2022

Reference Without Acquaintance: Naming and Thought in Fiction and Science

Yiğit Can Kahyaoğlu
Macalester College, yigitckahyaoglu@gmail.com

Reference Without Acquaintance: Naming and Thought in Fiction and Science

Yiğit Can Kahyaoğlu
Advised by Janet Folina
Philosophy Department
Macalester College
Spring 2022

Mill thought the semantic content of proper names is simply their referents. This is true even for objects we are not acquainted with, even though many contemporary Millian theories appeal to causal connections to account for the reference and thus the meaning of such names. However, in certain contexts, we refer to objects that we are not causally related to — moreover, we seem to be able to think about such objects, as well. Such cases appear in science and fiction, for example, when we talk about hypothetical objects and fictional characters. In this thesis, I will examine the naming practices in fiction and science, and look at how reference is fixed when speakers are not acquainted with the objects they are referring to. I will argue that reference can be fixed in many ways, and due to the close relation between reference and thought, we can think and talk about objects without acquaintance constraints, as long as we participate in the relevant social practices.

Acknowledgments

I’m extremely grateful to Janet Folina who guided me through this project, as well as my four years at Macalester. Her support and insights have been invaluable. I would also like to like to thank Hannah Kim and Joy Laine for being a part of my defense committee. Finally, I would like to express my appreciation to the rest of the philosophy department, and to my friends who were always there to help me think through philosophical problems.

Table of Contents

Introduction

Kripke's theory of reference gives a convincing account of the relationship between objects and names, basically as a causal relation between acts of naming and speakers. Moreover, Kripke rejects the Russell/Frege view of names that attribute descriptive semantic content to names. Instead, Kripke follows Mill's analysis of names according to which names have only reference, but no 'sense,' that is, their meaning is exhausted by their referents.

The causal theory of names is as convincing as it is influential. However, the way it is originally stated in *Naming and Necessity* does not give a full account of how we use proper names. Most cases Kripke discusses are generic cases of naming, which he calls 'initial baptisms,' where a person or object is ostensibly named. Those who were not present at the initial baptism get to use the name through causal chains that connect speakers to these acts of naming. Kripke thereby conceives of reference as a phenomenon that relies heavily on linguistic communities.

While this paradigmatic scenario demonstrates the basic tenets of the causal theory, this is not the only way in which speakers establish reference. Indeed, there are a lot of contexts where there is no strictly causal relation between speakers and the object. In this thesis, I aim to examine two of those contexts: science and fiction. By doing so, I aim to give an account of how we establish reference while retaining the basic ideas from the causal theory, namely the crucial role causal chains play in ordinary reference, and the idea that names are purely referential.

Throughout this thesis, I focus on two problems. The first one is that of how, if at all, we can have singular thoughts about objects we are not acquainted with. By singular thought, I mean thoughts about particular objects, like "Mary is a color scientist" or "Kamala Harris is the Vice President." Contrast these to *general* thoughts, such as "The person in the next room is annoying" or "All philosophy students are nerds." In both of these sentences, there is no reference to a

particular individual. Although the first sentence is about one person, the person in the next room, the person is not directly referenced. Rather, the sentence is about the person next door, whoever that person happens to be.

Since singular statements make apparent reference to a definite individual, some philosophers have the intuition that we must be acquainted with the object we are talking about. Evans (1982), for example, argued that in order to have a singular thought about an object, speakers must know which object they are talking about. Notice how this is not a problem about general statements — there is no particular individual the speaker has in mind when she thinks "The person in the next room is annoying." In contrast, in order to have a singular thought about Mary, it seems reasonable to assume that a speaker must know who Mary is, in some sense of 'know.' However, there are cases where speakers seem to have singular thoughts about objects they are not acquainted with. These cases generally arise when the reference of a term is established in a way that is different from the Kripkean baptism, so speakers gain the ability to refer without establishing a causal relation to the object. One way to name an object in this way is via stipulation, which I will discuss thoroughly in Chapter 2.

The second problem I will be discussing is that of empty names. Empty names are proper names that have no referents. Examples of empty names include 'Vulcan', 'Tyche', 'phlogiston[1]', 'Pegasus', etc. Empty names are problematic because they sometimes appear in our assertions (and thus thoughts). I may utter, for example, that "Vulcan is a planet." Whereas propositions that contain ordinary non-empty names like 'Kamala Harris,' 'Neptune,' etc. are about the referents of the names, Kamala Harris and Neptune respectively, propositions that contain empty names seem to be about nothing at all. What, then, do we say when we utter the sentence "Vulcan is a planet"?

[1] 'Phlogiston' is not a proper name, but its function as a natural kind term is similar to that of a proper name, at least according to Kripke.

What do these propositions express, if anything at all? Moreover, given that reference is a relation between a word and an object, how do we end up with empty names in our language in the first place?

I think a theory of reference should answer these questions. In this thesis, I will seek to answer them by contextualizing the problems. I will not be discussing these two problems in a detached, abstract way. Instead, I will look at some particular uses of language where such problems regularly arise. I chose science and fiction, because the naming practices in these two areas generate problems of empty names and singular thoughts without acquaintance. In science, the practice that gives rise to these problems is hypothesizing. When we hypothesize an object that is not yet observed, we are, in effect, having thoughts and beliefs about something with which we are not causally related. When hypotheses fail, our names turn out to lack referents, and our thoughts contain empty names. To answer these problems, I will argue that epistemic, perceptual, or causal acquaintance is not necessary to establish reference, and that reference can be fixed by stipulation. Furthermore, I will argue that empty names introduced this way are nonetheless proper names, due to the cognitive and semantic role that proper names play.

In fiction, we refer to fictional characters and objects, with which we can never be causally acquainted. We cannot perceive Harry Potter, for example, and yet we have thoughts about him regardless. How, then, do we manage to refer to fictional characters? In order to answer this question, I will first argue that fictional characters exist, and that they are abstract objects created by social practices. Then, I will argue that we can refer to fictional characters by participating in certain relevant social practices, or by standing in a causal relation with speakers who do. So having singular thoughts about Harry Potter does not require seeing, or otherwise perceiving, Harry Potter in any way.

Before starting the discussion, I would like to make some preliminary points to set off the discussion. First, throughout this thesis, I am taking for granted that we have singular thoughts. It is clear to me that we do, and I will not be arguing for this claim.

Second, there is a very intimate relation between reference and singular thought: if one can refer to an object, then one can have a singular thought about it. The ability to form singular thoughts about an object does not require anything more than the ability to refer. I will also take this for granted, although this point is more contentious than the former.

