those causal connections are not *safe* enough to constitute knowledge. It runs as follows:

> **Hallucinogen**: Jane is taking part in a trial of a new drug with powerful hallucinogenic properties. Sitting in her cubicle, she appears to see a number of people in front of her. In fact, there is only one other person in the cubicle, Helen, who is standing several feet away from her. Helen appears to Jane to be no more flesh and blood than the other figures that appear to her in the cubicle. Despite knowing that she will experience hallucinations while in the cubicle, Jane judges 'She is beautiful', 'She is freckled', 'She has green eyes', and so on, keeping her

> attention firmly on Helen the whole time. Each of these judgements is correct (McGlynn 2012: 398).

According to McGlynn, Jane's thoughts do refer to Helen, even though she is not in a position to know that Helen is beautiful, that she has freckles, that she has green eyes, and so on. Whether this is sufficient to falsify (1), let alone **Hypothesis**, is not clear; much more on the significance of **Hallucinogen** for the status of **Hypothesis** below. First, however, it should be briefly noted that it is not clear that **Hallucinogen** constitutes even a prima facie counterexample to McGlynn's intended target, the knowledge-maximisation principle of charity itself, even taken in unqualified form. For, according to the knowledge-maximisation principle, an assignment of reference is correct only if it maximises knowledge. Yet this does not entail that an assignment of reference is correct *if* it maximises knowledge. Indeed, this latter principle would be rather implausible, for a reason which is familiar from discussions of truth-maximisation principle of charity—namely, that principles of charity may underdetermine reference. More specifically, if any assignment of reference that maximised knowledge is correct, then, whenever more than one assignment of reference equally maximises knowledge, multiple assignments of reference are correct. But when multiple assignments of reference equally maximise knowledge, it still seems possible that at least some such assignments are *not* correct, on the basis that they fail to satisfy some further constraint on the assignment of reference, which either adjudicates between assignments of reference that equally maximise knowledge,

or counterbalances, in some more complicated way, the knowledge-maximisation principle of charity itself.[8]

What is worth noticing about **Hallucinogen**, then, is that an assignment of reference to Jane's beliefs on which 'she' refers to Helen does not clearly fail to maximise her knowledge; at least, there is no obvious alternative assignment of reference to Jane's thoughts on which Jane would know *more* than she would know on an assignment of reference to her thoughts on which 'she' refers to Helen.[9] Thus, **Hallucinogen** is perfectly consistent with the principle that any correct assignment of reference maximises knowledge. At best, the case shows that the knowledge-maximisation principle requires supplementation, on the basis that reference need not fail even when the subject lacks knowledge on any candidate assignment of reference. But McGlynn does not offer any reason to think that this cannot be understood as a corollary of the more general fact that some assignments of reference which maximise knowledge might yet fail to be correct. Of course, without further details about how the knowledge-maximisation principle of charity should be supplemented, a response to **Hallucinogen** which simply notes that the case is consistent with the principle is

[8] As mentioned in fn1, the most familiar further constraint is that the objects of reference should be relatively *natural*. Note that epistemicists may be particularly well-advised not to think that an any knowledge-maximising assignment of reference is correct. For, given epistemicism's commitment to at most one assignment of reference being correct, reference would then fail whenever no assignment was uniquely knowledge-maximising, which seems likely in at least some cases of partial stipulation. See Hawthorne (2006a) for a discussion of naturalness in the context of epistemicism.

[9] At any rate, McGlynn himself does not argue that there is. His objection to the knowledge-maximisation principle turns on the claim that Jane's thoughts do refer to Helen, even though she lacks knowledge of Helen, rather than that they do not refer to someone else, even though they would constitute knowledge if they did.

incomplete. We will to some extent return to whether these details can be provided in §1.4. For the rest of this section, however, we turn to the implications of **Hallucinogen** for **Hypothesis**, to which it at least seems to be more immediately relevant.[10]

Evidently, there are two rather straightforward responses that a friend of **Hypothesis** might give to **Hallucinogen**, particularly if she moreover accepts (1). First, she might argue that Jane *is* in a position to know at least some facts about Helen, despite the fact that is under the influence of a hallucinogen. Alternatively, she might deny that Jane is in a position to refer to Helen, despite the fact that Helen is, in some sense, in Jane's visual screen. Each of these responses has more to be said for it than it might seem. Let us take each in turn. After considering each option, we will turn to a more conciliatory response,

[10] One response to **Hallucinogen** on behalf of the friend of the knowledge-maximisation principle should be resisted, however. Namely, Manuel García-Carpintero takes McGlynn's case to show that 'Williamson's proposal is too blunt', before briefly suggesting that '[a]n adequate metasemantics should instead look in more detail at the role that knowledge plays in the semantic rules for particular expressions' (2018: 1125). The thought seems to be that the knowledge-maximisation principle might help to determine the *characters* of certain expressions, without playing as central role in determining their *reference* at a given context of use. For instance, 'she' should have a knowledge-maximising character, but considerations of knowledge need not always determine its referent. However, this response to **Hallucinogen** trades on a confusion between 'context' in the ordinary sense and 'context' in the technical, Kaplanian sense. Certainly, what determines the referent of 'she' at a given context of use, in the technical sense, is simply the semantic rule that 'she' denotes the relevant female at that context. However, what's to be explained is which features of the context, in the ordinary sense, determine it to be the case that the context, in the technical sense, is such that Helen is the relevant female at that context (if she is). For the friend of the knowledge-maximisation principle to concede that considerations of knowledge do not play a central role in this explanation would in effect be for her to concede the argument, since clearly *some* principle of charity is relevant to the contextual determination of reference (for relevant discussion, see Dorr 2014). Incidentally, it is not clear that Jane's use of the demonstrative 'she' is essential to **Hallucinogen**; at least intuitively, she could equally have used an arbitrary proper name.

which concedes that **Hallucinogen** falsifies (1), but denies that this result is in tension with **Hypothesis**.

1.3.1 Hallucinogen – Jane Knows

In effect, McGlynn's argument against **Hypothesis** from the possibility of **Hallucinogen** takes as its first premise that Jane is not in a position to acquire any knowledge about Helen. As McGlynn notes, the central basis for the judgement that Jane is not in a position to acquire any knowledge about Helen is that **Hallucinogen** is essentially a modified *fake barn* case.[11] In a standard fake barn case, the subject sees what is in fact a barn and truly judges that it is a barn. However, had she looked out her car window several seconds earlier or later, she would have seen what was in fact a barn facsimile, and falsely judged that this facsimile was a barn. Thus, according to orthodoxy, the subject's judgement that the object in front of her is a barn is too unreliable to constitute knowledge, given that she could so easily have had a relevantly similar false belief. Of course, a standard fake barn case would not invalidate (1), since intuitively the subject in a fake barn case *is* in a position to know some facts about the object that she sees in virtue of her causal connection to it, such as that it is in front of her, that it appears to be a barn, that it exists, and so on. Hence, we can infer (although McGlynn does not say this explicitly), the central modification that McGlynn

[11] McGlynn himself also emphasises—but is careful not to rest his argument on—the fact that Jane has defeaters for her belief (eg, she knows that she may be hallucinating). McGlynn's reluctance seems to me appropriate—any conception of knowledge on which **Hypothesis** is plausible is likely to be highly anti-defeat. However, the considerations in §1.4 may help indicate how **Hypothesis** can be rendered consistent with a defeat-friendly epistemology.

makes to the standard formula: given that any belief Jane would have formed had she not been staring at Helen would have suffered reference failure, she is putatively not even in position to know that Helen exists.

More specifically, in discussing **Hallucinogen**, McGlynn appeals to a safety-theoretic condition on knowledge, which he fairly assumes that Williamson would accept. For concreteness, here is Williamson's own version of the principle:

> **Safety:** If in a case α one knows **p** on a basis b, then in any case close to α in which one believes a proposition **p*** close to **p** on a basis close to b, **p*** is true (2009a: 325).[12]

As McGlynn correctly notes, **Safety** is importantly more demanding than the following, very simple safety principle:

> **Simple Safety:** If in a case α one knows **p**, then in any case close to α in which one believes **p**, **p** is true.

In particular, given **Simple Safety**, one's belief that **p** is unsafe only if there is some close case in which one falsely believes **p** itself. Familiarly, however, **Simple Safety** fails to explain the fact that, if one merely guesses that **p**, then even if **p** is a necessary proposition, so one could not falsely have believed **p** itself,

[12] Compare McGlynn's version of **Safety**: 'S knows that P only if in all the nearby possible worlds in which S forms a relevantly similar belief on a relevantly similar basis as S forms her belief that P in the actual world, that belief is true' (2012: 398). For further discussion of safety-theoretic conditions on knowledge, see Hawthorne & Lasonen-Aarnio (2009), Williamson (2009a) and (2009b), Kearns & Magidor (2008), and Goldstein & Hawthorne (ms).

one's belief that **p** itself may still fail to constitute knowledge. By contrast, **Safety** can explain one's ignorance, given that some false proposition is close to **p**. Thus, again as McGlynn correctly notes, it is implausible to resist the verdict that Jane's beliefs in **Hallucinogen** do not constitute knowledge on the basis that Jane could not easily have formed those very false beliefs. That Jane could have formed relevantly similar false beliefs, in attempting to refer to some hallucination, may be enough for her actual beliefs to be unsafe.[13]

However, McGlynn curiously fails to note that, as well as being more demanding than **Simple Safety** in one respect, **Safety** is also *less* demanding than **Safety** in another. In particular, unlike **Simple Safety**, **Safety** incorporates a relativisation to *bases*. The motivation for this relativisation is again familiar. Namely, if one forms the belief that **p** on the basis **b**, then the fact that in some close cases, one falsely believes **p** on the very different basis **b'** does not obviously undermine the good epistemic status of one's belief that **p**.[14] Thus, given **Safety**, Jane's beliefs fail to constitute knowledge in virtue of not being safe only if, in some close case in which she forms relevantly similar false beliefs, these beliefs are formed on a basis that is close to the one on which she forms her actual beliefs. Similarly, given **Safety**, in a standard fake barn case, the subject's belief that there

[13] In fact, there is a complication here: in cases close to **Hallucinogen** in which Jane attempts to refer to some object which does not in fact exist, her thoughts plausibly suffer reference failure, so that there is no *proposition* which she believes at all. For a fix appealing to the notion of counterpart *thoughts*, see Manley (2007). For present purposes, we bracket this issue, since the broader criticism of **Simple Safety** is unaffected.

[14] 'I may know that the Prime Minister was in Oxford today, because I happened to see him, even though he might easily have cancelled, in which case I would still have had the belief, on the basis of the newspaper announcement which I read this morning' (Williamson 2009b: 20).

is a barn in front of her fails to constitute knowledge in virtue of not being safe only if, in some close case in which she falsely believes that there is a barn in front of her, this belief is formed on a basis that is close to that on which she forms her actual belief. Must the friend of **Hypothesis** accept these judgements about **Hallucinogen** and fake barn cases, respectively?

It is not the task of this essay to attempt to settle this question. However, at least in **Hallucinogen**, it is hardly *clear* that the basis on which Jane forms her actual beliefs is relevantly similar to the basis on which she forms false beliefs in the close cases in which she does so. For, on one natural description of the case, Jane forms the beliefs 'she is beautiful', 'she has freckles', and 'she has green eyes', on something like the basis of *seeing Helen*. Indeed, for all that McGlynn says about **Hallucinogen**, it might be that Jane's vision itself is working quite well, at least in local respects. But then, it seems open to the friend of **Hypothesis** to deny that the false beliefs that Jane forms in close cases are in fact formed on a similar basis. For those beliefs are formed not on the basis of seeing Helen, but on the basis of *hallucinating*. Yet what makes beliefs that Jane forms on the basis of seeing Helen similar to beliefs that she would have formed on the basis of hallucinating something else? Of course, Jane herself might lack a reliable way of distinguishing between beliefs formed on the two bases. But what is at issue is whether the bases themselves are relevantly similar, not whether Jane could know that she was forming beliefs on one basis, rather than on the other.[15] That

[15] Similarly, on a natural description of a standard fake barn case, the subject forms the belief that there is a barn in front of her on something like the basis of *seeing a barn*. In close cases in which

one was dreaming the moment before is consistent with one's acquiring knowledge through one's causal connection to the external world given that one is now awake. At least in principle, the friend of **Hypothesis** could take a similar line towards **Hallucinogen**.

1.3.2 – Jane Does Not Refer

Nevertheless, since it *is* plausible that Jane's beliefs do not constitute knowledge in **Hallucinogen**, it would be risky to rest the defence of **Hypothesis** on contentious considerations about the basis on which Jane comes to form her beliefs. Consequently, let us turn to the second envisaged reply on behalf of the friend of **Hypothesis** to **Hallucinogen**, namely, denying that Jane's thoughts in fact refer to Helen. For ease of discussion, we assume that Jane is not in a position to acquire any knowledge about Helen in virtue of her causal connection to her. The important question is thus: why think that, *given* this assumption, Jane successfully refers to Helen in **Hallucinogen**? Or, more precisely, why think that, given that she is not in a position to have knowledge about Helen, Jane is in a position to refer to Helen *in virtue of* her causal connection to Helen?

Let us start by clarifying the distinction implicitly marked by the previous sentence. For, crucially, **Hallucinogen** is *not* a counterexample to (1) if Jane only successfully refers to Helen partly in virtue of first thinking to herself, for

her belief is false, she forms it on the basis of *seeing a barn facsimile*. Again, the subject might lack a reliable way of distinguishing between beliefs formed on the two bases. But what is at issue is whether the bases themselves are relevantly similar, not whether the subject could know that she was forming beliefs on one basis, rather on the other. For critical discussion of orthodoxy on fake barn cases, see Gendler & Hawthorne (2005).

instance, 'there seems to be a person in front of me; I will use "she" to refer to that person, if there is one'. To see that it is not, notice that Mary could in principle refer to Lizzy in just the same way in **Hallucination**. That is, Mary could think to herself 'there seems to be in a person in front of me; I will use "she" to refer to the scientist who caused me to have this visual impression of someone identical to herself, if there is one'. Of course, it would be strange if Mary did think this—but *if* she did, she would plausibly single Lizzy out as the referent of her term 'she'.[16] Nevertheless, this use of a uniquely identifying definite description to fix reference is presumably consistent with (1), on its intended reading; in some sense, it would constitute an extra-causal element with respect to the determination of reference. But then, that Jane refers to Helen in **Hallucinogen** is similarly not a counterexample to (1) if Jane is only in a position to refer to Helen in virtue of possessing a uniquely identifying definite description for Helen. While the fact that Jane possesses that definite description is partly to be explained by the fact that she bears some causal connection to Helen, that casual connection itself does not need to be a channel for reference in its own right.

I emphasise this possibility for two reasons. First, it illustrates one way, consistent with (1), that Jane might come to refer to Helen in **Hallucinogen**. But second, and more importantly, it makes salient that the notion of interest is that

[16] I rely here on the assumption that definite descriptions are easily converted into names; an alternative may be that Mary uses 'she' to express a disguised definite description. For relevant discussion, see Kripke (2011), Kaplan (2012), and Blumberg & Holguín (2018). A descriptivist alternative would similarly be consistent with **Hypothesis**; Williamson (2009c) makes a related point in defending the knowledge-maximisation principle against a possible counterexample due to Mike Martin (2009).

of a causal connection which is a channel for reference in its own right. Now, this notion is clearly vague. However, on closer inspection, this vagueness is a deeper problem for the objection to **Hypothesis** from **Hallucinogen** than it is for **Hypothesis** itself. For it serves to reinforce the arguments for **Hypothesis** advanced in the previous section. In particular, **Hypothesis** supports the following, less vague condition on whether Jane is in a position to refer to Helen in virtue of the causal connection that she bears to her:

(5) Jane is in a position to refer to Helen in virtue of her causal connection to Helen if and only if she is in a position to see that Helen exists.[17]

For vision is surely a causal process, so if one is in a position to refer to some object in virtue of seeing that it has certain features, then one is in a position to refer to that object in virtue of the causal connection that one bears to it. However, given that Jane is *not* in a position to know that Helen exists, then, by the fact that seeing that something is so entails knowing that it is so, Jane is not in a position to see that Helen exists. Thus, (5) cannot be leveraged to explain why Jane's causal connection to Helen constitutes a channel for reference, whereas, for instance, Mary's causal connection to Lizzy does not—at least, not without retracting the judgement that Jane is not in a position to acquire knowledge concerning Helen. In effect, this is the argument of the previous section, now applied to a specific case: given that one rejects **Hypothesis**, one loses a notable class of resources for explaining why the causal connection between Jane and

[17] (5) should be read as a material biconditional, unlike (1)–(4) or (6) below.