Third, there are no universal conditions a speaker needs to satisfy to be able to refer. This is not to say that there are no conditions whatsoever, but rather that the specific conditions a speaker needs to satisfy is often context-dependent. In some cases, perception of an object may suffice for a speaker to have the ability to refer. In others, speakers may need to be situated in the causal link for reference to take place. In some others, as I aim to show in this thesis, reference is established through participation in social practices. I think this latter case is often overlooked, but it accounts for a lot of cases, including in scientific practice and fiction.

Fourth, to understand a proper name is to be able to use it in the appropriate way. This is partly a consequence of my commitment to a Millian theory of names. Using a name appropriately means the ability to refer, and since names have no sense but only reference, then competently using the name implies that the speaker understands the name, even if the speaker is not acquainted with, or knowledgeable about, the referent. I will discuss this claim more in-depth in chapter 1.

Here's a roadmap for the rest of the thesis. In chapter 1, I will talk about some instances of naming in scientific practice. I will discuss how names are introduced to language in hypotheses, and how descriptive reference-fixing is an important part of this process. Then, I will consider some alternative views, mainly the abstract entity view and descriptivism, I will reject both.

Finally, I will explain how descriptive reference-fixing as a legitimate way to introduce proper names can accommodate intuitions about scientific practice, as well as answer both the problem of empty names and that of singular thought without acquaintance.

In the second chapter, I will discuss how names are introduced in fiction. First, I will present an argument for artifactualism, the idea that fictional characters are abstract objects that are created by speakers. My account of artifactualism will rely on social and linguistic practices of communities to bring about and maintain the existence of fictional characters. Then, I will discuss how reference to abstract objects is fixed in a similar fashion to scientific practice: participating in a linguistic community will allow speakers to reference.

Chapter 1: Reference and Thought in Science

Introduction

Scientific practice gives rise to the two main problems this thesis focuses on: singular thoughts without acquaintance, and singular thoughts about non-existent objects. The first problem concerns our ability to form singular, object-oriented thoughts without having been acquainted with said objects. The second problem is about names that fail to refer, but are nonetheless a part of language. Scientific practice generates these problems because hypothesizing often relies on descriptive reference-fixing (DRF) for terms whose referents are not observed. For example, as a part of a hypothesis, scientists may purport to name an object theretofore unobserved. As a result, our two interesting problems about reference and thought arise.

First, let us define descriptive reference-fixing, and distinguish it from other practices of naming. A name is descriptively introduced when a speaker utters a sentence of the form "N is the F," where 'the F' is a definite description. Paradigmatically, speakers introduce names this way when they are not acquainted with the object that they are naming — they fix the referent of the name via stipulation, so that the name 'N' refers to the object that uniquely satisfies the description 'the F.' DRF is distinct from other acts of naming, such as naming by ostension, because the object named is not 'present' when the act takes place. In other words, speakers are not acquainted with the object when they are naming it this way. DRF takes place in scientific hypothesizing exactly because hypotheses are sometimes about things that are not known (in a generic sense of 'known') to the scientific community. In this chapter, I will refer to names introduced by DRF in a scientific context as 'hypothetical names,' since they refer to hypothesized objects.

Since, by definition, hypothetical names are those whose referents are as yet unobserved, we are faced with the question of how we can come to have thoughts and beliefs that contain these

names, given that, in a sense, we are in the dark about their referents. To demonstrate the problem, let us consider a popular case of singular belief without acquaintance from the history of astronomy: the Neptune case. In the 19th century, French astronomer Urbain Leverrier had a theory about what caused some perturbations in Uranus' orbit. He hypothesized that there was a planet that was causing these perturbations; a planet he dubbed 'Neptune.' Later, German astronomer Johann Gottfried Galle observed Neptune, thereby confirming Leverrier's hypothesis[2].

This case evokes conflicting intuitions. On the one hand, it looks as though Leverrier, Galle, and others in the European astronomical community had beliefs that were *about* Neptune. On the other hand, many philosophers have strict conditions for our ability to have singular thoughts about objects, and many require that we be, in one way or another, acquainted with an object in order to have singular thoughts about them[3]. So how could members of the astronomical community have had beliefs about Neptune? Conversely, if their beliefs were not about Neptune, what were they about? Are they perhaps not even singular beliefs, but only have the appearance of singularity?

The second problem is raised by unsuccessful hypotheses, which generate names that lack referents. Similar to the singular belief without acquaintance problem, empty names appear in propositional attitude reports, even though they seemingly refer to nothing. A famous case of this, also due to Leverrier, is the Vulcan case. A few years after the success of his Neptune theory, Leverrier observed some perturbations in the orbit of Mercury, and hypothesized that there exists a planet, which he named 'Vulcan,' that was causing these perturbations. In the following years, astronomers looked for planet Vulcan to no avail. In the 20th century, the Vulcan theory was abandoned, and the perturbations in Mercury's orbit were explained by general relativity.

[2] This is not a historically accurate telling of the story; however, it makes a clearer example.
[3] See Russell (1985) and Evans (1981) for two prominent examples, both of which I will discuss later.

Although the Vulcan theory was eventually disproved, astronomers seemingly had beliefs about Vulcan for decades. At least they certainly thought they did — presumably, if we were to ask a French astronomer at the time what he believed was causing said perturbations, they would respond with something like "I believe that the perturbations are caused by Vulcan." Furthermore, we could probably question them further to elicit genuine reports of belief about Vulcan. So, at least in appearance, people used to have apparent beliefs about Vulcan. However, given that Vulcan does not actually exist, what were their beliefs *really* about? Certainly not an actual planet. What does it even mean to have beliefs about a non-existent object?

In the rest of this chapter, I aim to answer these questions. I will start by examining two alternative accounts. First, I will review a descriptivist account for the semantics of hypothetical names, and I will argue that it is ultimately inadequate, for reasons that are similar to Kripke's rejection of descriptivism. Then, I will discuss the abstract object theory — the idea that when we introduce a name by DRF, we consequently create an abstract object to which our name refers, at least until we confirm or disconfirm the hypothesis. I will reject this view due to its counterintuitive implications, such as the reference shift this view requires after scientific confirmation or refutation. Finally, I will present my own account, which admits DRF as a legitimate and simple way of introducing a name, and is unbothered by the lack of causal connections and acquaintance. Instead, in cases like DRF, the relevant community (in this case, scientific) will have the authority to fix the reference of the term, and speakers can successfully refer to objects insofar as they follow the community's linguistic practices. Although the reliance on community practices may bother proponents of direct reference, I will argue that this account successfully accommodates scientific language and practice.

Alternatives

Descriptivism

One way to account for hypothetical names is (at least a partial) return to Russellian descriptivism. For Russell, most names from ordinary language were basically abbreviated descriptions that are uniquely satisfied by the object to which they refer. (Russell 1914). The semantic analysis of a name like 'Joe Biden,' for example, is a description like 'the President of the United States in April 2022.' As a result, our beliefs about Joe Biden, and most ordinary objects, express *general* propositions, and not singular[4]. This view was later adopted by others, including Wittgenstein (1953) and Searle (1958), though it was later criticized by Kripke (1980) and others in the second half of the century for the sake of a Millian theory of names that denied denotative content to names. Nevertheless, one may adopt a partial descriptivism in order to account for hypothetical names. Since most hypothetical names are introduced to language by DRF, one may argue that the description by which a speaker purports to fix the reference of the name just is the meaning of the name. That is, when a speaker utters "N is the F," the meaning of 'N' is not the object that uniquely satisfies 'the F,' but rather the description 'the F' itself.

There are a couple of routes a partial descriptivist could take. In a stronger version of this view, one may argue that all names generated by DRF are descriptions in semantic content. An initial problem with this view is that, intuitively, we use the name 'Neptune' now as a Millian proper name, even if it was introduced by DRF. In order to accommodate that intuition, we can make the following addition: every name introduced by DRF is an abbreviated description, but once a speaker is acquainted with the object that uniquely satisfies the description, an implicit

[4] Technically, Russell did not consider these 'logically proper names,' since we are not directly acquainted with any ordinary objects. The only logically proper names for him were 'I,' 'this,' and any name one can give to a particular immediate sense-datum.

baptism takes place, and the name becomes a Millian proper name. So, once Le Verrier's hypothesis was confirmed, the semantic content of the term 'Neptune' went from 'the planet that causes certain perturbations in Uranus' orbit' to the planet. A version of this view is defended by Urtizberea (2005).

Alternatively, one can make the weaker claim that hypothetical names are descriptions only when the hypothesis is unsuccessful, that is, the hypothesized object does not exist. In this case, 'Neptune' is a Millian proper name whereas 'Vulcan' is a description. The appeal of this view is that it explains away the problem of empty names, or at least it explains away many instances of the problem. Sentences like 'Vulcan is a planet' or 'Vulcan does not exist' are not problematic due to a lack of referent, and a Russellian analysis can straightforwardly show that the former sentence is false while the latter is true. Similarly, it is impossible to formulate a singular negative existential statement under this view, since any name[5] that lacks a referent is a description, so cannot form a singular proposition in the problematic way.

This idea is controversial, however. A descriptivist about empty names must explain why empty names have the appearance of Millian proper names, and are sometimes introduced to language in the same way that proper names are, but are different from proper names with regard to their semantic content. I don't think there is a satisfactory explanation for this, for the following reason. The semantic content or function of a word is a fact about language, and it should be determined by linguistic practice. It is counterintuitive to say that, although 'Vulcan' is a description, it would have been a proper name if the planet existed, even though it was introduced to language as a name just like 'Neptune,' and used as such. Facts about language should not

[5] Perhaps this does not explain away all empty names, but a good chunk of them can be accounted for this way, making singular negative existentialism no longer problematic for those names.

depend on the way the world is; they depend on our linguistic practices, on how we use language, etc. For this reason, a theory that lets the world determine linguistic facts is unattractive.

A comparison between Vulcan and Neptune helps clarify this point. There are a lot of similarities between the cases of Neptune and Vulcan. They were introduced to language in the exact same way, not by a baptismal process, but by a description as a part of a hypothesis to explain some data. A descriptivist about empty names should explain why Neptune was a rigid designator, whereas Vulcan was not. The obvious answer is that because Neptune exists, and Vulcan doesn't. But the respective existence and non-existence of Neptune and Vulcan have nothing to do with linguistic practices. They are facts about the world, not about language. If we accept that 'Neptune' is a rigid designator but 'Vulcan' isn't, then we let facts about the world determine some facts about language, namely, the semantic category of the words 'Neptune' and 'Vulcan.[6]' Facts about language are determined by how we use language. A view that lets the world determine facts about language in this way is unacceptable[7].

A descriptivist can concede this point but defend her view in the following way. It is not that 'Neptune' is a proper name and 'Vulcan' is a description from the start. Rather, they both start asdescriptions. After Galle observed Neptune and confirmed the hypothesis, the name 'Neptune' became a proper name. Vulcan was never observed, so the name 'Vulcan' remains a description. In the case of 'Neptune,' there was a semantic shift: it entered language as an abbreviated definite description, and then became a proper name that refers to the object that was confirmed to satisfy the description. This view is plausible, primarily because it does not contradict the principle

[6] I hold that terms like 'Vulcan' are empty designators that are necessarily empty, but this I will not argue for this claim here as it is not relevant.

[7] Of course, there are some facts about language that are determined by the world, such as the meanings of referential terms. But these are facts about the content of sentences and terms, and not the structure. Structural facts about language are determined by how we use language.

loosely stated above, that facts about language are determined by linguistic practices. If the descriptivist is right, then the semantic type of the word ‘Neptune’ shifts because we change the way we use the word. Of course, this change is caused by a fact about the world, but it is the change in linguistic practice, and not some brute fact about the world, that determines the semantic shift.

However, this account still runs short of a satisfactory explanation of empty names, for reasons similar to Kripke’s (1980) argument against descriptivism. Braun (1993) provides an interesting counterexample to this view. Suppose a speaker associates the description “the planet closest to the Sun” with the name ‘Vulcan.’ If the meaning of ‘Vulcan’ was this description, then the utterance “Vulcan exists” would be true, whereas it is false, because there is a planet that is the closest to the Sun, namely, Mercury. In the case of non-empty names, the Searle/Wittgenstein version of descriptivism is unaffected by this argument, because according to this view, it is not a particular definite description that determines reference, but rather a majority of descriptions that apply to the object. This version of descriptivism is especially unsuitable for empty names, however, since there are no descriptions that can apply to the object; there is no object. Braun’s example hints at a fundamental problem about descriptive names, though it is not inadequate. Braun puts too much emphasis on the individual’s own attributed description, whereas a sufficiently clear example will take into account a whole linguistic community.