Helen is non-deviant. On the assumption that Jane is not in a position even to be *aware* that Helen has certain features in virtue of his causal connection to her, then what about her causal connection to Helen nevertheless suffices for Helen to be the object of her thoughts?[18]

1.3.3 Jane Doesn't Know, Does Refer

Thus far, we have considered two responses to **Hallucinogen**, both of which aim to show that the case is consistent not just with **Hypothesis**, but with (1). Nevertheless, let us suppose that *both* of the following judgements about **Hallucinogen** remain plausible: because her causal connection to Helen is not sufficiently reliable, Jane is not in a position to acquire knowledge about Helen; and because Jane seems, in some sense, able to visually attend to Helen, she is in a position to refer to Helen in virtue of her causal connection to her. If both of these judgements are correct, then (1) is false. But even if they are not, the friend of **Hypothesis** might still hope to explain the fact that they are jointly plausible. Thus, the remainder of this section sketches a more conciliatory response to **Hallucinogen**, which concedes that Jane *is* in a position to refer to Helen, despite not being in a position to have knowledge about her, but which argues that this can be explained on the basis that **Hypothesis** is true.[19]

[18] In principle, the opponent of **Hypothesis** could insist that Jane can see that Mary has certain features, even though she does not know that she does. But as Magidor and Hawthorne (2018) argue, of the three possible reactions to fake barn cases—accept that the subject knows that *p*; deny that the subject can see that *p*; reject the inference from seeing that *p* to knowing that *p*—the third reaction is the least plausible.

[19] Readers convinced by the foregoing considerations, if any, are encouraged to read this explanation as approximating an error-theory of the relevant pair of judgements.

Perhaps one remaining reason to think that Jane does refer to Helen in **Hallucinogen** is closely connected to the suggestion, which was made in discussing the basis of Jane's beliefs, that, despite the fact that she has ingested a hallucinogen, Jane's visual system has not itself been impaired. As it were, the cognitive damage of the hallucinogen is elsewhere, insulated from the part of her brain that Jane uses for seeing itself. Very roughly, we can assume that Jane's visual system is, in some relevant sense, modular—consequently, Jane receives roughly the same visual information in **Hallucinogen** as she would receive were she in the same situation, but not on a hallucinogen.[20] Thus, at least in *local* respects, Jane's causal connection to Helen is highly reliable. To the extent that it is not reliable enough for knowledge, then it is not reliable enough for knowledge in virtue of factors that are extrinsic to the causal connection itself. Plausibly, some set of assumptions like these underwrites fake barn cases generally, in the sense that fake barn cases are understood, either implicitly or explicitly, to be ones in which, *if* the subject fails to know, she fails to know in virtue of factors that are extrinsic to the causal connection between her and some relevant object.

Now, suppose that this schematic description of **Hallucinogen** is correct. Then it is not clear why the fact that Jane is in a position to refer to Helen in virtue of her causal connection to her, if she is, should not be perfectly explicable in

[20] Since this is presumably a nomologically possible case, we can set aside how far anything like this is true of the actual human visual system; in its intended reading, **Hypothesis** states a quite general thesis about one connection between reference and knowledge (Williamson 2007: 251). Again, if one thinks that this assumption suffices for Jane to see that Helen has green eyes, etc, then one should also think it suffices for her to have knowledge of Helen.

knowledge-theoretic terms. In particular, if this is the right way of describing the case, then the similarity of **Hallucinogen** to a case in which Jane *does* have knowledge is part of its very structure. For, on this way of conceptualising it, the case seems in effect to be stipulated to be one in which Jane is in a position to acquire knowledge about Helen *but for* the fact that she is in an extrinsically unfriendly environment. Details like the fact that Jane is on a hallucinogen function to make the case realistic, but otherwise seem to be inessential; the same effect is presumably achieved by the stipulation that, had Jane looked anywhere else besides at Helen, then a non-veridical visual impression would have been beamed directly into her brain. Alternatively, to the extent that the hallucinogen itself is essential to the case, for instance because it *does* affect the functioning of Jane's visual system, preventing her from visually attending to Helen (despite McGlynn's description), then the judgement that Jane is in a position to refer to Helen in virtue of her causal connection to her seems less plausible.

Roughly, then, the suggestion is that the friend of **Hypothesis** can explain the fact that Jane refers to Helen in **Hallucinogen** (if she does), on the basis of (6):

(6) A subject S is in a position to refer to an object O in virtue of some causal connection that obtains between S and O if and only if, in all relevant non-sceptical cases, S is in a position to know that O exists in virtue of that causal connection.[21]

[21] An alternative to 'relevant non-sceptical' is potentially 'normal'. However, the literature on normality contains a range of proposals for which cases are normal, on which I wish to remain

Non-sceptical cases are here understood as the converse of *bad* cases. A *relevant* non-sceptical case is one in virtue of which subjects in the bad case do know that they are in the bad case. For instance, classically, brains-in-vats do not know that they do not know that they have hands; the reason is that there are relevant non-sceptical cases at which they do know that they have hands. Similarly, given that one is in fake barn country, one does not know that one does not know that there is a barn; the reason is that there are relevant non-sceptical cases at which one does know that there is a barn. Arguably, the difference between brain-in-vat-style bad cases and fake barn-style bad cases partly consists in the fact that, in the latter, one does not merely have knowledge at relevant non-sceptical cases, but one has that knowledge in virtue of the very causal connection that obtains in the bad case. It is precisely this feature that makes **Hallucinogen** an apparently compelling counterexample to (1). Whether it *is* a counterexample is perhaps not clear. But even if it *is* one, the friend of **Hypothesis** is excellently positioned to explain why Jane is nevertheless in a position to refer to Helen in **Hallucinogen**. Namely, she is in a position to refer to Helen because the causal connection between her and Helen is effectively transplanted from a situation in which Jane

neutral. Further, the *ordinary* term 'normal' seems to equivocate between a normative and descriptive sense. In one sense, a normal heart is one which is capable of pumping blood, since hearts should pump blood. In another sense, a normal heart is one which is like most other hearts, and there are possible situations in which most hearts are dysfunctional. For relevant discussion of normality, see Goodman and Salow (2018), Carter (2019), and Goldstein & Hawthorne (ms). The more general strategy of taking some non-knowledge condition to obtain only because either it itself is similar enough to knowledge, or because it arises under conditions similar enough to ones which produce knowledge, is borrowed from Williamson (2000: §1) and (forthcoming-a) on belief; Unger (1975: §5) and Williamson (2000: §9) on justification; and, especially, Williamson (2017b), (forthcoming-b), (forthcoming-c), (forthcoming-d) on reasonable belief and rationality. This latter connection in particular is made more explicit in the following section.

could have knowledge in virtue of the fact that it obtains into one in which, for reasons quite extrinsic to that connection, she cannot.[22]

1.4 'Belief aims at knowledge'

In this final section, I relate the preceding discussion of **Hypothesis** to the slogan 'belief aims at knowledge', implemented within a broadly functional framework, before tentatively and briefly discussing some possible implications of that framework for the knowledge-maximisation principle.

According to a well-known slogan of Williamson's knowledge-first programme, 'belief aims at knowledge' (ibid: 48). In *Knowledge and its Limits*, this slogan is closely associated with two more specific claims. First, that 'various mental processes aim at more specific factive mental states', for instance '[p]erception aims at perceiving that something is so' and 'memory aims at remembering that something is so' (ibid: 48). Second, 'that knowledge is the evidential standard for the justification of belief' (ibid: 47), in roughly the sense that a belief is justified just to the extent that it is made probable by that which one knows. More recently, Williamson has integrated both of these ideas within

[22] Incidentally, Magidor (2018) argues that it cannot simply be assumed that brains-in-vats, for instance, are not in a position to know that they do not have hands; the same externalist considerations that support the claim that agents in the actual world know that they have hands support the possibility that some brains-in-vats are, in Magidor's terminology, in a position to know that they are envatted. One implication of this which Magidor emphasises is that 'care needs to be taken in showing what kind of counterexamples refute transparency and luminosity principles' (2018: 177). A stronger moral (not one which Magidor draws), however, is that our grasp of which situations are relevant to sceptical arguments is not independent of our grasp of which worlds are epistemically accessible from such situations. If correct, this moral arguably further supports the more specific assumption in the text that fake barn-style bad cases are naturally understood partly in terms of their relation to relevant non-sceptical cases.

a broadly functional framework, so that the *norm* on belief is to be knowledge in a sense at least analogous to the sense in which it is the norm on a heart that it is capable of pumping blood (see especially Williamson forthcoming-c). Roughly, cognitively speaking, all goes well *locally* when a particular belief constitutes knowledge, while all goes well *globally* when the cognitive system produces enough knowledge without producing too much false belief. Of course, since the cognitive system can only produce knowledge at the risk of producing false belief, and can only avoid producing false belief at the risk of not producing knowledge, things are rather unlikely to go perfectly well (ibid: 7).

Within this framework, (1) roughly comes to the claim that the cognitive system is only capable of producing *belief* in virtue of some causal connection if it at least *could* produce knowledge in virtue of that causal connection. For belief, like knowledge, is only possible given reference; one cannot believe that something exists without being able to refer to it. Correspondingly, the arguments in §1.2 attempted to provide support for this restriction on the possibility of forming beliefs about an object in virtue of one's causal connection to it, at least under ordinary circumstances. In turn, the question that is raised by **Hallucinogen** is whether belief-formation processes would, as it were, break down completely in some environment that is sufficiently inimical to the production of knowledge. For again, if Jane cannot *refer* to Helen in virtue of her causal connection to her because she cannot have knowledge in virtue of that connection, then she cannot have *beliefs* about Helen in virtue of that connection because she cannot have knowledge in virtue of it.

Roughly, then, the point of (6) is to concede the possibility of perceptual belief without the possibility of perceptual knowledge—but only on the condition that there is enough intrinsic causal similarity between a case in which one can have perceptual belief without perceptual knowledge and one in which one could have perceptual knowledge. For instance, if (6) is true, then Jane's perception *may* continue to function well enough for her to have beliefs about Helen without being in a position to have any knowledge concerning Helen, but only because the very perceptual process that she relies on to form beliefs about Helen in **Hallucinogen** *would* have produced knowledge about Helen had the non-sceptical case for which it evolved been the one which obtained. At least in principle, this does not seem like a consequence which should trouble a theorist who accepts the broadly functional account just described. For it arguably reflects the more general, and clearly acceptable, principle that belief-formation processes often function similarly across non-sceptical cases and corresponding bad cases. In other words, the cost of having belief-formation processes which are relatively simple, so treat like cases alike, and which produce knowledge in non-sceptical cases, is that they may produce false belief in cases relevantly similar to those non-sceptical cases. On the most conciliatory response considered in §2.3, **Hallucinogen** merely represents the limiting case of this idea; on the other responses we considered, it may not represent even that. Either way, it presents no deep threat to **Hypothesis**, so no deep reason to think that whether some causal connection is a channel for reference cannot be explained in knowledge-theoretic terms.

So much for the interpretation of **Hypothesis** within this functional framework. But where does 'belief aims at knowledge' leave the knowledge-maximisation principle? As far as I can tell, this depends. For, on one natural reading of the knowledge-maximisation principle, it seems roughly to fall out of the idea that belief aims at knowledge. That is, suppose it is true that the function of belief is to be knowledge. And suppose moreover that, if all is equal, then belief succeeds in this function. Then, at least on the reading of 'if all is equal' expressed by the previous sentence, if all is equal, then any correct assignment of reference to a subject's beliefs maximises knowledge, for, if all is equal, then subjects' beliefs will constitute knowledge. Further, the preceding considerations already seem to give us some grip on the ceteris paribus clause in this formulation. In particular, the fact that belief-formation processes may produce belief in unfriendly environments relevantly similar to environments in which those same processes produce knowledge partly explains why the correct assignment of reference to a subject's thoughts would not be one on which *all* of her beliefs constituted knowledge. For, contraposing, if the correct assignment of reference to a subject's thoughts were one on which *all* of her beliefs constituted knowledge, then belief-formation processes would *not* produce belief in unfriendly environments relevantly similar to ones in which those same processes produced knowledge. But then, they would have to be much more sophisticated than they actually are.

Of course, there may be other aspects to the knowledge-maximisation principle that are not naturally captured by the idea that belief tends to aim at

knowledge. For instance, there are considerations about how far the objects of reference should be relatively natural, which have been left largely unaddressed here. Further, there may exist a strong reading of the knowledge-maximisation principle on which it has, as it were, some degree of intercessory power—quite metaphorically, even if some perceptual process itself would have gone wrong had it been left by itself, the principle of charity suffices to ensure that reference is fixed to some object or property about which the subject does have knowledge. How plausible the knowledge-maximisation principle of charity should be considered to be, if it indeed does represent a commitment over and above the commitments implicit in the slogan that belief aims at knowledge, is a question that we leave open here.

The next chapter takes up the question of why, given the social determination of meaning, it should follow that there are no analytic sentences. As we shall see, the answer may be closely related to the fact noted above that a cognitive system which never produced false belief would, almost certainly, be one that also produced very little knowledge.

2 Analyticity and the Social Determination of Meaning

2.1 Introduction

Let a sentence **s** be epistemologically analytic (henceforth, simply 'analytic') if and only if, necessarily, anyone who understands **s** assents to **s**. That is:

(1) For any sentence **s**, **s** is analytic if and only if, necessarily, one understands **s** only if one assents to **s**.

Terms and sentences are assumed to have their meanings essentially, so that two sentences which differ in meaning count as distinct sentences, even if the sentences are homophonic. In this sense of 'analytic', Timothy Williamson has notoriously argued that there are *no* analytic sentences in natural languages. More precisely, Williamson rejects:

Analyticity: For some sentence **s**, necessarily, one understands **s** only if one assents to **s**.[23]

For present purposes, we count rules of inference as sentences. Thus, Williamson's rejection of **Analyticity** encompasses the claim that there is no inference rule that one understands only if one assents to that rule. The effect is to undermine a broadly inferentialist picture of meaning, on which one's

[23] See, especially, Williamson (2007: §4). The presentation here follows Williamson's own, albeit in a much less detailed way. In particular, Williamson also argues that *dispositions* on assent are never required on understanding. Since the central interest of this chapter is in understanding the failure of **Analyticity** given the social determination of meaning, I will partly be taking the fact that **Analyticity** *does* fail for granted. In any case, for further discussion, see Williamson (2003), (2011) and (2012), which reply to Boghossian (2003), (2011) and (2012), respectively.

understanding of a term consists in one's assent to certain sentences containing that term. Given that one may understand any sentence containing, for instance, the term 'if' without assenting to that sentence, then one's understanding of the term 'if' cannot simply consist in assent to certain sentences containing the term 'if'. More generally, for any sentence **s**, two speakers may both understand **s**, and thus all the terms in **s**, even though they do not mutually assent to **s**. Thus, disagreements between speakers, including disagreements about foundational logical principles, cannot uniformly be explained away as merely apparent disagreements. Often, if perhaps not always, speakers who assent to sentences that seem to be contradictory, assent to sentences that *are* contradictory—they simply disagree about whether those sentences are true.

This chapter proceeds as follows. In §2.2, I rehearse the two central components of Williamson's discussion of **Analyticity**. First, Williamson provides a direct argument against **Analyticity**, with the aim of showing that, for any sentence **s**, a clearly competent speaker may cease to assent to **s** in virtue of accepting some sophisticated logical or semantic theory which predicts that one should not assent to **s**. Second, Williamson proposes a non-inferentialist conception of understanding, which does not have the consequence of validating **Analyticity**. On this alternative picture, meaning is socially determined: roughly, for any term **t** in a public language, to understand **t** just is to be a member of a linguistic community that uses **t**. In §2.3, I note that, although this conception of understanding does not validate **Analyticity**, it does not immediately *invalidate* **Analyticity** either. Rather, it only does so given the further assumption that, for

any term **t**, there are no sentences containing **t** to which one must assent in order to be a member of a linguistic community that uses **t**. Although one might accept this assumption simply on the basis that it correctly predicts that no sentence is analytic, Williamson's own discussion suggests a broadly normative argument for it, from the premise that it would be dysfunctional for a linguistic practice to rule out speakers who have *good reason* to dissent.[24] §2.4 attempts to state this argument more precisely, as well as to fill in the conception of analyticity to which it stands opposed. In §2.5, however, I argue that two initially attractive motivations for this argument do not clearly succeed, particularly given the externalist assumption that beliefs are justified only if *true*. Finally, I conclude by suggesting that a similar argument, but which emphasises the unreliability of even the best cognitive dispositions, may be more compelling.

2.2 Williamson on Analyticity

For ease of discussion, I will assume that both friends and opponents of **Analyticity** accept the following principle:

(2) For some sentence **s**, necessarily, one understands **s** only if one assents to **s** if and only if for some term **t** in **s**, necessarily, one understands **t** only if one assents to **s**.[25]

[24] This is not to imply that Williamson himself endorses the argument. Even if he does, it clearly plays at best a supporting role in his broader argument against **Analyticity**.
[25] Quasi-symbolically, to clarify scope: $\exists \mathbf{s}(\Box(\mathrm{U}\mathbf{s} \rightarrow \mathrm{A}\mathbf{s}) \leftrightarrow \exists \mathbf{t}(\Box(\mathrm{U}\mathbf{t} \rightarrow \mathrm{A}\mathbf{s})))$.

The left-to-right direction of (2) is surely uncontroversial. If, for some term **t** in **s**, necessarily, one understands **t** only if one understands **s**, then, given that presumably one understands a sentence only if one understands every term in that sentence, necessarily, one understands **s** only if one assents to **s**. By contrast, the right-to-left direction of (2) may require some finessing. For instance, it seems possible in principle for one to *individually* understand every term **t** in a sentence **s** without assenting to **s**, without its being possible for one to *jointly* understand every term **t** in **s** without assenting to **s**, in which case it would not be possible for one to understand **s** without assenting to it, even though, for any term **t** in **s**, it is possible that one understands **t** without assenting to **s**. Nevertheless, (2) helpfully allows us to treat **Analyticity** as materially equivalent to the claim that for some sentence **s** and some term **t** in **s**, failure to assent to **s** implies failure to understand **t**:

> **Possession:** For some sentence **s**, for some term **t** in **s**, necessarily, one understands **t** only if one assents to **s**.