Here is another example, similar to Braun’s, that I think makes the point clearer. This is analogous to Kripke’s Gödel/Schmidt example. First, a recap of Kripke’s original case. According to Kripke, most people, if they are familiar with the incompleteness theorem at all, know that Gödel was the person who proved the incompleteness theorem. Suppose that it wasn’t actually Gödel, but some other person, Schmidt, who proved the incompleteness of arithmetic. Gödel copied Schmidt’s proof and published it under his name. Now, if the descriptivist theory is correct,

then the meaning of the name ‘Gödel’ is the description “the man who proved the incompleteness theorem” — remember that no other definite description is available to the community. So, if the hypothetical was in fact the case, then, when people say Gödel, they would actually refer to Schmidt, since Schmidt is the actual person who proved the theorem and Gödel is a mere impostor. But that can’t be right, Kripke argues. When people say ‘Gödel,’ they refer to Gödel, even if the only description of Gödel they have is satisfied by Schmidt. If the semantic contribution of a name was a definite description, then a speaker who utters the sentence “Gödel proved the incompleteness of arithmetic” would be expressing a true proposition about Schmidt, whereas it is a false proposition about Gödel, or so Kripke argues. Hence, he concludes, the semantic content of a name cannot be a description, and proper names are purely denotative.

Here is a scientific case similar to the Gödel/Schmidt example, loosely based on the Vulcan story. Suppose an astronomer notices some perturbations in the orbit of a planet P. He hypothesizes that there exists a heretofore unobserved planet, Q, which causes these perturbations. Astronomers look for Q for years with no success. As a matter of fact, what they did not know was this: the cause of the perturbations in the orbit of P was another planet, R. Astronomers were already aware of the existence of R, but R had some unknown property that caused the perturbations. Nonetheless, astronomers used the name ‘Q’ to refer to the planet that they thought satisfied the following description: “the planet that causes the perturbations in the orbit of P.” We may even assume that there was an explicit utterance such as “We shall refer to the planet that causes the perturbations in the orbit of P by the name ‘Q’.” In this case, did astronomers actually refer to R when they meant Q? This seems implausible. In this scenario, astronomers were aware of the existence of R, but they were unaware of some property of it. That is why they posited the existence of a *new* planet, as opposed to making new claims about R. After having discovered the new

property of R, there are two things the astronomers can say about the Q hypothesis. They can say that their hypothesis is confirmed and that they discovered Q, and that R is Q; or they can say that their hypothesis is failed, and that Q does not exist. I think it is clear that the latter is more plausible.

This raises an important point about hypothetical entities, entities that are posited as a part of a scientific hypothesis. Hypotheses, when positing the existence of such entities, will have some descriptions of that entity by which they fix the reference to the object, but they will never have a complete description of the hypothesized entity. In fact, it is impossible to have a complete description of an object. This has the consequence that the descriptions by which we fix the reference are always lacking, and that their reference is indeterminate. Kripke thinks this is a problem for fictional characters (Kripke 1970), I think it is also a problem for failed hypotheses. Failed hypotheses generate names like Vulcan, or Q from the above example, and the description associated with them may be satisfied by some real entities. In these cases, if we take descriptivism to be true, there is a referential ambiguity. It is not immediately clear, in the above case, whether the name 'Q' lacks a referent or if Q has the same referent as R.

Notice how the Q/R ambiguity is different from the Gödel/Schmidt ambiguity. The latter case provokes stronger intuitions about which person we are referring to. This is because, in that case, we are talking about existing entities, and who we intend to refer to is clear. This is not so in the Q/R case, because one possible interpretation suggests that the object we intend to refer to does not exist. Since it is a non-existent entity, it is not as clear what we intend to refer to with that name. This ambiguity is not by itself a conclusive argument against descriptivism, but it is a good reason to be skeptical against it. An advantage to treating empty names as proper names is that a Millian view would avoid this ambiguity. Moreover, I think a descriptivist about empty names

who would hold that the name ‘Q’ refers to R, but not that the name ‘Gödel’ refers to Schmidt (if that thought experiment was in fact the case) needs to justify this distinction.

Abstract Object Theory

Some, like Salmon (1998) argued that names like ‘Vulcan’ fail to refer to a physical object astronomers thought to exist, but they are nonetheless not devoid of denotation. Instead, their referents are abstract objects. Although there are different versions of this view, the main idea is that scientific hypothesizing generates abstract objects, and these objects are the referents of the terms introduced to the language as a part of that hypothesis.

There are two versions of this view, both of which have counterintuitive implications. The first view, supported by Ackerman (2016), has a two-fold claim. When the reference of a name is fixed in a hypothesis, if the hypothesized object exists, the name simply refers to the object. On the other hand, if there is no such object, then the name refers to an abstract object that is created by the scientific community. So, in unsuccessful cases, the community (or the person who introduces the name) creates an abstract object that becomes the referent of the name.

The problem with this view is that, at the time of hypothesizing, nobody knows whether the object actually exists or not. As a result, we do not know whether we create an abstract object. Ackerman does not think it is a problem that we can create abstract objects unintentionally in this way, but I see two issues with it. The first is that this view comes with a demanding ontology, where any time a speaker utters a sentence like “N is the F,” provided there is no unique F, we bring a new object into existence. There doesn’t seem to be any requirements on the end of the speaker to commit to the use of the name, or have a particular intention (such as explaining an observed phenomenon). Thus, Ackerman makes abstract objects a little bit too easy to create. This is not a contradiction in his view, but it is a disadvantage of it.

The second issue with this view is that the speaker who creates the abstract object seems not to be aware of their creation. It is a bizarre consequence of this view. My issue with it is not, however, that an agent needs to have the relevant intention in order to create an object. This is not always true. Many times, I had the intention to create an omelet but I instead created scrambled eggs, even though I had no such intention. Rather, the problem is that uttering very similar sentences ("N is the F") can have drastically different consequences, even though the utterances belong to the same kind of linguistic practice. One sentence can simply introduce a name into a language but another will create a whole new object, where the result fully depends on some unknown facts about the world. Again, this is not a view that is self-contradictory, but it requires accepting some counterintuitive assumptions about abstract objects and language.

The second version has a more unified story. According to this view, when we hypothesize about an object, we *always* create an abstract object. When the hypothesis is confirmed, the reference of the name shifts from the abstract object to the physical one. If a hypothesis is disconfirmed, then the referent of the name remains the same. This view avoids the second problem Ackerman's view has, though it still admits a lot of abstract objects into our ontology[8].

However, the reference shift from the abstract to the physical object presents a problem. Since the referent of the name is different before and after the observation of Neptune, the utterance "Neptune is a planet" expresses different propositions at different times. Indeed, we can make a distinction between Neptune_1 and Neptune_2, where the former refers to the abstract object and latter to the physical. Then, it seems as though discourse about Neptune_1 and Neptune_2 can be wholly independent of one another. But this is obviously not the case, since everything we want

[8] This is not necessarily the case, however. There may be variants of this view that put more demanding conditions on creating abstract objects (for example, by tying it to established social practices). This would indeed make a stronger view, but I have independent objections to the view that I will discuss.

to say about Neptune_2 should be true of Neptune_1. On the other hand, this can only be true if Neptune_1 and Neptune_2 are identical, which they are not. So, there must be some other sort of relation between Neptune_1 and Neptune_2 that holds which can justify discourse about Neptune, and I am not sure if we can find one unless we make up an extremely *ad hoc* relation.

Note also that the names Neptune_1 and Neptune_2 function differently in modal contexts. For example, it is true that "Neptune_2 could have been located between Saturn and Venus," but this would have been false of Neptune_1, since abstract objects are not spatially located. Moreover, according to this view, "Vulcan does not exist" is a false sentence, because the abstract entity created by Leverrier (presumably) still exists. These are counterintuitive consequences of the abstract object view which make it unsuitable for scientific hypotheses.

Reference without Acquaintance

In the rest of this chapter, I will argue that reference can be established without an acquaintance relationship, and as a result, we can have singular thoughts without acquaintance, as well. I will start by reviewing some cases where acquaintance conditions are challenged, but are generally accepted as situations where speakers successfully refer. Then, I will discuss how DRF is used in scientific practice, and how accepting DRF as a legitimate way to coin a proper name helps us account for scientific discourse. I will support this claim by appealing to linguistic practices and community, and how speakers must gain the ability to refer by participating in them. Finally, I will consider what it means to understand a referential term, and argue that semantic understanding of a name is the ability to use the name appropriately, and that we should not confuse it with understanding a theory, phenomenon or object.

The main feature of descriptive reference-fixing seems to be that we refer to objects that we are not *acquainted* with. A lot of the time, we don't even know a whole lot about the object

that we are referring to. The intuition driving the case against DRF is that, in order to refer to an object, a speaker must stand in some relation to the object, either causal, perceptual, or epistemic. There are varying levels of expectations of rigor for relation among different philosophers. Some think that reference (and singular thought) requires direct acquaintance with the object of thought. This was certainly Russell's (1918) motivation to deny singular thoughts about anything other than sense-data and the self, since, strictly speaking, we have no direct acquaintance with anything else[9].

Russell's rigid conditions for singular thought have been criticized, but some expectation for acquaintance remained for many. Evans (1982), for example, argued (in line with Russell) that in order to have a singular thought about an object, one must have the capacity to distinguish the object from everything else. This capacity can be acquired causally or epistemically, but without it, according to Evans, we cannot have singular thought. I will not argue against Evans in this section. Instead, I will discuss some familiar ways of referring to objects without acquaintance, and motivate intuitions towards loosening acquaintance conditions.

Causal Chains. First, let's examine a standard case of singular beliefs where no one who currently holds this belief is acquainted with its object. I believe that Aristotle was an Ancient Greek philosopher. The content of this belief is expressed by the proposition that "Aristotle was an Ancient Greek philosopher." More formally, we can analyze this proposition as such:

<Aristotle, being an Ancient Greek philosopher>

Now, I am not personally acquainted with Aristotle, but it is possible for me to have this belief about him. The standard Kripkean explanation for this is that I am causally connected to Aristotle, that is, there is a causal chain of speakers that connect me to Aristotle. This chain enables

[9] We may even question what it means to be directly acquainted with 'the self,' so Russell's theory of logically proper names is even more strict than it initially appears.

me to refer to him, and as a result, also enables me to have singular thoughts, and utter sentences that are about him. I can do this without direct acquaintance; the causal chain helps me outsource the acquaintance to some other speaker.

The Aristotle case is unproblematic, because the referent of the name 'Aristotle' enters into the causal explanation of the name from the beginning. Aristotle was named *baptismally*, that is, he was present when the act of naming took place. So, whoever named Aristotle, and the immediate community of speakers who started using the name 'Aristotle,' were acquainted with him from the beginning. The rest of us are causally connected to Aristotle through linguistic communities, and we gain our ability to talk and think about Aristotle through that connection, despite our lack of direct acquaintance.

However, the Aristotle case by itself is not enough to convince anyone of abandoning an acquaintance condition. It may be argued that what happens in this case is that our causal connection helps us *outsource* our acquaintance condition to those who were directly acquainted with him. So, an objector may argue, this is not a case where there is a lack of direct acquaintance, but speakers may defer to others in the community to refer. As long as there exists *some* member of the community that is directly acquainted, we can refer to Aristotle. The key is that there must be at least one.

By and large, for this case, this is correct. I don't think direct acquaintance is a necessary condition for all reference or singular thought, but it definitely explains a majority of the cases. This is one of them. However, there is an important point to be made. Although the causal chain allows us to outsource the acquaintance in this case, the perceptual content of acquaintance does not transfer through the chain. So, insofar as the acquaintance condition involves perceptual experience, the Aristotle case *does* pose a counterexample, because I, for example, do not have

access to the perceptual content to anyone's direct acquaintance with Aristotle, although I am a link in the causal chain. This is a good reason to be skeptical about a condition on singular thought and reference that relies heavily on perceptual experience, like that of Russell.

***Abstract Objects*.** The second chapter of this thesis discusses how we refer to abstract entities (as well as how we create them), but I would like to make some remarks here about the viability of the basic idea. We talk about abstract objects all the time. Of course, there are some who deny their existence altogether, and I will not attempt to argue for their existence here. Instead, I will talk about a few different kinds of abstract objects, with the hope that the reader will accept at least some of them into their ontology. Take, for example, numbers. 'Two', 'three,' and so on are referential terms for numbers. Since numbers are (arguably) abstract objects, we cannot causally interact with them. As a result, mathematical discourse cannot require a causal acquaintance condition.

If there would be any sort of acquaintance condition for mathematical objects at all, it would have to be epistemic. For instance, one may argue that in order to have a singular thought about the number 2, you must have some discriminating knowledge about it. Given the simplicity of this case, we may say that when someone understands basic arithmetic, one can have discriminating knowledge about a number. Suppose someone tells you that they believe that 2 is the smallest prime number, and you want to see whether they have the ability to tell the number 2 from, say, the number 3. In that case, they can tell you that 2 is even while 3 is odd, so they are different numbers. The person can generate answers for each number using basic arithmetic.