Very roughly, the idea is that analytic sentences encode conditions on possessing concepts, albeit with the qualification that distinct terms in distinct languages may express the same concept; a French speaker can clearly have the concept *if* despite not understanding 'if'. The assumption that both the friend and the enemy of **Analyticity** accept (2) reflects the fact that much of the debate is carried out in such terms. In particular, **Possession** might be regarded as *explaining* **Analyticity**; the reason that there are analytic sentences is that some concepts in

effect encode conditions on possession. Further, more nebulously, the commitment to **Possession** might itself be explained by the commitment to some constitutive claim about understanding, such as (3):

(3) For some sentences $\mathbf{s}_1, \ldots, \mathbf{s}_n$, for some term **t** in $\mathbf{s}_1, \ldots, \mathbf{s}_n$, for one to understand **t** just is for one to assent to $\mathbf{s}_1, \ldots, \mathbf{s}_n$.

For instance, for one to understand 'and' just is for one to assent to the arguments 'A; B; therefore A and B', 'A and B; therefore A' and 'A and B; therefore B'; or for one to understand the 'is' of identity just is for one to assent to 'x is y only if x and y have all the same properties'; or, to take a non-logical example, for one to understand 'vixen' just is to assent to 'vixens are female foxes'. Ostensibly, that linguistic understanding consists in assent to certain sentences demystifies how it is that subject understand their terms. Although theorists who accept **Analyticity** on the basis of accepting **Possession** and (3) might also think that **Analyticity** is independently plausible, they need not; in principle, they might regard it as an unattractive prediction of an account of linguistic understanding that should nevertheless be accepted on the basis that to reject it would be to accede to a thorough-going scepticism about meaning. It is in large part against this background that Williamson's argument against **Analyticity** proceeds. First, Williamson provides a formula for constructing counterexamples to putatively analytic sentences, which he applies himself to several cases. The effect is to show that **Analyticity** should only be accepted, if at all, if there is no alternative account of how speakers understand those terms that the friend of (3) takes speakers to

understand in virtue of assenting to certain sentences containing them. Second, Williamson offers an alternative account of linguistic understanding, which does not entail **Analyticity**, thus undermining residual support that the principle might have possessed. The rest of this section rehearses these two components of Williamson's discussion.

Williamson's discussion is largely restricted to inference rules, as the best candidates for being analytic. As Paul Boghossian (2010) helpfully puts it, Williamson's strategy is to given a general 'recipe' for constructing counterexamples to the claim that some inference rule is analytic. Namely, one imagines a clearly competent speaker, who comes to accept some sophisticated logical or semantic theory according to which one ought not to assent to the candidate inference rule. Given that it is implausible that such a speaker thereby loses her competence with the rule in question, she constitutes a prima facie counterexample to the analyticity of that rule. For instance, consider the following two schemata, where '⊢' expresses logical consequence:

(4) A and B ⊢ A.

(5) If A, B; A ⊢ B.

It is natural to think that any speaker who understands 'and' would assent to (4), while any speaker who understands 'if' would assent to (5); failure to assent to (4) or (5) suggests a failure to understand 'and' or 'if'. That is, at least initially, (4) and (5) seem to be analytic—and thus to support **Analyticity**, independently of

any broader theory of linguistic understanding. However, according to Williamson, the judgement that it is *necessary* for understanding (4) and (5) that one assents to (4) and (5) represents, at best, an over-extrapolation from the data. Even if speakers who do not assent to (4) and (5) typically do not assent to them because they do not understand them, it should not be inferred that the only possible reason for failing to assent to either sentence is that one does not understand it. For consider the following two speakers:

> **Simon:** Simon takes 'A and B' to be indefinite (neither true nor false) whenever either A or B is indefinite. Simon also assents to all and only sentences that have designated truth-values; he takes the designated truth-values to be 'true' and 'indefinite', but not 'false'. Thus, given that A is false and B is indefinite, Simon assents to 'A and B', but not to A. For instance, Simon assents to 'Jones is tall and the sun is green', because Jones is a borderline case of tallness, but of course not to 'the sun is green'. Simon does not assent to (4).[26]

> **Vann:** Vann assents to 'if a Democrat wins the 2020 presidential election, then if it is not Joe Biden who wins, it will be Bernie Sanders', on the basis that Biden and Sanders are the only Democrats still officially running for president. He also assents to 'a Democrat will win the election', on the basis that Biden is significantly outpolling Donald Trump. However, he does not

[26] See Williamson (2007: 95–6). As noted in fn.23, I leave out a significant amount of Williamson's detail, which he provides partly in order to show that the relevant speakers also fail to be *disposed* to assent to the relevant sentences.

assent to 'if it is not Biden who wins, it will be Sanders', on the basis that Trump is more likely to win than Sanders. Vann does not assent to (5).

Plausibly, Simon is a counterexample to the analyticity of (4), and Vann a counterexample to the analyticity of (5). To deny that Vann continues to understand 'if' despite his failure to assent to (5) would be to deny understanding of 'if' to any number of experts on the semantics of 'if', not least, of course, Vann McGee.[27] This is a high price to pay. Admittedly, fewer theorists hold the same package of views as Simon. But could the natural thing to say about Simon really be that he fails to understand (4)? As Williamson suggests, 'once we become aware of Simon's deviant theory of logic, an explanation of his unwillingness to [assent to (4)] in terms of linguistic incompetence looks much less attractive' (2011: 499). In any case, here is another example:

Sarah: Unlike Simon, Sarah believes that one should only assent to a sentence if one knows that it is true. However, Sarah is interested in the sentence 'Conj', where 'Conj' is stipulated to be true if and only if it occurs embedded in a conjunction. Thus, given that she knows A, Sarah assents to 'A and Conj', but not to 'Conj'. Further, Sarah is an alethic relativist, happily willing to multiply the parameters relative to which sentences express propositions. In particular, along with the more familiar parameters of

[27] See McGee (1985). Kratzer (1981) also invalidates modus ponens. Mandelkern (forthcoming) offers putative counterexamples to weakened versions of modus ponens, while Mandelkern (ms) shows that McGee and Kratzer's theories also invalidate '⊢ if A, A'. Regardless, Angelika Kratzer understands 'if' if anyone does.

world and time, she thinks that there is a parameter for sentential context. Consequently, she regards 'Conj' as having the same meaning throughout the argument from 'A and Conj' to 'Conj'. Sarah does not assent to (4).[28]

Sarah has an even more specific reason than Simon for not assenting to (4). However implausible her background relativism might be, it seems highly unattractive to explain her unwillingness to assent to (4) by claiming that she does not understand (4). Rather, Sarah merely seems to be a theorist who has discovered a hard case for (4), about which most ordinary speakers might not know what to say, and resolved it in a natural way, given her other theoretical commitments. Once again, then, it is a rather high cost to deny that Sarah understands (4). While she is manifestly in error, she hardly seems to have lost competence with 'and'.

Suppose, then, that it is conceded that examples like **Simon**, **Vann**, and **Sarah** undermine the attempt to justify **Analyticity** on pretheoretical grounds, given that they suggest that (4) and (5) are not analytic, that (4) and (5) seem to be as strong candidates as any for being analytic, and that the cases are instances of a more generally applicable formula. Then this brings us to the second component of Williamson's discussion—how the commitment to **Analyticity**, once **Analyticity** has been conceded to be implausible, can be avoided. Clearly,

[28] Compare Russell (2018: 316), which gives a parallel (highly debatable) counterexample to conjunction-introduction. For details on alethic relativism, see MacFarlane (2007), as well as Cappelen & Hawthorne (2009).

since (3) validates **Analyticity**, some other conception of linguistic understanding is required to take its place. As Williamson writes:

> Despite all the evidence against the existence of links from understanding to assent or dispositions to assent, it can be hard to resist the idea that there *must* be such links, otherwise the distinction between understanding and not understanding would dissolve: speakers who all understood the same term might have nothing substantive in common to constitute its shared meaning (2007: 121; emphasis in original).[29]

Characteristically, the alternative picture of understanding that Williamson proposes is socially externalist. Roughly, whereas (3) suggests that the meaning of at least some terms that a speaker uses depends primarily on factors internal to that speaker—namely, the sentences to which she assents—Williamson takes the meaning of terms in a public language to be determined by patterns of use across the relevant linguistic community. More importantly, with respect to **Analyticity** itself, Williamson takes one to *understand* such terms in virtue of being a member of that linguistic community, rather than in virtue of assenting to sentences that contain them. For concreteness, we can state Williamson's

[29] It is interesting to compare this observation to the remark by David Kaplan that '[t]he notion of a historical chain of acquisition by which a name is passed from user to user, was first used to facilitate abandonment of the classical, description theory of proper names found in Frege and Russell' (1989b: 602). Indeed, in both its strategy and its conclusions, Williamson's attack on analyticity closely parallels Kripke's attack on descriptivism in *Naming and Necessity*. Both authors first use a series of similarly constructed cases to show that understanding cannot merely consist in assent to certain sentences, before appealing to social externalism to explain how understanding could *fail* to consist in assent to certain sentences. Again as Kaplan remarks: 'It is hard to see how to avoid some version of a description theory of proper name… if one maintains a subjectivist semantics. Thus the attack on description theory… is *a fortiori* an attack on subjectivist semantics' (1989b: 603n87).

alternative picture of linguistic understanding, which gives centre stage to speakers' participation in the social practice of using certain terms, as follows:

> **Social Determination of Meaning (SDM)**: For any term **t** in a public language, for one to understand **t** just is for one to fully engage in the public practice of using **t**.[30]

In broad outline, because the social determination of meaning imposes less demanding conditions on linguistic understanding than some condition like (3), it allows meaning to be more robust over differences in use. Thus, at least in principle, it conduces to the possibility that speakers might be using the very same terms, despite each failing to assent to central aspects of the other's doctrine—more strongly, to the possibility that, for *any* specific difference in assent, this difference reflects a difference in theory, rather than in meaning. Certainly, given that, according to **SDM**, linguistic understanding simply does not consist in assenting to certain sentences, **SDM** does not *validate* **Analyticity**. The rest of this chapter explores the question of why, given **SDM**, it would be expected that **Analyticity** *fails*.

2.3 No Common Creed

Take again **Possession**:

[30] For the clearest recent statement of this claim, see Williamson (2011a: 483), although similar remarks are scattered throughout Williamson (2007: §4), some of which will be discussed below. A similar conception of understanding of is articulated in Williamson (1994: 209–212).

Possession: For some sentence **s**, for some term **t** in **s**, necessarily, one understands **t** only if one assents to **s**.

We have assumed that **Analyticity** holds if and only if **Possession** holds, in virtue of assuming (2). Thus, **SDM** entails that **Possession** is false if and only if it entails that **Analyticity** is false. However, it should be clear that, taken by itself, **SDM** is not at all logically inconsistent with **Possession**. For technically, **SDM** is consistent with the following claim:

Common Creed: For some sentence **s**, for some term **t** in **s**, necessarily, one fully engages in the public practice of using **t** only if one assents to **s**.

But taken together, **Common Creed** and **SDM** entail **Possession**, and so entail **Analyticity**. Conversely, letting **No Common Creed** be the negation of **Common Creed**, **No Common Creed** and **SDM** entail that **Possession** is false, and so that **Analyticity** is. That is, given **No Common Creed**, there is no sentence **s** such that, for some term **t** in **s**, one must assent to **s** in order to fully engage in the public practice of using **t**. Hence, given that, by **SDM**, one fully engages in the public practice of using some term only if one understands it, there is no sentence **s** such that, for some term **t** in **s**, one must assent to **s** in order to understand **t**. In short, **No Common Creed** is essential to the explanation of why **Analyticity** fails, given the premise that meaning is socially determined, in the sense of **SDM**. But why believe **No Common Creed**?

An initial, and arguably quite sufficient, answer to this question is that one should believe **No Common Creed** because it is part of our evidence that **Analyticity** is false, where this is shown by the possibility of cases like **Simon**, **Vann**, and **Sarah**. Since **SDM** and **No Common Creed** together entail that **Analyticity** is false, we have good evidence for their conjunction. Equally, given that **SDM** and **Common Creed** entail **Analyticity**, then, if it is part of our evidence that **Analyticity** is false, we have decisive evidence that **No Common Creed** is true if **SDM** is. In effect, **No Common Creed** might be motivated on abductive grounds, as one component of a natural package of claims that together predict what is independently part of our evidence. Nor, it should be emphasised, would this argument for **No Common Creed** illicitly beg the question against **Analyticity**. In particular, any such objection would misrepresent the dialectical situation. Given that **Analyticity** is independently implausible, the residual justification for accepting it is that it is entailed by (3), which, in the absence of an alternative picture of linguistic understanding, stands as the only substantive alternative to scepticism about understanding. But if it is legitimate for the friend of (3) to insist that (3) must be true on the basis that it correctly predicts that many speakers do understand their terms, then presumably it is legitimate to insist that **SDM** and **No Common Creed** must be true on the basis that they jointly predict, what also seems to be the case, that speakers like Simon, Vann and Sarah understand 'and' and 'if'. More basically, even if it is reasonable for the friend of **Analyticity** to insist that apparent counterexamples to **Analyticity** be accompanied by a substantive alternative

picture of linguistic understanding, **SDM** and **No Common Creed** arguably *do* comprise a substantive alternative.[31]

Nevertheless, even given that one accepts **No Common Creed**, it is interesting to see whether the principle can also be motivated top-down, from more theoretical considerations; the aim might simply be to learn more about why **Analyticity** fails, on the assumption that it does. And indeed, Williamson himself seems to offer two suggestive lines of argument in defence of **No Common Creed**. Roughly, the first proposes to subsume the question of what unites members of a linguistic practice under the broader question: 'what binds together different events into the history of a single complex object, whether it be a stone, a tree, a table, a person, a society, a tradition, or a word?' (2007: 123). Correspondingly, **No Common Creed** is explained by the fact that:

> Rarely is the answer to such questions the mutual similarity of the constituents. Almost never is it some invariant feature, shared by all the constituents and somehow prior to the complex whole itself—an indivisible soul or bare particular. Rather, it is the complex interrelations of the constituents, above all, their causal interrelations (ibid).

From this perspective, **Common Creed** is claimed to amount to 'one of the crudest and least plausible answers to the question of what makes a unity out of diversity', that 'what animates a word is a soul of doctrine' (ibid). And elsewhere,

[31] Compare Williamson (2011b: 503–4) in response to Boghossian (2011: 496), who advances essentially the objection in the text, claiming that the social determination of meaning 'repeats the rejection of inferentialism but without providing a substantive alternative' (2011: 496)

Williamson claims that, given that the unification of a linguistic practice 'consists in causal interrelations of speakers, not merely in their shared monadic properties', an inferentialist picture of understanding 'has the wrong structure', since it does not 'capture the causal interrelations' (2010: 504). Indeed, an extension of this point might be that, given a picture on which the meaning of a speaker's terms is constituted by just those conditions that she herself accepts, it is not clear that speakers *ever* proceed to the point of using terms with genuinely shared meanings. Rather, so-called linguistic communities remain collections of speakers with their own idiolects; when these idiolects coincide, it is only because speakers have somehow succeeded in bringing them into perfect synchronisation.

This argument for **No Common Creed** at once has the advantage and the disadvantage of being highly general. On the one hand, it has the exciting feature of implying that central questions in the philosophy of language partly reflect broader issues in the metaphysics of composition, with **Analyticity**, in particular, partly reflecting a local scepticism about the possibility of emergent linguistic practices. However, on the other hand, it has the mildly ironic feature that, if one were to expect *any* unity to be held together by 'a soul of doctrine', one might expect a social practice to be. After all, some social practices and communities *do* impose conditions on participation and membership, so that it is only possible for one to participate in the relevant practice, or belong to the relevant community, if one satisfies specific, invariant criteria. And prima facie, if linguistic understanding really does consist in full engagement in a social

practice, it seems intelligible to wonder why there are no barriers to entry in that practice, or necessary conditions on continued participation, which take the form of sentences to which one must assent. For instance, Williamson characterises two of his speakers as not having been 'expelled' 'from the club of "every"-users' for their heterodoxy (2007: 91). But, if this kind of talk makes sense in the first place, why *not* have linguistic practices which expelled members who held highly heterodox beliefs?

In fact, this is a question to which Williamson does seem to suggest an answer, which it is worth quoting at length:

> We cannot anticipate all our disagreements in advance. What strike us today as the best candidates for analytic or conceptual truth some innovative thinker may call into question tomorrow for intelligible reasons. Even when we hold fast to our original belief, we can usually find ways of engaging rationally with the doubter. If a language imposes conditions of understanding that exclude such a doubt in advance, as it were in ignorance of its grounds, it needlessly limits its speakers' capacity to articulate and benefit from critical reflection on their ways of thinking. Such conditions are dysfunctional, and natural languages do better not to impose them. Similarly, conceptual practices do better not to restrict in advance their capacity for innovation (2007: 126).