However, individual numbers may be deceptively simple examples, and epistemic acquaintance conditions are complicated when we look at slightly more abstract mathematical objects. Suppose John has a very basic understanding of mathematics that consists mostly in

middle school arithmetic and algebra. Suppose further that, although John understands what natural numbers are, and can describe the set of natural numbers, he does not understand the distinction between a set and a sequence. In that case, he would not have the ability to distinguish between the set of natural numbers from the sequence of natural numbers (the difference, of course, is that the latter is ordered, unbeknownst to John). With the level of understanding John has, I believe it is possible that he has singular thoughts about the set of natural numbers, though he lacks the relevant distinguishing knowledge.

We can make similar observations about other, socially-constructed abstract entities, like states. Suppose John also lacks geographical competence, so that he does not have the ability to distinguish between Latvia and Lithuania. He understands perfectly well that they are two European countries, but he does not know anything about them. Say John has a friend, Stella, who tells John that her ancestors came from Latvia. Now John can have the belief that "Latvia is the country that Stella's ancestors lived in." He may have this belief regardless of his inability to distinguish Latvia from Lithuania. Of course, we may say that this piece of information about Stella's ancestors *is* the distinguishing knowledge, but this is not necessarily the case. For all we know, Stella has ancestry in Latvia *and* Lithuania, so what John knows may not distinguish the two countries at all. This is simply not relevant to his ability to refer to Latvia, and have singular thoughts about it.

Hypothetical Objects

In the last two sections, I gave some examples of cases where reference and thought are successful despite lack of acquaintance. It is important to note that the two cases discussed above are by no means conclusive evidence for reference without acquaintance. For the first case, one may argue that acquaintance *is* necessary, just not for everyone in the community, and that the

participation in the relevant causal chain allows speakers to bypass the acquaintance requirement. For the second case, one may deny the existence of abstract objects altogether, or even say that, since abstract objects are so unlike physical objects, it is only natural that how we refer to them and think about them are different also. Now, I will make a case for why we can still refer to objects, even physical ones, as long as we fix their reference in the correct way.

In Neptune-like cases which occur in science, the contested point is whether the reference of a term can be fixed stipulatively, using a description. Presumably, the reference of 'Neptune' was fixed with an utterance such as "Let Neptune be the planet that causes the observed perturbations in the orbit of Uranus.[10]" The issue here is whether DRF is a legitimate way of introducing referential terms into language. However, it is important to note that this problem is not only a problem in philosophy of language, it is also a problem about the mind. Arguably, if DRF is legitimate, then we can refer to objects that we are not at all acquainted with. Moreover, if we can *refer* to objects, then we can *think* about them also. In other words, we can have singular thoughts about them. So descriptive reference-fixing is significant in drawing the limits to singular thought.

The main argument for allowing DRF is that it is consistent with how things appear to be. When astronomers were talking about Neptune before it was observed, it looked like they had singular thoughts, and that they were referring to Neptune. So simply accepting DRF as a legitimate way to generate proper names gives us an intuitive explanation of hypothetical names.

Jeshion (2002) gives a few conditions by which we can judge the success of a name introduction. These include a sincerity condition which entails a commitment to using the name, a psychological neutrality condition that requires the speaker to token the name without a particular

[10] Probably nobody said this, but assuming somebody uttered something like this makes it clearer to analyze.

mode of presentation, etc. Jeshion's argument is that DRF is legitimate, but it is parasitic on 'regular' naming conventions, like ostensive naming. When ostension is not possible, DRF works only if these conditions hold. Although I generally agree with her conclusion, I think her argument has a problem.

Jeshion is committed to the following principle:

> *Primacy of Ostension*: Fix the reference of a term descriptively just in case, so far as you know, you cannot do it ostensively.

To be exact, Jeshion sees *Primacy of Ostension* not as a strict condition that must hold, but rather as a general rule for good practice. Although her motivations are clear for adopting this principle, it is ambiguous what 'just in case you *cannot* so it ostensively" here means. There are some objects that are unobservable by nature, such as quarks. Should we follow this principle only in cases like this, where it is *impossible* for us to observe the object? Or is it a weaker principle that asks us to fix the reference of a term descriptively only when it is inconvenient to do it by ostension. When Leverrier named Neptune, it wasn't impossible for him to name the planet until it was observed, but it was way more convenient to do it descriptively, prior to observation. *Primacy of Ostension* is a good principle only when we interpret the 'cannot' in a context-dependent way: in scientific hypothesizing, it is often more convenient to fix reference descriptively, and hence it is an appropriate way to introduce a term.

Otherwise, Jeshion states a few conditions that must hold for a naming to be successful. Most of them apply to ostensive naming as well as descriptive. I will not discuss all of her conditions in-depth; however, her primary one, the sincerity condition, is inadequate. She formulates the sincerity condition as follows:

> *Sincerity* (Description): S intends for 'N' to name the F, whatever object it is, and to use 'N' as a name for it.

This is an important point. When a speaker introduces a name, whether by ostension or description, she must be committed to actually using the name to refer to the object. Otherwise, the act of naming is pointless. However, *Sincerity* underestimates the importance of community. A speaker must also intend to establish a linguistic practice in her community.

I suggest that what makes DRF legitimate is, firstly, the speaker's compliance with generally accepted norms of scientific practice (in terms of how hypotheses are made), and secondly, the speaker's genuine intention to refer to an object in the world. In the context of science, a speaker who uses DRF to introduce a proper name must do so with the intention to establish a use practice with the name in the scientific community. For instance, if a speaker utters a sentence like "Titan is the planet causes the perturbations in the orbit of the Earth" in a random manner, with no intention to actually corroborate the theory or even use the name consistently to refer to some entity that exists, the speaker fails to introduce the name 'Titan.'

Notice that both 'Neptune' and 'Vulcan' are names that are introduced this way. Leverrier had the intention to refer to an actual object in both cases, and his use generated a linguistic practice: astronomers used the two names in their research programs. The problem here is that the success of the Neptune case may not convince readers of the success of the Vulcan case. After all, Vulcan does not exist.

In the rest of this section, I will respond to concerns. First, I address worries about reference-fixing and empty names. Second, I will discuss the cognitive significance of proper names, and explain why the way in which a proper name is introduced should not matter for cognitive purposes.

Cases of Non-existence Not all definite descriptions are uniquely satisfied by an object – some have null extensions. Russell famously discusses the 'the present king of France.' Russell's

analysis of descriptions gave him a way of analyzing non-referring expressions. But what if we stipulate that the name ‘Paul’ refers to the present King of France. If we accept that DRF is possible, then we just introduced an empty name into our language. DRF easily allows such cases, so, one may argue that DRF is a problematic idea that we should reject.

However, given that there are a lot of successful cases of DRF, simply disallowing it would be too extreme a response. A more subtle point would be that DRF is only successful when there actually is an object uniquely satisfying the description. When naming is unsuccessful, Donnellan (1977) remarks that the naming “has been an unhappy one and [is] not to be taken as being in effect.” There are two things that may have failed to happen here: either we failed to fix the reference of the term, or we failed to introduce the name into language altogether (the latter includes the former).

I believe the former is true, but that the latter is not. The reference has not been fixed, since there is no referent. Alternatively, we may argue that the term is designated so as not to refer. Either way, what we have is an empty name. But I think to argue that we failed to introduce a name comes with a big problem: although a term may be non-referring, agents can still use the name in felicitous utterances. The Vulcan case is a good example of this. Vulcan was stipulated to be the planet that caused the perturbations in Mercury’s orbit, but it was later discovered that there is no such planet. So, the name ‘Vulcan’ was empty all along, and the scientific community’s beliefs about Vulcan were wrong. There are two things worth noting in this case. First, the scientists’ utterances about Vulcan before the Vulcan theory was abandoned were felicitous, that is, the scientists were not making claims about a planet they knew didn’t exist. Rather, they genuinely believed at least that it was possible that there was a planet to be causing said perturbations. They were trying to confirm the theory. So, up until disconfirmation, the name ‘Vulcan’ was used as a

proper name. Second, the scientists' beliefs about Vulcan were structured to be singular thoughts, and they were organized as such, just as they would have been in a case where the reference-fixing is successful. Cognitively speaking, there is no difference in the psychology of speakers whether the object of their singular thoughts exists.

Here's what I mean by this. There is a distinction between the form of a thought (or the corresponding sentence) and its content. For example, a thought can be singular or general, which concerns its form. A singular thought is one that refers to a particular object, while a general thought makes a claim about an object that satisfies a description, *whatever that object is*. A singular thought is something like "Joe Biden is Catholic." The subject of this sentence is the object picked out by the name 'Joe Biden.' On the other hand, a general thought is something like "The President of the United States is Catholic." Although (in the actual world) this thought *is* about Joe Biden, there is no direct reference to him in this sentence. The description "the President of the United States" happens to be satisfied, in this world, by Joe Biden, but that is not necessarily the case.

There is a familiar distinction between general propositions and singular propositions that regards their truth conditions with respect to possible worlds. Although both propositions above are about Joe Biden, this is a contingent feature of the actual world. There are possible worlds where the description "the President of the United States" is satisfied by some other person, say by Donald Trump in a possible world where he won reelection. Conversely, the referent of the name 'Joe Biden' is the same person with respect to all possible worlds (where Joe Biden exists). Hence, there are possible worlds in which the truth values of the two propositions are different. So, whether a proposition is singular or general has implications for its truth conditions.

However, there is also a cognitive distinction between singular and general propositions. General thoughts make general claims about the world — particular objects are not mentioned in such claims. Singular thoughts, on the other hand, are claims about particular objects. As a result, proper names[11] have the cognitive function of organizing information about particular objects (Jeshion 2002). The thoughts "Joe Biden is Catholic," "Joe Biden is from Delaware," and "Joe Biden is the logician who proved the incompleteness of arithmetic" are all thoughts about Joe Biden, and the singular structure of the thoughts work so as to group these thoughts together. In this way, the proper name 'Joe Biden' works together with a mental file by which our information about Joe Biden is collected together. Note that the mental file contains *all* thoughts about Joe Biden, and not just the true ones. The truth or falsity of our information about Joe Biden is not relevant to the structure with which we organize that information, in the way that our ability to refer to him does not hinge on our knowledge of any true facts about him (Recanati 2013).

This cognitive function of a proper name is crucial in understanding why DRF generates proper names. In general, when we introduce a proper name, we also create the corresponding mental file. The mental file allows us to represent information in a way that we could not have done had we only had descriptive (or general) information.

Say James has the belief "The President of Nauru is the President of Nauru." This is a tautologous belief, and it is not about a particular individual. At any rate, no particular individual is referenced. Now, let James fix the reference of a name, say 'David,' to be the object that satisfies the description 'the President of Nauru.' He can utter a sentence like "From here on out, 'David' will refer to the President of Nauru." And now he can hold the belief "David is the President of Nauru." Note that the two propositions now have different modal statuses — the former is a

[11] Not all singular thoughts employ proper names, they may also employ pronouns, demonstratives, etc. The focus here is proper names however.

tautology but the latter is contingent on the outcome of Nauru's presidential elections. Now, James is able to collect his beliefs about David in a way that is centered around David the individual. Note that this shift does not require an acquaintance with David. James does not have to see David, or be otherwise perceptually acquainted with him, to undergo this cognitive change.

Now, let's go back to cases where scientific proper name fails to refer. In the Vulcan case, when the name 'Vulcan' was introduced, the speakers consequently created mental files by which they organized their beliefs about Vulcan. Notice that, despite their collective falsity, all their beliefs purported to be about the same object. Given that their falsity has no bearing on the way they are organized cognitively, they may still be organized together, that is, they had the same purported object.

In fact, the idea that proper names are the main organizational principle of mental files allows us to differentiate between different sorts of false beliefs about non-existent objects. There are different proposals for the semantics of empty names, but one that I find plausible is the gappy proposition view (Braun 1993). According to this view, empty names are proper names that fail to contribute to propositions. So, whereas a proposition with a non-empty name like "Joe Biden is Catholic" can be analyzed like this:

<Joe Biden, is a Catholic>

Where the first position stands for the object and the second stands for the predicate, a proposition with an empty name like "Vulcan is a planet" is analyzed like this:

<, is a planet>

Such propositions are aptly called 'gappy,' since the place reserved for the object is not filled.