Whereas Williamson's first argument for **No Common Creed** appealed to general characteristics of composition, the argument for **No Common Creed** suggested by this passage seems broadly *normative*, appealing to functional norms on linguistic and conceptual practices. In particular, the suggestion seems

to be that for a linguistic practice to impose conditions on understanding would be dysfunctional, since it would necessarily be for it to rule out some case of *reasoned* non-assent, given that even the most compelling of principles can be rejected on a principled basis. Officially, Williamson only identifies two ways in which the expertise of the speakers in his counterexamples is relevant: 'First, it blocks the objection that the protagonist is no expert. Second, it helps explain how the protagonist can be a fully paid-up member of the speech community, speaking the public language and interpretable in the normal way' (2020a: 228–9). But in light of this second argument for **No Common Creed**, it is tempting to think that the expertise of his speakers plays a third, indirect role in Williamson's argument against analyticity: his speakers end up constituting not only evidence for the claim that any sentence can be denied, but for the claim that, for any proposition, one could have good reason for not assenting to that proposition, and thus to make vivid the cost to the public if it imposes a common creed.

Incidentally, note that Williamson is not the only theorist who seems to suggest an explanation of **No Common Creed** in terms of the possibility that innovative thinkers might someday disagree with widely accepted beliefs. For instance, in responding to Frank Jackson's moral functionalism (1998), Stephen Yablo offers the following suggestive explanation of why the fact that speakers often do not assent to the same sentences containing evaluative terms does not entail that they do not mean the same thing by their terms:

> [W]e refuse to attach 'right' to *any* descriptive property so tightly that moral dissidents, even hypothetical ones, come out as simply misusing the term. Similarly, it seems part of what we have in mind by 'reasonable' that the door is left open to the brilliant iconoclast who gets us to see that we have all along been acting contrary to reason. To dismiss such a person as meaning something else by 'reasonable' strikes us as dogmatic (2002: 19).[32]

In the following section, I attempt to reconstruct what might be called the *normative argument* for **No Common Creed** more precisely, prescinding away from Williamson and Yablo's own formulations of it. In fact, this task is not entirely straightforward, for it cannot easily be done in a dialectically neutral way. Nevertheless, addressing the initial difficulty helps to clarify the conception of analyticity that the normative argument might be understood to target.

2.4 The Normative Argument

This section attempts to reconstruct the normative argument for **No Common Creed** slightly more precisely. As a crude initial attempt, consider the following pair of claims, which can easily be seen to jointly entail **No Common Creed**:

[32] Of course, taken literally, Yablo's suggestion cannot be right: If it is part of what we mean by 'reasonable' that we could always be wrong about what is reasonable, then someone who denies 'we could always be wrong about what is reasonable' does not mean what we do by 'reasonable'. Yet if someone who denies 'we could always be wrong about what is reasonable' does not mean what we do by 'reasonable', then we could not be wrong that reasonableness is something we could always be wrong about, so we could not always be wrong about what is reasonable, contrary to the initial supposition. Thus, it is not part of what we mean by 'reasonable' that we could always be wrong about what is reasonable. Perhaps unsurprisingly, the argument that there are no analytic sentences cannot in fact proceed from the premise that 'no sentences are analytic' is analytic.

(6) For some sentence **s**, for some term **t** in **s**, if, necessarily, one fully engages in the public practice of using **t** only if one assents to **s** then, necessarily, if one does not assent to **s**, one does not have good reason not to assent to **s**.[33]

(7) For any sentence **s**, possibly, one does not assent to **s** and one has good reason not to assent to **s**.

Roughly, (6) says that if participation in the public practice of using some term requires assenting to some sentence, then not assenting to that sentence strictly implies that one lacks good reason not to assent to it, while (7) says that no sentence satisfies this condition: for any sentence, it is possible that someone who does not assent to that sentence has good reason not to assent to it. However, it should be clear that, as the premises of an *argument* for **No Common Creed**, (6) and (7) are not very good. For, given the plausible assumption that one has good reason not to assent to a sentence only if one understands that sentence, (7) already implies that there are no analytic sentences, a consequence which was supposed to *follow* from the normative argument; while (6) reduces, on the same assumption, to the triviality that, if assent is required for having good reason not to assent, then one lacks good reason not to assent if one doesn't.

Arguably, the lesson from the failure of (6) and (7) to constitute a non-question-begging argument for **No Common Creed** is that the normative

[33] Quasi-symbolically, to clarify scope: $\exists \mathbf{s} \exists \mathbf{t}(\Box(\mathrm{E}\mathbf{t} \rightarrow \mathrm{A}\mathbf{s}) \rightarrow \Box(\neg \mathrm{A}\mathbf{s} \rightarrow \neg \mathrm{R} \neg \mathrm{A}\mathbf{s}))$.

argument cannot rely on the premise that, for any *sentence*, it is possible that someone who does not assent to it has good reason not to.[34] Rather, in place of (7), what is required is a premise closer to the following:

(8) For any meaning **m**, possibly, for some sentence **s** which semantically expresses **m**, one does not assent to **m** and one has good reason not to assent to **m**.[35]

Roughly, (8) says that, for any meaning—even if it is necessary to assent to whichever sentences actually express that meaning in order to understand them—it *could* have been the case that, for some sentence which expressed that meaning, one had good reason not to assent to that sentence. Thus, (8) provides a more neutral basis for an argument against **No Common Creed**, given that it is at least consistent with the supposition that there are analytic sentences. Indeed, in principle, (8) is consistent with the claim that, for some special class of meanings $\mathbf{m}_1$, …, $\mathbf{m}_n$, all of the sentences which *actually* express $\mathbf{m}_1$, … , $\mathbf{m}_n$ are analytic. Crucially, however, (8) is not completely dialectically neutral, for it still entails the following claim:

---

[34] What about just the premise 'for any sentence **s**, possibly, one has good reason not to assent to **s**'? To be interesting, good reason not to assent to **s** should be sufficient for some speakers to feel compelled not to assent to **s**; otherwise they are not excluded from the practice even if **s** is analytic. In any case, there remains the commitment discussed below, so long as a speaker who has good reason not to assent to **s** is *capable* of not assenting to **s** while continuing to understand the meaning that it semantically expresses.

[35] I use 'meaning', rather than 'proposition', since it is not completely clear that inference rules like (4) and (5) express propositions. In the intended sense, meanings are the worldly content of sentences. Correspondingly, valid [invalid] inference rules and true [false] propositions will both be subsumed under the category of correct [incorrect] meanings

(9) For any meaning **m**, possibly, for some sentence **s** which semantically expresses **m**, one does not assent to **s** and one understands **s**.

On the assumption that if one understands and assents to some sentence **s** which expresses **m**, then one understands and assents to **m**, what (9) says is that there is no meaning **m** for which, understanding **m** requires assenting to **m**. In short, there are no analytic *meanings*: meanings which can only be grasped given that one assents to those very meanings. Thus, (8) is not *fully* neutral with respect to **No Common Creed**, for it rules out the possibility that some linguistic practices require their members to assent to a certain sentence in order to understand it in virtue of the fact that they have no other option, given that they have a sentence which expresses that meaning at all.

However, by the same token, that the normative argument carries a commitment to (9) may indirectly serve to clarify the context for the normative argument in the first place. For, given (9), if there are analytic sentences, then it is not the case that there are analytic sentences because there *must* be analytic sentences—that is, because certain thoughts could not be expressed except by sentences which speakers could only understand if they assented to those sentences. But then, why else might languages encode analytic sentences? One suggestion, perhaps, is just this: if there are analytic sentences, then there are analytic sentences because linguistic practices themselves have good reason to exclude speakers who do not assent to certain sentences in the relevant language. Very roughly, this is analyticity understood as a matter of social convention,

rather than analyticity understood, with the inferentialist, as a matter of conceptual necessity—in other words, analyticity that ostensibly obtains in virtue of the social determination of meaning, not despite it.

In turn, the normative argument can be seen as aiming to preempt this defence of analyticity. For the sake of argument, it concedes that **Common Creed** is true if and only if it *should* be true, in some broadly functional sense of 'should':

> (10) For some sentence **s**, for some term **t** in **s**, necessarily, one fully engages in the public practice of using **t** only if one assents to **s** if and only if it should be that, for some sentence **s**, for some term **t** in **s**, necessarily, one fully engages in the public practice of using **t** only if one assents to **s**.[36]

But it argues that (10) turns out to support **No Common Creed**, rather than **Common Creed**, since linguistic practices would be dysfunctional if they excluded reasoned disagreement:

> **Anti-Dogmatism:** If it should be that, for some sentence **s**, for some term **t** in **s**, necessarily, one fully engages in the public practice of using **t** only if one assents to **s**, then, for some meaning **m**, necessarily, for any sentence **s**

[36] Quasi-symbolically, to clarify scope: ∃**s**∃**t**□(E**t** → A**s**) ↔ □$_S$∃**s**∃**t**□(E**t** → A**s**). It is not incidental that '□$_S$' scopes *over* the quantifiers, which should themselves be read as restricted to actual terms and sentences. (If some sentence is necessarily a given way, then it is arguably redundant to say that it *should* necessarily be that way. Similarly, while it is an interesting question whether, given that **s** is analytic, it is necessarily analytic, it is safer to simply assume that, while it is not necessary *that* there are analytic sentences, any analytic sentence **s** is necessarily analytic.)

which semantically expresses **m**, if one does not assent to **s**, one does not have good reason not to assent to **s**.[37]

In short, given (10), **No Common Creed** is true unless **Common Creed** *should* be true. Given **Anti-Dogmatism**, it should be that **Common Creed** is true only if, for some meaning, no one who did not assent to a sentence which expressed that meaning could have good reason not to assent. Thus, since, given (8), there is no such meaning, **No Common Creed** is true.

Since (10) represents a concession to the friend of **Common Creed**, the normative argument rests on (8) and **Anti-Dogmatism**. The last section of this chapter examines whether these are jointly plausible. In particular, I suggest that two initially attractive reasons for accepting the normative argument are not clearly compelling, especially by the lights of epistemological externalism, which takes truth to be a necessary condition on epistemic justification. Fortunately, however, weaker premises arguably suffice for the conclusion that **No Common Creed** is true, while better illuminating it.

3.5 Good Reasons, Good Reasoner

Why think that it would be dysfunctional for a linguistic practice not to allow participants not to assent to a sentence expressing a given meaning, given that

[37] Quasi-symbolically, to clarify scope: $\Box_S \exists \mathbf{s} \exists \mathbf{t} \Box(E\mathbf{t} \rightarrow A\mathbf{s}) \rightarrow \exists \mathbf{m} \Box \forall \mathbf{s}(\neg A\mathbf{s} \rightarrow \neg R \neg A\mathbf{s})$. While **Anti-Dogmatism** does not technically say that one should only be required to assent to **s** if **s** itself expresses a meaning to which one could have good reason not to assent, if (8) and **Common Creed** are true, then it's still the case that *some* sentence expresses a meaning to which one could have good reason not to assent. Since what's at issue is whether there are *any* common creeds, and so analyticities, this consequence is already strong enough to be interesting.

speakers could have good reason not to assent to some sentence expressing the same meaning? An initially attractive answer perhaps runs as follows: There is some good epistemic status that certain beliefs have. A linguistic practice is dysfunctional if it does not allow participants to have certain beliefs which, if they were able to have those beliefs, would have that status. If one has good reason not to assent to something, then one's belief that one should not assent to it has that status. Thus, given that one could have good reason not to assent to the meaning expressed by a certain sentence, it is dysfunctional for a linguistic practice to require that members assent to the sentence which it uses to express that meaning. For instance, it is possible for speakers to have good reason not to assent to conjunction-elimination, so it is possible for one's belief that one should not assent to conjunction-elimination to have this good epistemic status, so it is dysfunctional for a linguistic community only to admit members on the condition that they assent to the sentence which that linguistic community uses to express conjunction-elimination.

Let us call the relevant epistemic status *justification*. Then, if it is motivated by the foregoing line of reasoning, the normative argument essentially relies on the following assumption:

(11) For any sentence **s**, if one has good reason not to assent to **s**, then one is justified in not assenting to **s**.

Given (11), **Anti-Dogmatism** seems fairly plausible. For it is implied by the claim that it should be the case that engagement in the public practice of using some

term requires assenting to some sentence only if there is some meaning to which one could not be epistemically justified in not assenting. And it seems plausible that if one *could* be justified in not assenting to some meaning, then, if all else is equal, one should not be *required* to assent to it in order to participate in a given linguistic practice. An analogous principle for action is arguably plausible as well: if one *could* be justified in not performing some action, then, if all is equal, one should not be *required* to perform that action in order to participate in a given social practice. Just as a community which required its members to perform some action that they might be justified in not performing could reasonably be accused of dogmatism (or worse), a linguistic community which required its members to have some belief that they might be justified in not having could reasonably be accused of dogmatism. For instance, if conscientious objectors are *justified* in their refusal to be drafted, this is compelling reason to think that they should not be drafted. Similarly, if speakers who do not assent to conjunction-elimination are *justified* in not assenting to conjunction-elimination, this is compelling reason to think that they should not be required to assent to conjunction-elimination.

The problem with a version of the normative argument that is motivated by these considerations is that, while **Anti-Dogmatism** might be plausible on the assumption that (11) is true, it is far from clear that (8) is plausible on this assumption. For, given (11), (8) entails:

(12) For any meaning **m**, possibly, for some sentence **s** which semantically expresses **m**, one does not assent to **s** and one is justified in not assenting to **s**.

Now, the point of saying that a speaker has good reason not to assent to **s** is presumably not just that she has good reason to suspend belief in **s**. Rather, at least if Williamson's heterodox logicians are representative, the point is moreover that the speaker has good reason to believe that **s** does not express a correct meaning. Correspondingly, the claim that speakers should be permitted not to assent to some meaning because they may be justified in not assenting to it looks most compelling if they are not just justified in suspending belief in that meaning, but justified in believing that it is not true. But, given this reading of 'justified in not assenting to **s**', it is not clear that the failure to assent to inference rules like conjunction-elimination (or, for instance, to the proposition that vixens are female foxes) *is* possibly justified. More concretely, epistemological externalists are likely to endorse at least the following norm on belief:

(13) It ought to be that (one believes **m** only if **m** is correct).

But on the assumption that justification, in the relevant sense, corresponds to (13), speakers like Simon, Vann, and Sarah do *not* constitute good evidence that one can be justified in not assenting to modus ponens or conjunction-elimination. Consequently, holding (11) fixed, they do not constitute good evidence that one can have good *reason* not to assent to the relevant principles. For, in the relevant sense, Simon, Vann, and Sarah arguably only have what *seems* to them to be good

reason not to assent to those principles. For we might ourselves justifiably think that the reasons they have for denying (4) and (5) are inadequate for the conclusions which they draw on the basis of those reasons. Indeed, I take it that we are in a position to *know* that Simon, Vann, and Sarah are wrong—albeit less clearly Vann than Simon and Sarah.

In short, (12) fails, at least given (13). We can perhaps build on the significance of this fact further by considering a second motivation for accepting the normative argument. In particular, consider Williamson's suggestion that a language that imposes conditions on understanding 'needlessly limits its speakers' capacity to articulate and benefit from critical reflection on their ways of thinking'. As before, there is an analogy with action: an agent who irrevocably commits to performing a certain action thereby limits her capacity to modify her course of behaviour in the future. Thus, certainly, if all is equal, one should not irrevocably commit to performing certain actions. Similarly, it seems plausible that, if all is equal, then a linguistic community should not circumscribe the sentences to which its members can assent.

However, at least in the case of action, it is often the case that not all else *is* equal. For there is an important difference between an agent who commits to performing some action in the knowledge that she ought to perform that action, and one who has good but misleading evidence that she should perform that action. The latter agent certainly ought not to have incurred the commitment that she did. But in principle, the former agent might have been justified in

committing to performing that action, since she might have been aware that if she did *not* commit to performing it, she would have failed to do so. Indeed, she might have known that she would fail to perform it precisely in virtue of the fact that her friends and family would have provided her with putative reason after putative reason that she should not perform it, until, unable to answer their sceptical challenges, she eventually lost her belief that she should perform it.

In other words, further critical reflection does not always conduce to more justified action; it may have the opposite result. But then, given that speakers like Simon, Vann, and Sarah are arguably pressing sceptical challenges to beliefs which *are* justified, rather than themselves justified in their beliefs, it seems rather optimistic to be altogether confident that the epistemic value of their objections themselves could not turn out to be similarly ambiguous. Of course, Simon, Vann and, Sarah themselves might be scrupulous and honest philosophers, always reasonable in the face of objections. But what about their followers? The consequences of bad ideas often ramify, given that one reasons from those incorrect meanings to other incorrect ones; correspondingly, a sub-community of speakers who falsely believe theories that invalidate (4) and (5) may not be the most helpful members of their broader linguistic community.

So far, we have considered two motivations for the normative argument. According to the first, linguistic practices should not require assent to sentences which express meanings to which it is possible that one is justified in not assenting. According to the second, linguistic practices should not limit critical

reflection on the sentences accepted in that practice. Neither was clearly compelling. For the first motivation to fully vindicate the normative argument, it would have to be the case that it is possible to be justified in not assenting to any meaning, but it is not. For the second motivation to fully vindicate the normative argument, it would be preferable if we were not familiar with cases in which further critical reflection had less than salutary consequences, but we are.

Perhaps, however, there is a simpler justification for **No Common Creed** in the vicinity, which does not require it to be the case that one could have *good reason* not to assent to any meaning. For as Williamson himself emphasises, there is an important distinction between having *good reason* for a belief, and being a *good reasoner* who comes to have that belief.[38] Roughly, in the intended sense, one has good reason for a belief just in case it is supported by one's evidence, whereas one is a good reasoner who comes to have that belief just in case one comes to have it in virtue of employing cognitive dispositions that produce justified beliefs—in particular, let us assume, which produce knowledge.

With this distinction in mind, the most striking feature of Williamson's cases is not that Simon, Vann and Sarah could have good reason not to assent to modus ponens and conjunction-elimination. Rather, it is that even reasoners as good as Simon, Vann, and Sarah could come to believe that they should not assent to principles like (4) and (5), which are liable to seem obviously true. But

[38] See Williamson (2017), (forthcoming-b), (forthcoming-c), and (forthcoming-d) for further development and justification of the distinction.

then, if Simon, Vann, and Sarah ceased to understand the terms 'if' and 'and' because they failed to assent to (4) and (5), because they were thereby expelled from their linguistic practice, one moral should seem to be that almost *anybody* could be expelled, and so lose their understanding of 'if' and 'and'. In effect, Williamson's speakers end up constituting a particularly vivid reminder of the constant possibility of error. Not even the best cognitive dispositions are liable to keep one perfectly safe, in part since epistemological holism suggests that new evidence about one subject matter can induce changes almost anywhere in one's web of beliefs. While (8) might be doubtful, (14) is true:

(14) For any meaning **m**, possibly, for some sentence **s** which semantically expresses **m**, one does not assent to **m** and one is a good reasoner.

And (14) seems to provide the basis for a compelling case for **No Common Creed** by itself, despite being weaker than (8).[39] It suggests that the dysfunction of a linguistic practice that excludes speakers if they fail to assent to some analytic sentence may be akin the dysfunction of a library full of difficult books which expels its members if they misread one of them. Since any of its books can be misread, any of its members can be expelled. Correspondingly, the other members of that library are deprived of the knowledge of the member who is expelled (not just of the chance to come to understand his misreading), while the member who is expelled—even if he retains his body of past knowledge—is

[39] Together with **Anti-Dogmatism** suitably reformulated.

deprived of future engagement with the library's accumulated resources, in virtue of having committed a single infraction.

In an encyclopaedia entry on reference, Williamson writes: '[f]or realists, the pre-condition [on thinking about something] is minimal. It permits us no end of ignorance and error' (2005: 789). But the point of having cognitive, conceptual and—if the speculations of this section are along the right lines—social practices which impose a minimal pre-condition on understanding is of course not to permit the ignorance and error. Rather, the point is that, without the possibility of error and of ignorance, there is no possibility of knowledge, at least not for creatures whose cognitive processes are less than perfectly reliable.[40] Sometimes, the error and ignorance are moreover valuable, as when they conduce to critical reflection, which in turn generates knowledge. If so, that is a further benefit.

The last chapter of this thesis argues that fictionalism, a popular semantic theory in the philosophy of mathematics, is false—more specifically, that it is false even if some version of mathematical scepticism is correct. The interest of the conclusion is primarily in suggesting that the logical space of responses to what Williamson (2007) calls judgement scepticism—according to which, very roughly, our judgements are no more likely to be true than false—is more restricted than it might initially have seemed.

[40] See, of course, Williamson (2000: §4) on anti-luminosity.

4 Fictionalism as Judgement Scepticism

4.1 Introduction

One is less than two. If one is less than two, then the numbers one and two exist. If the numbers one and two exist, then numbers exist. So numbers exist.

Call this the *Moorean argument* for the existence of numbers. It presents an obvious problem for philosophers who deny that anything is a number, given that it is valid and that its premises seem true. Similarly, it presents an obvious problem for philosophers who deny that it is known whether anything is a number, given that it is valid and that its premises seem known. Let it be sufficient for being a *mathematical sceptic* that one denies that it is known that numbers exist.

In broad outline, there are two general ways that mathematical sceptics have responded to the Moorean argument for the existence of numbers. First, some have simply denied that it *is* known that one is less than two, on any admissible interpretation of the English sentence 'one is less than two'. Call such mathematical sceptics *radical*. Radical mathematical scepticism may be implausible, but it is clear enough. By contrast, *non*-radical mathematical scepticism encompasses a complex and constantly shifting set of positions. For instance, some non-radical mathematical sceptics deny that 'one is less than two' logically entails 'numbers exist', even holding context fixed. Call such mathematical sceptics *non-classical*. Alternatively, some mathematical sceptics

accept that the Moorean argument is sound, holding context fixed, but think that there are multiple admissible interpretations of its constituent terms, such that 'one is less than two' is only known on some, but only inconsistent with (the best version of) mathematical scepticism on others. Call such mathematical sceptics *contextualist*. And lastly, some mathematical sceptics simply accept that numbers exist, on every admissible interpretation of 'numbers exist', but deny that this fact is inconsistent with (the best version of) mathematical scepticism. Call such mathematical sceptics *anti-radical*, to mark the distance between their view and radical mathematical scepticism.

In this essay, I will be concerned with hermeneutic fictionalism (henceforth simply *fictionalism*), a semantic theory popular among non-radical mathematical sceptics. Almost all of the details of fictionalism are contested. As I will discuss below, on different characterisations, it may support different versions of non-radical mathematical scepticism. However, perhaps the dominant strain of fictionalism is defined by the fact that it draws a distinction between two *kinds* of content that are possessed by sentences containing terms which apparently quantify over numbers. Call such sentences *number-sentences*. Then fictionalists draw a distinction between what I will call the *semantic content* and the *assertive content* of number-sentences.[41] Roughly, let the semantic content of a sentence at

[41] In this broad sense of 'fictionalism', views that take assertions of sentences in a given domain to be implicitly modified by an 'according to the fiction' operator are merely one way of developing fictionalism. For instance, Yablo (2001) develops a popular version of fictionalism called 'figuralism', which instead proceeds by analogy to figurative language. Recent versions of fictionalism, in the intended sense, also include Yablo's 'presuppositionalism' (2014) and Daniel Hoek's closely related theory of conversational exculpature (2018). The term 'assertive content' is

a given context of utterance be the proposition which is *semantically expressed* by that sentence at that context; and let the assertive content of a sentence be the proposition that speakers *assert* or *communicate* in uttering that sentence at that context. Then the characteristic fictionalist claim is that, at least at most ordinary contexts, number-sentences have *distinct* semantic and assertive contents. In effect, according to fictionalism, what is semantically expressed by number-sentences systematically differs from what speakers use number-sentences to assert.

With the distinction between semantic and assertive content in hand, fictionalists subsequently aim to explain away the appearance that the Moorean argument is inconsistent with mathematical scepticism, without having to *deny*, with the radical mathematical sceptic, that one is less than two. Here we face a choice point in the development of fictionalism. On the version of fictionalism that will be primarily under discussion here, fictionalism will be understood as a version of *anti-radical* mathematical scepticism.[42] Thus, we shall understand fictionalists as responding to the Moorean argument in the following way. First,

borrowed from Yablo (2017a). Although Yablo's own contrast is between assertive content and 'compositional content', this is unhelpful, for it is not clear in virtue of what assertive content itself is *not* compositionally determined. Incidentally, a similar distinction is sometimes drawn in the literature on loose talk; for instance, Carter (forthcoming) distinguishes between 'literal' and 'communicated' content of certain sentences. (Carter notes that, technically, 'it is particular utterances of sentences, rather than the sentence-types themselves which can be attributed communicated content' (2019b: 23n1). However, while it is unclear whether this is intended to mark a contrast between assertive and semantic content, there is no such contrast. Given the presence of indexical expressions in some sentence-types, technically, it is only particular utterances of sentences which can be attributed semantic content.)

[42] Alternative ways of developing fictionalism will be discussed in the following section.

fictionalists offer the following (non-equivalent) metalinguistic reformulation of the Moorean argument:

(1) The content of 'one is less than two' is true.

(2) If the content of 'one is less than two' is true, then the content of 'the numbers one and two exist' is true.

(3) If the content of 'the numbers one and two exist' is true, then the content of 'numbers exist' is true.

Therefore,

(4) The content of 'numbers exist' is true.

Fictionalists then argue as follows. Given a uniform interpretation of 'content', this argument is indeed valid, and (2) and (3) are true on both candidate interpretations of 'content'. However, it is only the case that (1) is known if 'content' as it occurs in (1) means 'assertive content', while it is only the case that (4) is inconsistent with mathematical scepticism if 'content' as it occurs in (4) means 'semantic content'. In turn, this equivocation in this metalinguistic variant on the Moorean argument tracks a conflation in our judgements about the Moorean argument itself. We judge it to be sound because the argument it asserts *is* sound, but judge its conclusion inconsistent with mathematical scepticism because the semantic content of 'numbers exist' *is* inconsistent with mathematical scepticism—and, crucially, we do not realise that number-sentences *don't* assert their semantic contents. Although fictionalism thus developed clearly has

important similarities to contextualist mathematical scepticism, unlike the contextualist, the fictionalist does not claim that there is any ambiguity in the terms which constitute the Moorean argument itself. One way of making this difference more precise will be suggested in the next section; for now, the point is that the fictionalism *accepts* that one is less than two, so that numbers exist.

For those inclined to mathematical scepticism, fictionalism seems to offer a number of theoretical benefits. First, and perhaps most importantly, it does not entail that mathematical practice is in error. If fictionalism is true, our practice of using number-sentences does not require revision, because number-sentences make true assertions. Second, and relatedly, it is consistent with norms on assertion which imply that one should only assert a proposition if that proposition is true. Numerical claims are not merely 'useful fictions', in some purely pragmatic sense of 'useful'. Indeed, for all that has been said, it may be that a number-sentence is assertible only when its content is *known*—so long as 'content' means 'assertive content'. And third, fictionalism purports to explain why the Moorean argument *seems* to be a compelling argument against mathematical scepticism. It seems compelling because (i) it is sound and (ii) it is not pre-theoretically obvious that number-sentences do not assert their semantic contents. At the same time, fictionalism seems to come at minimal theoretical cost. Namely, its defining theoretical commitment—to the existence of *two* contents systematically associated with number-sentences—might seem relatively harmless, especially if the distinction can indeed be put to theoretical work.

However, I wish to argue that fictionalism's distinction between semantic and assertive content is theoretically unmotivated. The rest of this chapter proceeds as follows. In §3.2, I explain the treatment of fictionalists as *anti-radical* mathematical sceptics, specifically, rather than non-classical or contextualist sceptics. In §3.3, so that the discussion is less schematic, I introduce a counterfactual semantics for number-sentences, and defend that semantics against the *ad hominem* aspect of an objection to it due to Timothy Williamson. In §3.4, I suggest that the counterfactual semantics for number-sentences can be simplified; and note that, particularly in its simplified form, it can naturally be used by non-fictionalist anti-radical mathematical sceptics as well. In §4.5, I argue that fictionalism does not better explain the appeal of the Moorean argument as an argument against mathematical scepticism than non-fictionalist anti-radical mathematical scepticism, and that two residual motivations that the fictionalist might have for nevertheless drawing a distinction between the semantic content and assertive content of number-sentences are unsuccessful.

3.2 Anti-radical fictionalism

In §3.1, I interpreted fictionalism as a version of *anti-radical* mathematical scepticism. In this section, I clarify the reasons for this choice. While fictionalism perhaps could be interpreted as a version of non-classical or contextualist mathematical scepticism, it constitutes a less interesting hypothesis on either of these interpretations. More specifically, on these alternative characterisations, the motivation for the fictionalist's characteristic distinction between semantic and

assertive content becomes quite obscure. I take this to constitute defeasible reason to reject these characterisations of fictionalism, since it seems worthwhile to try to get as clear as possible about what *might* be distinctive about the fictionalist approach. In fact, I will argue in §3.5 that, even as fictionalism is characterised in §3.1, the fictionalist is unable to explain what theoretical work is done by her distinction between semantic and assertive content as such; the considerations in this section may thus perhaps be seen as prefiguring this conclusion. However, as §3.5 also attempts to bring out, this chapter's version of fictionalism may conduce to a cleaner diagnosis of why one might be tempted by the fictionalist's distinction between semantic and assertive content, even granted that it is not easy to identify exactly what that distinction explains.

4.2.1 Non-classical fictionalism?

According to non-classical mathematical scepticism, the Moorean argument is not sound, because 'one is less than two' does *not* logically entail 'numbers exist'. Some discussions of fictionalism give the initial appearance of claiming precisely this. For instance, in an influential exposition of fictionalism, Stephen Yablo characterises fictionalism as a response to the following 'predicament': 'One, we find ourselves uttering sentences that seem on the face of it to be committed to so-and-so's—sentences that could not be true unless so-and-so's existed. But, two, we do not *believe* that so-and-so's exist' (2001: 72; emphasis in original). For instance, it seems that the sentence 'one is less than two' could not be true unless numbers existed; but we do not seem to (fully) believe that numbers exist. As

Yablo presents fictionalism, it constitutes a reply to this 'predicament' which requires neither paraphrasing away the relevant commitment; nor rejecting the apparently committed sentences; nor accepting the commitment (ibid.). In fact, the apparently committed sentences never carried the relevant commitment at all. Or perhaps better, since it is presumably *speakers* who incur commitments, our acceptance of, for instance, 'one is less than two' never committed us to the putatively problematic consequence that numbers exist.

Clearly, at least the *spirit* of Yablo's characterisation of fictionalism is consistent with anti-radical mathematical scepticism—in accepting that one is less than two, we do not accept anything that is inconsistent with (the best version of) mathematical scepticism. However, note that Yablo seems to characterise the problematic commitment as being simply that so-and-so's exist; and gives no indication that he is using 'exists' in some special sense. But then, it seems that what fictionalism denies is that 'numbers exist' is a logical consequence of 'one is less than two', since presumably one is committed to the logical consequences of one's beliefs. Thus, fictionalism does not support anti-radical mathematical scepticism after all. Instead, the fictionalist seems to be a non-classical mathematical sceptic: she neither denies that one is less than two nor accepts that numbers exist, but rather regards the inference from the former to the latter as invalid.

I will shortly discuss one way that the fictionalist might nuance this claim (namely, by appealing to context-sensitivity in the assertive content of number-

sentences). For the moment, however, let it suffice to note that, given this interpretation of fictionalism, the central work in defusing the Moorean argument seems to be done merely by rejecting the classically valid inference from 'one is less than two' to 'numbers exist'. But if *that* is the strategy, then it is obscure what function the fictionalist's distinction between semantic and assertive content serves. That is, why distinguish between what number-sentences semantically express and what speakers use them to assert? Of course, the rhetorical force of this question is presumably acknowledged by most fictionalists, as evidenced by their use of operators like 'really' and 'literally'. For when stating their view more carefully, fictionalists (including Yablo himself) are much more likely to say that 'one is less than two' does not entail that it is *really* the case that numbers exist or that is *literally* the case that numbers exist, than that it does not logically entail that numbers exist. The crucial point, then, becomes that the fictionalist cannot regard her use of these operators as somehow redundant. If the fictionalist semantic apparatus is to be interesting, then the claim that numbers do not *really* exist cannot be seen as a mere restatement of the ordinary claim that numbers do not exist.

Relatedly, for much the same reason, fictionalists should not gloss their position as being that, although one is less than two, this does not entail that what number-sentences say is *true*, in some pretheoretically clear sense of 'says'. For the fictionalist has as much reason as anyone to accept that if one is less than two, then 'one is less than two' is true, as well as that the nature of truth can be partly characterised by the disquotational principle:

(5) If a sentence says that something is so, then it is true if and only if that thing is so.[43]

Consequently, fictionalists have good reason to think that there is a sense of 'says' such that what a number-sentence says is true if and only if its assertive content is true; since if there is no such sense of 'says', then (5) will fail to explain why it is valid to infer from the fact that one is less than two to the fact that 'one is less than two' is true. Of course, fictionalists might also be expected to contend that there is *some* important sense of 'says' on which what a number-sentence says is true if and only if its semantic content is true; and thus, a corresponding sense of 'true' on which the fact that one is less than two does not entail that 'one is less than two' is true. But again, it is necessary to resist the temptation to identify the truth of a sentence's semantic content, which we might call *truth*$_S$, with truth, in some more ordinary sense of 'true'.[44]

4.2.2 Contextualist fictionalism?

I turn now to the idea that, rather than constitute a version of anti-radical mathematical scepticism, fictionalism constitutes a version of *contextualist* mathematical scepticism. Indeed, note that it is not clear that anything in my development of fictionalism thus far is inconsistent with contextualist mathematical scepticism. Thus, I shall first explain why the fictionalist might

[43] Andeljković & Williamson (1998: 215).

[44] In light of the preceding two points, let me note that, when I use number-sentences in this essay, I intend to assert their assertive content. Similarly, in calling a sentence 'true', I intended to say that its assertive content is true.

wish to distinguish her view from the contextualist's, before indicating how the distinction between anti-radical and contextualist fictionalism might be implemented.

Given contextualist fictionalism, it is easy to give a true reading to Yablo's expository claim that speakers can be justified in accepting that one is less than two, without thereby being committed to believing that numbers exist.[45] Namely, there simply occurs a context change halfway through the sentence, so that 'numbers exist' asserts its semantic content, while 'one is less than two' does not. On this version of fictionalism, speakers can, and perhaps often do, admissibly assert the semantic content of some number-sentences. Correspondingly, the conclusion of the Moorean argument may *not* be consistent with mathematical scepticism, given that 'numbers exist' may be appropriately used to assert its semantic content, and that its semantic content is *ex hypothesi* inconsistent with mathematical scepticism. Presumably, on the most natural way of continuing to develop this view, the Moorean argument is valid, so long as the context of utterance is held fixed. However, provided that context shifts can occur surreptitiously, it may illicitly be used to assert an argument which is not in fact sound; it is in virtue of exploiting this context-sensitivity that platonists (and radical mathematical sceptics) make their view seem compelling.

The general approach is of course familiar. For present purposes, the question is why, given a distinction between different contents that speakers may

[45] This reconstruction seems to fit especially well with Yablo (2014: 486–7).

use number-sentences to assert, it is necessary to posit a *further* distinction, as the fictionalist does, between two kinds of content that number-sentences have. In particular, we can assume that contextualist fictionalists do not think that, when number-sentences are used to assert their semantic content, this is an incorrect usage—as with other applications of contextualist ideology, the idea is that, given the actual meaning of number-sentences, speakers may *correctly* use number-sentences to assert different propositions relative to different contexts of utterance. But then, it is again obscure what theoretical work is done by the fictionalist's special classification of *one* of the propositions that speakers can correctly use a given number-sentence to assert as that sentence's semantic content. As Cian Dorr notes about theorists who maintain that belief-ascriptions have a unique semantic content, but may be used to assert more than one proposition, '[t]hey will need some theory about the shifting but systematic standards of assertability that govern this practice. I suspect that this theory will end up isomorphic in relevant respects to a contextualist semantics' (2011: 947).

Perhaps this challenge can be met, or fails to be decisive. Nevertheless, I think it's worth exploring what at least *seems* to be a more distinctive view—that the distinction between assertive and semantic content can do explanatory work as such, independently of shifting standards of assertability. Consequently, let us add to our version of fictionalism a commitment to the claim that it is *inconsistent* with the ordinary meaning of number-sentences that they be correctly used to assert their semantic contents. In a different, influential exposition of fictionalism, Yablo writes '[l]eave out the "really" and there's no philosophy; the ordinary

judgment that there exists a city called Chicago stands unopposed' (1998: 258). Prima facie, transposed to the case of numbers, this is anti-radical mathematical scepticism. The ordinary judgement that numbers exist is correct; but through the use of operators like 'really', we can access the semantic content of 'numbers exist', those truth we might *not* be in a position to know.

4.3 The counterfactual semantics and mathematical necessity

So far, the presentation of fictionalism has been highly general. Nothing has been said about what, according to fictionalism, the assertive content of number-sentences *is*. But if fictionalism is to be plausible, it requires an account of what number-sentences in fact assert, if not what mathematical realists (and, for that matter, radical mathematical sceptics) take them to assert.[46] Taking up this question in slightly more detail will eventually allow us to state an apparently natural alternative to fictionalism, at least from the perspective of the anti-radical mathematical sceptic, and to compare that alternative to fictionalism in a relatively concrete way. Thus, in this and the following section, I introduce one popular account of the assertive content of number-sentences, and discuss some notable problems that arise for it.

Although there are a very large number of such accounts on offer, to my knowledge, the simplest and most intuitive extant version of fictionalism takes

[46] Compare Williamson (2020: §14.7). This is why Matti Eklund's 'indifferentism' (2007), according to which speakers are indifferent to the 'ontologically committing aspects' of certain sentences' literal content is not an 'alternative' to fictionalism, but rather a prolegomenon to fictionalism.

the assertive content of an arbitrary number-sentence to be the semantic content of some corresponding counterfactual conditional.[47] For instance, on a simple version of what may be called the *counterfactual semantics*, the assertive content of 'one is less than two' is identical to the semantic content of 'if there were numbers, then two plus two would equal four'; the assertive content of 'there are numbers' is identical to the semantic content of 'if there were numbers, there would be numbers'; and so on. In the next section, I will discuss one reason, which I will suggest is not compelling, to think that this simple version of the counterfactual semantics is *too* minimal. But first, there is a more pressing problem. Although the counterfactual semantics is attractive not least for relying on ideology that is relatively familiar—the counterfactual conditional, rather than some special 'according to the fiction', 'true-in-reality', or 'fundamentally' operator[48]—it is also, for the same reason, especially vulnerable to the objection that pure arithmetical claims seem to be *necessary*.

More specifically, Timothy Williamson (2017c; 2020) has recently developed at length an objection to the counterfactual semantics from the following two attractive premises, where 'necessary' expresses metaphysical necessity:

(6) If it is necessary that not-A, then, for all C: if it were that A, it would be that C.

[47] This approach is developed most clearly in Dorr (2005: §3) and (2008: §2). (The name 'counterfactual semantics' below is also due to Dorr.) But see also, for instance, Horgan (1984) and Putnam (1987).

[48] Of course, especially given Lewis (1978), that 'according to the fiction' operators represent an alternative to the use of counterfactual conditionals is far from clear.

(7) If numbers do not exist, then it is necessary that numbers do not exist.

In short, all counterpossibles are true and if arithmetic is false, it is necessarily false. The reason that this presents a dire objection to the counterfactual semantics is that, if all counterpossibles are true, then the necessary falsity of arithmetic threatens to trivialise the truth of number-sentences. Presumably, the anti-radical mathematical sceptic does *not* think that 'two plus two equals five' has true assertive content, for it is a trivial mathematical fact that two plus two does *not* equal five; given that she does not regard mathematical discourse as in need of revision, predicting that two plus two equals five would be no better by the lights of the anti-radical mathematical sceptic than predicting that two plus two does not equal four. Yet if the semantic content of 'there are numbers' is metaphysically impossible, then the assertive content of 'two plus two equals five' is identical to the semantic content of some counterpossible, so given that all counterpossibles are true, the counterfactual semantics predicts that two plus two equals five!

What a philosopher who is *not* a fictionalist should think about this objection to the counterfactual semantics is a question that is beyond the scope of this chapter, since the concern here is primarily to argue that one should not be a fictionalist *even if* one is an anti-radical mathematical sceptic. Thus, the question for us is what one should say about this objection if one *is* a fictionalist. And in fact, whether or not it constitutes a plausible response simpliciter, fictionalists

arguably have a ready-made reply to Williamson's objection.[49] In particular, fictionalists might offer the following (non-equivalent) metalinguistic reformulation of (6) and (7):

(8) If it is necessary that the content of 'A' is false, then, for all 'C': 'if it were that A, it would be that C' is true.

(9) If the content of 'numbers exist' is false, then it is necessary that the content of 'numbers exist' is false.

And as before, the fictionalist will think that 'content' in (8) and (9) can be precisified as either meaning 'assertive content' or as meaning 'semantic content'. Now, let us assume that fictionalists accept (8) on both disambiguations. Williamson, among others, has defended the truth of all counterpossibles at length; and, perhaps more importantly, it is not clear that fictionalists have any independent reason at all to deny either version of (8). Nevertheless, the fictionalist potentially *can* insist that (9) is only true given that 'content' means 'assertive content', but only implies that all number-sentences are true given that 'content' means 'semantic content'. Further, the fact that (9) is true given that 'content' means 'assertive content' again explains the temptation to judge that (9) is true given that 'content' means 'semantic content'. Slightly less schematically, let the *modal rigidity thesis* be that the *semantic* content of 'there are numbers' is either necessarily true or necessarily false. Then fictionalists *deny* the modal

[49] This reply is developed at greater length in Yablo (2002b), and noted by Hoek (ms); I just sketch its outline here.

rigidity thesis, but of course accept that the *assertive* content of 'there are numbers' is necessarily true or necessarily false. After all, given the counterfactual semantics, the assertive content of 'there are numbers' is the semantic content of (10):

(10) If there were numbers, then there would be numbers.

And, given that (10) is a harmless logical truth, its semantic content is necessarily true. Further, since (10) is necessarily true$_S$,[50] then the assertive content of (7) is true by the very fact that any conditionals with a necessarily false antecedent is true, since its assertive content is the semantic content of (11):

(11) If it is not the case that (if there were numbers, then there would be numbers), then it is necessarily not the case that (if there were numbers, then there would be numbers).

Incidentally, in objecting to the counterfactual semantics, Williamson partly justifies (7) by the remark that 'nominalists typically regard Platonism as not just *false* but *metaphysically impossible*' (2020: §14.7; italics in original). In context, the suggestion seems to be that, if all counterpossibles are true, then, given that the same philosophers who might accept the counterfactual semantics also accept the modal rigidity thesis, the counterfactual semantics fails to capture the assertive content of number-sentences by their own lights. However, it is not clear that this is right. For instance, Dorr, who explicitly defends the counterfactual semantics,

[50] Recall that S is true$_S$ just in case S's semantic content is true.

and Yablo, who was once described as 'the arch-hermeneutic fictionalist about abstract objects' (Stanley 2001: 38), both express at best ambivalence about it.[51] Relatedly, while Williamson (2017c: 200) is correct to note that Dorr assumes that at least some counterfactuals with metaphysically impossible antecedents are true (as indeed does Yablo), this observation somewhat misrepresents the dialectical role of this assumption in Dorr and Yablo's discussions. In both cases, the move to discussing modalities weaker than metaphysical possibility occurs centrally to shift the weight of the argument off the claim that the modally rigidity thesis is false.[52] Thus, it does not represent a clear commitment to the modal rigidity thesis, especially not if all counterpossibles *are* true.

Now, none of this is to say that fictionalists are *justified* in rejecting the modal rigidity thesis; from a more neutral perspective, it clearly seems to constitute a costly commitment. Indeed, the next section notes just how strong the claim is to which the mathematical sceptic is committed. Let '(there are numbers)$_S$' express the semantic content of 'there are numbers'. Then, in order to account for the truth of claims of applied arithmetic, the fictionalist must think not just that it is contingent whether (there are numbers)$_S$, but that, in some sense,

[51] See Dorr (2010: 142–3) and Yablo (2002b), (2012), and (2017c: 222). Hartry Field (1989) is another obvious case of a nominalist who denies the modal rigidity thesis, in some form. Although Williamson (2017c) notes that Field treats it as *logically* contingent whether 'there are numbers' is true$_S$, rather than *metaphysically* contingent, it is not clear how far this merely reflects a difference in theoretical apparatus. At any rate, much subsequent discussion (eg Wright & Hale 1992 and Colvyan 2003) seems to treat Field as making a much more controversial claim than that 'there are numbers' could be false in the same purely logical sense that 'Socrates is Plato' could be true.
[52] See Dorr (2010: 143–4) and Yablo (2012: 1013).

whether (there are numbers)$_S$ is independent of how anything else is in the world.[53]

4.4 Applied arithmetic and the modal semantics

Anti-radical mathematical sceptics are typically inclined not just to defend the truth of claims of pure arithmetic as consistent with mathematical scepticism, but also to defend the truth of claims of applied arithmetic as consistent with mathematical scepticism. For instance, we can assume that their explanation of how 'one is less than two' could be true consistently with mathematical scepticism should also explain how 'the number of planets is eight' could be true consistently with mathematical scepticism. Given the simple counterfactual semantics introduced in §3.3, the natural idea would that the assertive content of (12) is identical to the semantic content of (13):

(12) The number of planets is eight.

(13) If there were numbers, then the number of planets would be eight.

Given that (13) is true$_S$, (12) has true assertive content. However, this explication of the assertive content of claims of applied arithmetic raises both a metaphysical and a more narrowly semantic worry, on the further assumption that (13) has true assertive content if and only if there are eight planets. For then, (13) must also be true$_S$ if and only if there are eight planets. Consequently, the number of

[53] Williamson has elsewhere indicated that he finds this implausible (2016b: 575); indeed, perhaps he refrained from imputing it to nominalists out of charity.

planets must, in some sense, be held fixed when evaluating the truths of (13); otherwise (13) could be falses even though there actually are eight planets because *if* (there are numbers)s were true, then there would *not* be eight planets. Yet if (13) were falses even though there actually are eight planets, then it would not accurately capture the assertive content of (12), given that (12) has false assertive content only if there are not actually eight planets.

The metaphysical worry presented by the fact that the number of planets must be held fixed in evaluating the truths of (13) is simply that, at a minimum, this seems to imply that it is *possible* that (there are numbers)s and that the number of planets is the same as it actually is. Further, given that 'the number of planets is eight' is an arbitrary example of the kind of claim whose truth-conditions the counterfactual semantics is intended to capture, the more general commitment seems in effect to be to the metaphysically compossibility of the truth of (there are numbers)s with any non-numerical state of affairs.[54] Even for those potentially willing to deny the modal rigidity thesis, this degree of modal independence between the numerical realm and the rest of reality may still seem rather implausible, a projection of conceptual contingency onto metaphysical reality. Indeed, the postulated modal independence of the numerical from the non-numerical threatens to undermine one of the central advantages claimed for the counterfactual semantics, namely that it employs independently understood

[54] 'Non-numerical' is vague, but better than 'concrete', for a reason I am about to discuss. Note that fictionalists widely accept that fictionalist semantics incurs this commitment (see, eg, Yablo 2012 and Hoek 2018).

modal ideology. For, as Wright and Hale observe in reply to Field, the idea that there are no 'absolutely insular contingencies'—that is, contingencies which turn on nothing else and on which nothing else turn—'seems to go deep enough in our ordinary thinking to ensure that it will be very difficult to vindicate the conception of mathematical objects as strongly contingent merely by description of anything we already find intuitive' (1992: 134).

As we shall see in §3.5, fictionalism may seem to offer a solution to this problem unavailable to non-fictionalist anti-radical mathematical sceptics. For the time being, however, let us turn to a more narrowly semantic objection to the simple counterfactual semantics. As already noted, truths about, for instance, the number of planets must be 'held fixed' in evaluating these counterfactual conditionals—that is, in terms of possible worlds, 'if there were numbers, there would be eight planets' should not be false$_S$ because the relevant worlds at which its antecedent is true$_S$ are worlds at which there are more or fewer than eight planets. Yet at the same time, on the popular Lewis-Stalnaker semantics for counterfactual conditionals, for an arbitrary counterfactual 'if it were that A, it would be that C' to be true, it must be that the *closest* worlds at which A is true are worlds at which C is true. Thus, given the Lewis-Stalnaker semantics, it must be that the *closest* worlds at which (there are numbers)$_S$ is true are also worlds at which there is the same number of planets. However, as Yablo (2017b: 121) notes, this seems to be an even stronger assumption than that there are *some* worlds at which (there are numbers)$_S$ and the number of planets is the same as it actually is. To the extent that it is stronger, Yablo suggests that it introduces an unwanted

element of 'risk' into the counterfactual semantics, for how can we be sure that the closest worlds at which (there are numbers)$_S$ is true are worlds at which there are as many planets as there actually are? Indeed, Yablo gives this 'riskiness' as one reason for rejecting the counterfactual semantics in favour of one that employs the resources of truthmaker semantics. Less drastically, Dorr (2008) builds into the antecedent of the relevant counterfactuals that 'the concrete world is just as it actually is', as does Williamson in his discussion of fictionalism. Thus, the assertive content of (12) is treated as identical to the semantic content of the more complicated sentence (14), and mutatis mutandis for other number-sentences:

(14) If there were numbers and the concrete world were just as it actually is, then the number of planets would be eight.

However, if the fictionalist is worried about this objection to the counterfactual semantics at all, then she should arguably endorse a semantics for number-sentences that is *less* complicated than even the simple counterfactual semantics. For, to the extent that the objection turns on features ostensibly possessed by natural-language counterfactual conditionals, it is arguably possible for the fictionalist to sidestep it entirely. In particular, rather than take the assertive content of number-sentences to be equivalent to their semantic content embedded under a counterfactual conditional, let us suppose that the fictionalist take their assertive content to be equivalent to their semantic content embedded

under a necessity operator '$\Box_N$', so that, for instance, the assertive content of (12) is the semantic content of:

(15) $\Box_N$(The number of planets is eight).

More precisely, let N be an accessibility relation between worlds, so that $\Box_N$S is true at a world *w* just in case S is true at every world to which *w* has N; and let N(*w*) be the set of worlds accessible from *w*. Consequently, $\Box_N$S is true$_S$ at a world *w* just in case S is true$_S$ at every world in N(*w*). Further, let it be stipulated that, for any worlds *w* and *v*, *v* is in N(*w*) only if (there are numbers)$_S$ is true at *v*. Thus, if there is some world *w* at which (there are numbers)$_S$ is false, then *w* does not have N to itself, so the T axiom for $\Box_N$ fails. Otherwise, we do not attempt to further specify the details of the accessibility relation, given that exactly which facts one thinks should be held fixed by the accessibility relation will partly depend on one's other metaphysical commitments.[55] Presumably, for instance, if there are eight planets at *w*, then *v* will be in N(*w*) only if there are eight planets at *v*; this reflects the idea that how many planets there are must be held fixed in evaluating a given number-sentence. More generally, the crucial question for the fictionalist becomes whether it is plausible for her to suppose that facts about

[55] For instance, it is conceivable that some mathematical sceptics also accept the A-theory and permanentism in the philosophy of time, e.g. as developed by Williamson (2013) and Sullivan (2012), so (i) *do* think that there really are past and future objects, but (ii) do not think that the number of past and future objects supervenes on the *concrete*. Thus, if 'the number of past people is greater than the number of present people' is not true$_S$, it is only not true$_S$ because (there are numbers)$_S$ is not true. On this view, facts about how the concrete world actually *was* and *will be* must be held fixed, not just facts about how it actually *is*.

neutral subject matters (such as how many planets there are) are held fixed by the accessibility relation; but this seems as much a question for her metasemantics and her metaphysics as for her semantics itself. To be clear, the present approach in no way helps address the metaphysical worry that the postulated modal independence does not obtain.[56] Rather, its advantage is merely that it sidesteps potentially orthogonal issues about the semantics of counterfactuals, without thereby introducing any obviously non-modal semantic ideology. For instance, note that, if there are no worlds at which (there are numbers)$_S$ is true, then N(w) is empty, so the truth of (12) and other number-sentences immediately follows.[57]

Call this the *modal semantics* for number-sentences: the assertive content of an arbitrary number-sentence S is identical to the semantic content of $\Box_N$S. In other words, $\Box_N$S semantically expresses that which S asserts. Yet a natural question seems to arise at this point, which the last section of this chapter will

[56] Indeed, note that the present semantics for number-sentences may be truth-conditionally equivalent to the simple counterfactual semantics on the assumption that counterfactuals are in fact context-sensitive strict conditionals, rather than variably strict conditionals (see von Fintel 2001 and Williamson 2020). On such a semantics, 'if there were numbers, then there would be numbers' semantically expresses $\Box$(there are numbers $\rightarrow$ there are numbers), where '$\Box$' is a contextually-restricted necessity operator and '$\rightarrow$' is the material conditional. Yet this is truth-conditionally equivalent to interpreting 'there are numbers' as restricting '$\Box$'; the effect, on the assumption that the context suffices to determine that the same other facts are held fixed, is to make '$\Box$' equivalent to '$\Box_N$'. Incidentally, Dorr himself implies that the point of including 'the concrete world is just as it actually is' in the antecedent of the relevant counterfactuals is just to make the counterfactual analysis 'more explicit and less context-dependent' (2008: 37).

[57] It may be helpful to compare '$\Box_N$' to 'would', as it occurs in sentences like 'I would buy that book'. Williamson (2020: esp. §10) compellingly argues that 'would' as it occurs in such sentences functions as a context-restricted necessity-like operator, which need not satisfy the T axiom (that I *would* buy that book need not imply that I *do* buy the book, since I might actually lack the money). Importantly, the ordinary use of 'would' arguably shows that it is quite common to use modal operators that hold fixed much about how the world actually is, even without the explicit specification of what exactly is to be held fixed.

pursue in greater detail. Namely, suppose that the anti-radical mathematical sceptic accepts that $\Box_N S$ semantically expresses that which S asserts. This commitment still does not suffice for her to be a fictionalist; to be a fictionalist, she must also think that S and $\Box_N S$ have distinct semantic content. But there at least seems to be another theoretical option available: that S asserts that which $\Box_N S$ semantically expresses because S also *semantically expresses* the same thing as $\Box_N S$. Slightly more concretely, the anti-radical mathematical sceptic might think that the existential quantifier, as it occurs in the Moorean argument, has the same semantic value as '$\Box_N\exists$', where '$\exists$' is stipulated to be a purely extensional existential quantifier; the hypothesis is hence that, despite appearances, '$\exists$' does not correctly translate the ordinary English word 'exists'.

On this proposal, which we might call *modalism*, the anti-radical mathematical sceptic's diagnosis of the Moorean argument is that philosophers have over-interpreted 'exists', wrongly taking it to mean '$\exists$', when in fact it means '$\Box_N\exists$'.[58] Crucially, given the modal semantics, there does not seem to be

[58] This version of anti-radical mathematical scepticism is due to Dorr (2002: 17). See also Dorr (2005). It is an interesting question whether one might endorse a variant of this hypothesis on which 'exists' only takes this modal meaning when it occurs in sentences which explicitly quantify over numbers; when it occurs in contexts that quantify over, say, mereological simples, it simply means '$\exists$'. This would be a version of contextualist non-radical mathematical scepticism, and I believe is at least closer to Dorr's own view. However, for simplicity, modalism should be understood as the view that the ordinary word 'exists' has a univocal, modal sense—although philosophers can introduce a *new* term 'exists' which means '$\exists$', '$\exists$' itself is not an admissible interpretation of 'exists'. (Of course, considerations from other areas of metaphysics might show that '$\Box_N\exists$' is not the best translation of 'exists', even if the modalist is correct as far as numbers are concerned. For instance, a mereological nihilist like Dorr might take 'exists' to be equivalent to '$\Box_N\exists$' composed with a *further* modal operator '$\Box_C$', such that the only accessible worlds are ones at which '$\exists x$(x is a composite object)' is true. Alternatively, the permanentist mathematical sceptic

much mystery, at the level of the compositional semantics, about how number-sentences could have the semantic content that modalists take them to have, and which fictionalists identify with their assertive content. For instance, it is not clear that one can reasonably object to modalism on the basis that the compositional semantics should not give a very unnatural interpretation to 'exists', with naturalness understood in terms of definability in some fundamental language (see Lewis 1983). While the concept expressed by '$\Box_N\exists$' is not, in this sense, quite as natural as the concept expressed by '$\exists$', it is not obvious that it is especially gruesome, either. In any case, setting these slightly nebulous considerations aside, I argue in the final section that fictionalism's diagnosis of the Moorean argument is no better than modalism's; that fictionalists are no better positioned that modalists to explain their disagreement with mathematical realists; and that, if motivated from specifically metasemantic considerations, fictionalism faces a dilemma which suggests an underlying instability .

5 Semantic content as 'an explanatorily idle theoretical construct'

According to the proposed version of fictionalism, an arbitrary number-sentence S asserts the proposition semantically expressed by the corresponding sentence $\Box_N$S, but does not semantically express that proposition. By contrast, according to modalism, the semantic content of S and $\Box_N$S are truth-conditionally

described in fn.55 might think, roughly with Williamson (2013: §1), that the ordinary term 'exists' also carries an implicit *restriction* to not contingently non-concrete objects. Further variations on these themes can be played; they do not affect the arguments below.)

equivalent. Thus, it is not necessary for the anti-radical mathematical sceptic to postulate that S does not assert its semantic content, since S's semantic content is already consistent with mathematical scepticism—which we can now concretely state to be the view that *∃x(x is a number)* is not known. The modalist may in effect be seen as presenting the following challenge to the fictionalist: Why think that, as well as the content which speakers use number-sentences to assert (often knowledgeably), there is also a *further* kind of content that number-sentences have, which is properly called their semantic content? Cian Dorr puts this objection acutely, observing that, if fictionalism is true, then:

> speakers utter context-insensitive sentences which semantically express propositions they do not believe, despite the fact that the utterances in question seem both literal and sincere by ordinary standards. Without doing philosophy of language, no one would think to classify these utterances with paradigms of non-literal speech like metaphor, exaggeration, and sarcasm, or with paradigms of insincere speech like lies and deliberate attempts to mislead. Any [such] view… faces the objection that it makes the notion of semantic content into an explanatorily idle theoretical construct (2014: 31).

Put another way: fictionalism severs 'central connections between the technical expression "semantically express" and other expressions we understand' (ibid). The rest of this section attempts to make good on Dorr's objection, by showing that fictionalism has no clear explanatory advantage over modalism: either the modalist can give the same explanation as the fictionalist, or the fictionalist's

explanation fails to show that her distinction between semantic and assertive content can be understood in independent terms.

3.5.1 Fictionalism and externalism

The first motivation for fictionalism's distinction between semantic and assertive content that we shall consider is the one advertised in §3.1: that it explains why the Moorean argument *seems* to be a compelling argument against mathematical scepticism (even though it is not). To recall our earlier exposition, the fictionalist claims that the Moorean argument seems to be compelling because subjects do not realise that its conclusion does not state its semantic content—which *is* inconsistent with mathematical scepticism. The metalinguistic reformulation of the Moorean argument, in terms of the 'content' of the premises and conclusion of the Moorean argument, aims to bring this out. In particular, (1) is known only if 'content' means assertive content, but (4) is inconsistent with mathematical scepticism only if 'content' means semantic content:

(1) The content of 'one is less than two' is true.

(4) The content of 'numbers exist' is true.

According to fictionalists, our judgements about the Moorean argument itself, which is stated in the object language, reflect the conflation of these two kinds of content; this conflation explains why we (correctly) judge the argument sound, as well as why we (incorrectly) regard its conclusion as inconsistent with mathematical scepticism. The claim is *not* that shifting standards of assertability

govern the sentences that comprise the Moorean argument. Rather, it is that the judgement that mathematical scepticism is inconsistent with the knowledge that there are numbers is simply incorrect—the correct description of the situation is that mathematical scepticism is inconsistent with *the knowledge that* $\exists x(x$ *is a number,* which is a different, albeit related, claim.

Whether some strategy along these lines is the *best* way for the non-radical mathematical sceptic to diagnose the appeal of the Moorean argument is beyond the scope of this chapter. Some alternative strategies—denying that 'one is less than two' does entail 'numbers exist'; denying that its premises and conclusion express the same proposition relative to all contexts of utterance—were indicated in §3.2. Rather, for present purposes, the question is whether an error-theory along these lines requires—or indeed is aided by—the fictionalist's distinction between the semantic and assertive content of number-sentences. And there is no good reason to think that it does.

More specifically, if the motivation for fictionalism is merely that it conduces to the possibility of explaining the appeal of the Moorean argument in terms of a conflation between two contents somehow associated with an arbitrary number-sentence S, one of which is the content that speakers assert in uttering S and one of which they are tempted to believe that they assert in uttering S, then fictionalism has clear explanatory advantage over modalism supplemented by a minimal semantic externalism. To see this, let **Barrier** be whatever reason the fictionalist has for thinking that we conflate the assertive content of S, $\mathbf{p}_{\square}$, with its

semantic content, $\mathbf{p}_{\exists}$. That is, **Barrier** is the reason that, at least prior to theoretical reflection, speakers do not realise that $\mathbf{p}_{\Box}$—rather than $\mathbf{p}_{\exists}$—is asserted by S, at least under the guise '$\mathbf{p}_{\Box}$ is asserted by S'.[59] For instance, perhaps **Barrier** is that S does not *seem* to be implicitly modal, so does not seem to assert $\mathbf{p}_{\Box}$.

Then, so long as **Barrier** is not directly stated in terms of the notion of semantic content itself, the modalist can straightforwardly mimic the fictionalist's explanation of the appeal of the Moorean argument. For she can simply let her view be that, although $\mathbf{p}_{\Box}$ *is* the semantic content of S, given that **Barrier** obtains, it is not knowable that $\mathbf{p}_{\Box}$ is the semantic content of S. In effect, she can take the distinction the fictionalist is trying to draw to simply be one between, on the one hand, what S means, and, on the other, what S cannot be known not to mean (at least so long as **Barrier** obtains). In no sense is the latter proposition the *semantic* content of S; to make this identification would be to collapse the distinction between the semantic content of a sentence and other contents that are, in some sense, epistemically possible semantic contents.

However, what if the fictionalist *does* characterise **Barrier** in terms of the notion of semantic content? For instance, suppose that **Barrier** is simply that $\mathbf{p}_{\Box}$ is not the semantic content of S, and speakers are typically not in a position to know before theoretical investigation that sentences assert propositions which are not their semantic content. On this supposition, the problem is that the

[59] Of course, speakers *can* know that $\mathbf{p}_{\Box}$ is the proposition asserted by S under the guise 'the proposition asserted by S is the proposition asserted by S'.

fictionalist's notion of semantic content seems to be defined in a tight circle. For her *motivation* for introducing a distinction between semantic and assertive content was supposed to be that the distinction could be used to explain why the Moorean argument seems inconsistent with mathematical scepticism. But, given this account of **Barrier**, that explanation turns out to be that, in virtue of the very distinction between semantic and assertive content, the Moorean argument seems inconsistent with mathematical scepticism.

The following subsections argue that two alternative motivations for fictionalism over modalism face analogous dilemmas: On the first horn, the modalist can give the same explanation as the fictionalist. On the other, the fictionalist's explanation only succeeds given that one already understands the notion of semantic content. Thus, since the question at issue is precisely whether the fictionalist's distinction between assertive and semantic content can be explained in independent terms, these alternative motivations provide no evidence that Dorr's objection can be met.

4.5.2 Fictionalism and unintelligibility

Recall Wright and Hale's objection to the perfect modal independence between the numerical realm and the non-numerical realm for the counterfactual (and modal) semantics to avoid triviality—that the idea that there are no absolutely insular contingencies of the required kind 'seems to go deep enough in our ordinary thinking to ensure that it will be very difficult to vindicate the

conception of mathematical objects as strongly contingent merely by description of anything we already find intuitive (1992: 134).

Clearly, this is quite vague. However, we can perhaps make the lurking worry somewhat more vivid by recalling that, given the modal semantics, *even* 'it is metaphysically necessary that numbers exist' is true, for, treating $\Box_N$ as a quantifier over accessible worlds, this sentence ostensibly asserts that it is metaphysically necessary that: $\exists x$(x is a number) is true at all accessible worlds, where all accessible worlds (if any) are *ex hypothesi* worlds at which $\exists x$(x is a number) is true. Given this consequence of the modal semantics, it seems natural to worry that—whether or not she is moreover a fictionalist—the anti-radical mathematical sceptic's purely existential quantifier, i.e. '$\exists$', is itself an idle theoretical posit. In effect, we might wish to similarly ask the anti-radical mathematical sceptic to explain the distinction between '$\exists$' and '$\Box_N\exists$' in independent terms. For instance, why think that the judgement that numbers are not necessary genuinely tracks some objective modality?[60]

In other words, the initial promise of the counterfactual/modal semantics was that it seemed to rely exclusively on an intensional ideology, of which speakers might be expected to have pretheoretical grasp, rather than on some

[60] Note moreover that, if one's judgement is specifically that numbers cannot be *actual*, then it is not clear that the modal semantics vindicates this judgment—as Wright and Hale note in reply to Field's nominalism, given the postulated radical contingency of numbers, the more natural conclusion would seem to be *agnosticism* (1992: 114). Indeed, Yablo's official view is that it is *indeterminate* whether there are really numbers (1998; 2014). Although Rayo (2013) is nominally a platonist, his background global relativism—according to which identity-statements are only true relative to one's conceptual scheme—induces a similar result.

more theoretical 'according to the fiction', 'in-reality', or 'fundamentally' operator. Given that, on closer inspection, it is rather less clear whether this benefit is genuine, it seems fair to ask the anti-radical mathematical sceptic in what exactly her disagreement with the mathematical *realist* consists. And here the fictionalist may seem to have an advantage over modalism. Given modalism, the content of 'numbers exist' is necessarily true, on all admissible interpretations of 'content'. Whereas, given fictionalism, the anti-radical mathematical sceptic can give a further characterisation of her disagreement with the realist: it concerns whether number-sentences *semantically express true propositions*. In short, the notion of semantic content does the theoretical work of explaining the anti-radical sceptic's disagreement with the realist.

Unfortunately, this argument for fictionalism over modalism begs exactly the question which is at issue: whether the fictionalist's notion of semantic content is itself an explanatory idle theoretical posit. Analogously to before, the fictionalist is presented with the following dilemma. On the one hand, either there *is* some explanation, in terms independent of the notion of semantic content, of how the distinction between '$\exists$' and '$\Box_N\exists$' can be understood or there is not. If there is, then the modalist can equally use it in explaining her disagreement with the mathematical realist, so the fictionalism's distinction between semantic and assertive content is theoretically redundant. Alternatively, if there is not, then the fictionalist's notion of semantic content seems to be defined in a tight circle: roughly, the semantic content of number-sentences is explained as being the content about which anti-radical mathematical sceptics

and mathematical realists disagree, while the content about which anti-radical mathematical sceptics and mathematical realists disagree is explained as being the semantic content of number-sentences.

The final subsection examines a final motivation for fictionalism over modalism, and argues that it too collapses under closer scrutiny. Although we have seen no reason to think that *either* the fictionalist or the modalist can characterise their disagreement with the realist in independent terms, we continue to bracket the potential implausibility of both views for the sake of argument.[61]

4.5.3 Fictionalism and judgement scepticism

Thus far, we have considered two *positive* arguments for fictionalism, from phenomena that fictionalism can putatively help to explain—namely, the fact that the Moorean argument seems inconsistent with mathematical scepticism; and the nature of the anti-radical mathematical sceptic's disagreement with the mathematical realist. But it is worth considering the possibility that the motivation for fictionalism is, as it were, reactionary. In particular, perhaps fictionalism is in effect motivated by some version of the thought that number-sentences just *couldn't* have implicitly modal semantic content. After all,

[61] That said, compare Kripke on the danger for 'deflationary' sceptics: 'Whenever our opponent insists on the perfect propriety of an ordinary form of expression… we can insist that if these expressions are properly understood, we agree. The danger comes when we try to give a precise formulation of exactly what it is that we *are* denying—*what* "erroneous interpretation" our opponent is placing on ordinary means of expression. It may be hard to do this without producing yet another statement that, we must admit, is *still* "perfectly all right, properly understood"' (1982: 70; emphasis in original).

mathematical discourse clearly *seems* extensional;[62] presumably mathematicians never conspired to ensure that their proofs would remain true even if there were not *actually* numbers. So how could number-sentences have ended up with semantic content that is true even if $\exists x$(x is a number) is not? [63]

If fictionalism *is* thus motivated, it would arguably be in good company (or at least, a lot of company). For consider the position that Williamson (2007) calls *judgement scepticism*. Roughly, the judgement sceptic is an anti-realist by disposition, inclined to doubt, for any relevant domain of discourse, that the sentences which comprise that domain are more likely to be true than false. In describing *naturalistic* judgement sceptics, for example, Williamson writes that '[t]hey present themselves as identifying ways in which our conceptual practices need, or may need, revision in the light of scientific advances those practices failed to anticipate' (2007: 220). But the thought that if our conceptual practices fail to anticipate a certain discovery, then they have not miraculously made provisions for that discovery is not restricted to naturalists. One might equally think that, if our conceptual practices have, in some sense, failed to anticipate certain metaphysical or logical discoveries, then they would need revision were those discoveries to be made.

---

[62] Compare Williamson: 'It is worth noting that some philosophers of mathematics interpret the language of mathematics itself as implicitly modal… Despite taking such views seriously, for present purposes I prefer not to rely on philosophical interpretations so distant from the way mathematicians explicitly talk and think' (Williamson 2016a: 488n5).

[63] Of course, one might reply to this line of thought be insisting that for there to *actually* be numbers just is for it be true that $\Box_N \exists x$(x is a number), in the ordinary sense of 'actually'. However, one should resist the temptation to take this reply too far, in light of the preceding section.

A wide range of philosophical positions conduce to judgement scepticism, including the analyticity-friendly conceptions of understanding discussed in Chapter Two. For, given a theory of understanding which predicts that there are preconditions on understanding, if one fails to satisfy those preconditions, then one is locked out of understanding. Thus, if analyticities are very widespread, then inevitably so too is ignorance—unless, perhaps, the friend of analyticity attempts to fudge the results. Conversely, a wide range of philosophical positions aim in part to neutralise judgement scepticism, including Williamson's own picture of reference-determination, discussed in Chapter One. Williamson explicitly motivates the knowledge-maximisation principle of charity as a response to judgement scepticism. Given that subjects' beliefs have some tendency to constitute knowledge, they have some tendency to be true. Thus, while our beliefs may yet be riven with error—and, if we are very unlucky, our environment might be inimical to the acquisition of much knowledge at all—there is at least some bias in the direction of truth.

Perhaps fictionalism should be understood as constituting an uneasy compromise with judgement scepticism. In one way, the fictionalist sides with the judgement sceptic. Given that mathematical practice has not anticipated the truth of mathematical scepticism, fictionalists accept that number-sentences cannot be known to be *literally* true. However, fictionalists avoid the unhappy consequence that our practice of using number-sentences needs, or may need, revision, by postulating that number-sentences have a further kind of content, which may be known even if their semantic content cannot.

Of course, the problem for fictionalism, on this interpretation, is still roughly the same. Either it *is* possible for speakers to use number-sentences to assert propositions that are true even if $\exists x$(x is a number) is false, despite the fact that the practice of using number-sentences does not seem to have been safeguarded in advance against the falsity of $\exists x$(x is a number), or it is not. If it is not possible, then it is the *radical* mathematical sceptic who is correct if $\exists x$(x is a number) is false, not the non-radical mathematical sceptic. While if it *is* possible, then the question is why the propositions asserted cannot constitute the *semantic* content of numbers-sentences. For instance, if some principle of charity is operative, then why does that principle not apply in the first instance to semantic content, rather than to assertive content? More generally, if the fictionalist can explain how number-sentences come to have modal assertive content without invoking the notion of semantic content, then it is obscure why the modalist could not employ that explanation herself, but at the level of semantic content. Alternatively, if the fictionalist's explanation is, for instance, that it is essential to the notion of semantic content, but not to the notion of assertive content, that it be tightly constrained by speakers' intentions, then the modalist seems under no obligation to accept this stipulation about semantic content.[64]

At its core, fictionalism, in the sense of this chapter, is committed to the claim that, despite the fact that number-sentences are semantically associated

[64] Modalists might also call into question whether the content that fictionalists regard as the semantic content of number-sentences satisfy the postulated transparency conditions, whatever those should turn out to be. (Indeed, in many respects, the dialectic between the modalist and fictionalist is similar to the dialectic between the epistemicist and supervaluationist.)

with propositions that are not known if mathematical scepticism is correct, speakers nevertheless often use number-sentences to assert propositions that *are* known even if mathematical scepticism is correct. This pair of commitments distinguishes fictionalism from a position like modalism, on the one hand, and from both mathematical realism and radical mathematical scepticism, on the other. Correspondingly, the modalist is right about what number-sentences say, in one sense of 'says', while the realist and radical mathematical sceptic are right about what number-sentences say, in another sense of 'says'. This may seem like an attractive compromise, to negotiate an interesting middle ground between two plausible (or not so plausible) perspectives.

Well—bad compromises often seem that way.

Bibliography

1. Andjelkovic, M. and Williamson, T. 2000. 'Truth, falsity, and borderline cases'. *Philosophical Topics*, 28: 211–244.
2. Boghossian, P. 2003. 'Blind reasoning'. *Proceedings of the Aristotelian Society*, 77: 225–248
3. Boghossian, P. 2011. 'Williamson on the *a priori* and the analytic'. *Philosophy and Phenomenological Research*, 82(2): 488–497.
4. Boghossian, P. 2012. 'Inferentialism and the epistemology of logic: reflections on Casalegno and Williamson'. *Dialectica* 66: 221–236.
5. Boghossian, P. and Williamson, T. 2020. *Debating the A Priori*. Oxford: Oxford University Press.
6. Blumberg, K. and Holguín, B. 2018. 'Ultra-liberal attitude reports'. *Philosophical Studies*, 175 (8): 2043–2062.
7. Cappelen, H. and Hawthorne, J. 2009. *Relativism and Monadic Truth*. Oxford: Oxford University Press.
8. Carter, S. 2019. 'Higher order ignorance inside the margins'. *Philosophical Studies*. 176, 1789–1806.
9. Carter, S. Forthcoming. 'The dynamics of loose talk'. *Nous*.
10. Colvyan, M. 2003. 'Conceptual contingency and abstract existence'. *Philosophical Quarterly*, 50: 87–91.
11. Dorr, C. 2002. *The Simplicity of Everything*. Princeton: PhD thesis.
12. Dorr, C. 2005. 'What we disagree about when we disagree about ontology'. In *Fictionalism in Metaphysics*, ed. M.E. Kalderon. Oxford: Clarendon Press.
13. Dorr, C. 2008. 'There are no abstract objects'. In *Contemporary Debates in Metaphysics*, eds. T. Sider, J. Hawthorne and D.W. Zimmerman. Oxford: Blackwell.
14. Dorr, C. 2010. 'Of numbers and electrons'. *Proceedings of the Aristotelian Society*, 110: 133–181.
15. Dorr, C. 2011. '*De re* a priori knowledge'. *Mind*, 120: 939–91.
16. Dorr, C. and Hawthorne, J. 2013. 'Naturalness'. In *Oxford Studies in Metaphysics Vol. 8* (eds. K. Bennett and D. Zimmerman). Oxford: Oxford University Press.
17. Dorr, C. 2014. 'Transparency and the context-sensitivity of attitude reports'. In *Empty Representations: Reference and Non-existence*, ed. M. Garcia-Carpintero and G. Martí. Oxford: Oxford University Press.
18. Eklund, M. 2007. 'Fiction, indifference, and ontology'. *Philosophy and Phenomenological Research*, 71: 557–579.
19. Field, H. 1989. *Realism, Mathematics and Modality*. Oxford: Blackwell.
20. von Fintel, K. 2001. 'Counterfactuals in a dynamic context'. In *Ken Hale: A Life in Language*, ed. M. Kenstowicz. Cambridge, MA: MIT Press.

21. García-Carpintero, M. 2018. 'The Mill–Frege theory of proper names'. *Mind*, 127: 1107–1168.
22. Gendler, T.S. and Hawthorne, J. 2005. 'The real guide to fake barns: a catalogue of gifts for your epistemic enemies'. *Philosophical Studies*, 124: 331–52.
23. Goldstein, S. and Hawthorne, J. ms. 'Counterfactual contamination'.
24. Goodman, J. and Salow, B. 2018. 'Taking a chance on KK'. *Philosophical Studies*, 175(1): 183–96.
25. Hawthorne, J. 2006. 'Epistemicism and semantic plasticity'. In his *Metaphysical Essays*. Oxford: Clarendon Press.
26. Hawthorne, J. 2007. 'Craziness and metasemantics'. *Philosophical Review*, 116(3): 427–440.
27. Hawthorne, J. and Lasonen-Aarnio, M. 2009. 'Knowledge and objective chance'. In *Williamson on Knowledge*, eds. P. Greenough and D. Pritchard. Oxford: Oxford University Press.
28. Hoek, D. 2018. 'Conversational exculpature'. *Philosophical Review*, 127: 151–196.
29. Hoek, D. ms. 'Mathematical fictionalism'.
30. Horgan, T. 1984. 'Science nominalized'. *Philosophy of Science*, 51: 529–49.
31. Hyman, J. 2014. 'The most generative factive stative attitude'. *Analysis*, 74: 561–565.
32. Jackson, F. 1998. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.
33. Kaplan, D. 1989a. 'Demonstratives'. In *Themes From Kaplan*, ed. J. Almog, J. Perry and H. Wettstein. Oxford: Oxford University Press.
34. Kaplan, D. 1989b. 'Afterthoughts'. In *Themes From Kaplan*, ed. J. Almog, J. Perry and H. Wettstein. Oxford: Oxford University Press.
35. Kaplan, D. 2012. 'An idea of Donnellan'. In *Having In Mind: The Philosophy of Keith Donnellan*, eds. J. Almog & P. Leonardi. Oxford: Oxford University Press.
36. Kearns, S and Magidor, O. 2008. 'Epistemicism about vagueness and meta-linguistic safety'. *Philosophical Perspectives*, 22: 277–304.
37. Kratzer, A. 1981. 'The notional category of modality'. In *Words, Worlds, and Contexts: New Approaches in Word Semantics*, eds. H. Eikmeyer and H. Rieser. New York: de Gruyter
38. Kripke, S. 1972/1980. *Naming and Necessity*. Oxford: Blackwell.
39. Kripke, S. 1982. *Wittgenstein on Rules and Private Language*. Cambridge, M.A.: Harvard University Press.
40. Kripke, S. 2011. 'Unrestricted exportation and some morals for the philosophy of language'. In *Philosophical Troubles: Collected Papers*. Oxford: Oxford University Press.

41. Kripke, S. 2019. 'Naming and Necessity Revisited - Prof. Saul Kripke'. *YouTube*, https://www.youtube.com/watch?v=3zazonG6zBk.
42. Lewis, D. 1978. 'Truth in fiction'. *American Philosophical Quarterly*, 15: 37–46.
43. Lewis, D. 1983. 'New work for a theory of universals'. *Australasian Journal of Philosophy*, 61(4): 343–77.
44. Lewis, D. 1984. 'Putnam's paradox'. *Australasian Journal of Philosophy*, 62(3): 221–36.
45. MacFarlane, J. 2007. 'Relativism and Disagreement'. *Philosophical Studies*, 132: 17–31.
46. Magidor, O. 2018. 'How both you and the brain in a vat can know whether or not you are envatted'. *Proceedings of the Aristotelian Society*, 42: 151–181.
47. Magidor, O. and Hawthorne, J. 2018. 'Reflections on the ideology of reasons'. In *The Oxford Handbook of Reasons and Normativity*, ed. D. Star. Oxford: Oxford University Press.
48. Mandelkern, M. forthcoming. 'A counterexample to modus ponenses'. *The Journal of Philosophy*.
49. Mandelkern, M. ms. 'If p, then p!'.
50. Manley, D. 2007. 'Safety, content, apriority, self-knowledge'. *The Journal of Philosophy*, 104: 403–23.
51. Martin, M.G.F. 2009. 'Reupholstering a discipline: comments on Williamson'. *Philosophical Studies*, 145: 445–453.
52. McGee, V. 1985. 'A counterexample to modus ponens'. *The Journal of Philosophy*, 82: 462–471.
53. McGlynn, A. 2012. 'Interpretation and knowledge maximization'. *Philosophical Studies*, 160(3): 391–405.
54. Putnam, H. 1967. 'Mathematics without foundations'. *The Journal of Philosophy*, 64: 5–22.
55. Putnam, H. 1975. 'The meaning of "meaning"'. Minneapolis: University of Minnesota Press.
56. Rayo, A. 2013. *The Construction of Logical Space*. Oxford: Oxford University Press.
57. Russell, G. 2018. 'Logical nihilism: could there be no logic?'. *Philosophical Issues*, 28: 308–24.
58. Stanley, J. 2001. 'Hermeneutic fictionalism'. *Midwest Studies in Philosophy*, 25: 36–71.
59. Unger, P. 1975. *Ignorance*. Oxford: Oxford University Press.
60. Williams, J. Robert G. 2007. 'Eligibility and inscrutability'. *Philosophical Review*, 116 (3): 361–99.
61. Williamson, T. 1994. *Vagueness*. London: Routledge.

62. Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
63. Williamson, T. 2003. 'Understanding and inference'. *Aristotelian Society*. 77, pp. 249–293.
64. Williamson, T. 2005. 'Reference'. In *The Oxford Companion to Philosophy* (ed. Ted Honderich). Oxford: Oxford University Press.
65. Williamson, T. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.
66. Williamson, T. 2009a. 'Reply to Hawthorne and Lasonen-Aarnio'. In *Williamson on Knowledge*, eds. P. Greenough and D. Pritchard. Oxford: Oxford University Press.
67. Williamson, T. 2009b. 'Probability and danger'. *The Amherst Lecture in Philosophy* 4: 1–35. http://www.amherstlecture.org/williamson2009/.
68. Williamson, T. 2009c. 'Replies to Ichikawa, Martin and Weinberg'. *Philosophical Studies*, 145: 465–476
69. Williamson, T. 2011a. 'Reply to Peacocke'. *Philosophy and Phenomenological Research*, 82: 481–487.
70. Williamson, T. 2011b. 'Reply to Boghossian'. *Philosophy and Phenomenological Research*, 82: 498–506.
71. Williamson, T. 2012. 'Boghossian and Casalegno on understanding and inference'. *Dialectica*, 66(2): 237–247.
72. Williamson, T. 2013. *Modal Logic as Metaphysics*. Oxford: Oxford University Press.
73. Williamson, T. 2016a. 'Modal science'. *Canadian Journal of Philosophy*, 46: 453–492
74. Williamson, T. 2016b. 'Reply to Fine'. *Canadian Journal of Philosophy*, 46: 571–583.
75. Williamson, T. 2017a. 'Acting on knowledge'. *Knowledge First: Approaches in Epistemology and Mind* (eds. J. Carter, E. Gordon, and B. Jarvis). Oxford: Oxford University Press.
76. Williamson, T . 2017b. 'Ambiguous rationality'. *Episteme*, 14(3): 263–274.
77. Williamson, T. 2017c. 'Counterpossibles in semantics and metaphysics'. *Argumenta*, 2: 195–226.
78. Williamson, T. 2020. *Suppose and Tell*. Oxford: Oxford University Press.
79. Williamson, T. forthcoming-a. 'Knowledge, credence, and the strength of belief'. In *Expansive Epistemology: Norms, Action, and the Social World*, eds. A. Flowerree and B. Reed.
80. Williamson, T. forthcoming-b. 'Justifications, excuses, and sceptical scenarios'. In *The New Evil Demon* (eds. J. Dutant and F. Dorsch). Oxford: Oxford University Press.
81. Williamson, T. forthcoming-c. 'Epistemological ambivalence'. In *Epistemic Dilemmas*, ed. N. Hughes.

82. Williamson, T. forthcoming-d. 'E=K, but what about R?'. In *The Routledge Handbook of Evidence*, eds. M. Lasonen-Aarnio and C. Littlejohn.
83. Wright, C. and Hale, B. 1992. 'Nominalism and the contingency of abstract objects'. *The Journal of Philosophy*, 89: 111–135.
84. Yablo, S. 1998. 'Does ontology rest on a mistake?'. *Proceedings of the Aristotelian Society*, 72: 229–283.
85. Yablo, S. 2000. 'A paradox of existence'. In *Empty Names, Fiction and the Puzzles of Non-Existence*, eds. A. Everett and T. Hofweber. Stanford: CSLI Press.
86. Yablo, S. 2001. 'Go figure: a path through fictionalism'. *Midwest Studies in Philosophy*, 25: 72–102.
87. Yablo, S. 2002a. 'Red, bitter, best'. *Philosophical Books*, 41: 13–23.
88. Yablo, S. 2002b. 'Abstract objects: a case study'. *Nous*, 36: 220–240.
89. Yablo, S. 2012. 'Explanation, extrapolation, and existence'. *Mind*, 121: 1007-1029
90. Yablo, S. 2014. 'Carnap's paradox and easy ontology'. *Journal of Philosophy*, 111: 470–501.
91. Yablo, S. 2017a. 'Precis of *Aboutness*'. *Philosophical Studies*, 174: 771–7.
92. Yablo, S. 2017b. 'If-thenism'. *Australasian Philosophical Review*, 1: 115-132.
93. Yablo, S. 2017c. 'Replies to comments on if-thenism'. *Australasian Philosophical Review*, 1: 212–227.

Declaration of Originality of Authorship of a thesis submitted for examination in the BPhil in Philosophy in the academic year 2019-2020

I, Daniel Kodsi, affirm that my submitted thesis entitled 'Charity and Analyticity' is my own original work, except where otherwise stated, and that I have only shown drafts of it to my supervisor, Prof Ofra Magidor, in approximately five hours of supervision.

Signed: Date: 07/07/2020