The problem with this view, however, is that the sentences “Vulcan is a planet” and “Tyche is a planet” seem to express the same proposition, that is:

<, is a planet>

Intuitively, we would say that the sentences “Vulcan is a planet” and “Tyche is a planet” express different propositions. However, the gappy proposition view implies that they express the same incomplete one. Of course, the gappy proposition view is not the only way to analyze these sentences, and there may be more appropriate methods that respect this difference. Nevertheless, what the gappy proposition view presupposes is that proper names do not have descriptive content, so Millians have a reason to worry about the semantic content of empty names, even if they do not subscribe to the gappy proposition view, for the idea that they contribute nothing is at least *prima facie* plausible for a Millian. However, we are now left with this unsavory consequence of being unable to distinguish between these seemingly different sentences.

This is where we can use mental files to solve this problem. Since information is stored in mental files proper name-wise, the sentence “Vulcan is a planet,” although false and possibly gappy, is stored under the file that corresponds to ‘Vulcan,’ and similarly for Tyche. Hence, the difference between ‘Vulcan is a planet’ and ‘Tyche is a planet’ does not have to be one of propositional content; it is a difference in how these beliefs are represented in the mind[12].

In order to account for these cases, we need two things. One, proper names organize our thoughts and beliefs around objects, so they correspond to mental files. Two, DRF successfully introduces a proper name, and so is sufficient to ‘open’ a mental file, even in cases where the name fails to denote an object.

[12] This is also the way the mental files theory responds to Frege’s puzzle.

Confirmation and Understanding

So far, I have drawn a picture that illustrates how proper names are introduced to language in scientific hypotheses, and why the mental files theory is a good way to think about the relation between reference, beliefs and objects, especially in the context of science. Now, I will talk about whether our understanding of the name changes after the confirmation of a scientific theory.

As stated above, the distinguishing feature of DRF is that it is done in the absence of the object named. However, in many cases of DRF, speakers get acquainted with the object eventually. For example, although Le Verrier started using the name 'Neptune' before having been acquainted with Neptune, the planet was later observed, thereby confirming Le Verrier's hypothesis. So, the object eventually entered into the causal explanation of the name. I will argue that, in terms of the semantics (or the broader linguistic features) of the name, confirmation/perceptual acquaintance does not cause a change. To support this thesis, I will argue for a minimalistic approach for understanding referential terms. I will argue that confirmation may increase our knowledge and understanding of certain objects, but our semantic understanding of the term is unaffected by such advances.

I will argue that what it means to understand a referential term is context-dependent, and speakers can grasp referential terms in a variety of ways. Understanding a proper name requires that the speaker is in *some* relation to the denoted object, but there is no particular relation that must hold. In fact, the different cases of reference without acquaintance from above demonstrate this.

Perceptual Acquaintance This is perhaps the simplest case. A speaker *S* understands a referential term *t* which denotes object *o* if *S* is perceptually acquainted with *o*, and *S* knows that *t* refers to *o*. If I introduce my friend Jane to my spouse, and utter something like "This is Jane," my friend will

now have an understanding of the name ‘Jane.’ Here, it is useful to include knowledge derived from memory as well, even though memory is not perception *per se*. That is, after departing from the scene, if my spouse has the memory of Jane, he understands the name ‘Jane.’

Participation in Causal Chain Another way to grasp a proper name is to be a part of the relevant causal chain. As I discussed above, participation in a causal chain gives an agent the *ability* to refer to an object. In most of these cases, what will become clear is that to understand a name is the ability to use the name to refer to its referent. As such, so long as I am able to refer to an object, I do not need anything further to understand the name. Like Kripke and others have shown, everything I know about, say, Gödel can be false, but I can still successfully refer to Gödel.

Participation in Linguistic Practice This third way of understanding is related both to abstract objects, and also hypothetical objects.

I will now discuss both these cases. Like I argue in the other chapter, some abstract objects depend on certain social and linguistic practices for their existence. Examples of this are states, works of art, and fictional characters. Likewise, to understand a term that refers to an abstract object, one can simply participate in the social practice. Take Sherlock Holmes for example. Since Sherlock Holmes is not a physical object, we cannot be perceptually acquainted with it. So, the relation between a speaker and Sherlock Holmes has to be a different one. Remember that fictional characters are created by communities of speakers. Their existence is maintained also by the continuance of the linguistic practice. That is, for Sherlock Holmes to persevere, there needs to be communities of readers of the Sherlock Holmes stories who generate discourse about Sherlock Holmes.

What, then, does a speaker need in order to understand the name ‘Sherlock Holmes’? Given that the only ‘interactions’ we have with abstract entities are through social practices, participation

in these practices are sufficient for a speaker to understand the term. After all, a reader of the stories will be perfectly able to refer to Sherlock Holmes using the name ‘Sherlock Holmes’. A reader will have the ability to contribute to discourse about Sherlock Holmes, which in turn allows her to refer to him, and thus we can ascribe understanding.

Note, however, that participation in the relevant social practice is not the only way to gain understanding of the name. Someone who has not read the Sherlock Holmes stories can still appropriately use the name ‘Sherlock Holmes,’ if the person is a part of the relevant causal chain. Causal chains apply to abstract objects, as well. I, for example, have never read Anna Karenina, but through my friends who have read the novel and who are thus participants in the social practice, I gain the ability to refer to Anna Karenina. I can do this even though I know very little about Anna Karenina. My epistemic status about Anna Karenina is irrelevant to my ability to refer, and it is sufficient for my semantic understanding of the name ‘Anna Karenina.’

In a somewhat similar way to abstract objects, understanding of terms that refer to hypothetical objects may also come from following linguistic practices. Hypothetical names are often generated by DRF, so that not only are speakers not perceptually acquainted with the object, they are often unsure whether the object exists. None of this matters for the understanding of names. In fact, I will argue that requiring either a perceptual acquaintance condition or a knowledge of existence condition contradicts how scientific inquiry is undertaken.

First, what does it mean to say that understanding a hypothetical name requires participation in some social practice? More specifically, what are the relevant social practices one needs to participate in? To answer these questions, let’s look at why speakers choose to introduce hypothetical names. Usually, the existence of an object is hypothesized in order to explain a certain set of data, or certain phenomena. This is what happened with the Neptune case, where Le Verrier

concluded that there must exist a planet in order to explain the perturbation in the orbit of Uranus, which he observed and recorded. Similarly, the kind phlogiston was postulated to exist so as to explain how combustion takes place. So, these names are parts of different theories. In this context, we should understand theories not as sets of propositions, but as research projects that a scientific community undertakes[13]. The relevant social practice, then, is the research project of which the hypothetical object is a part.

A speaker understands a hypothetical term when she participates in the research project of which this name is a part. Once again, the ability to successfully use the name comes from the speaker's participation in the scientific practice. The ability, in turn, is what gives the speaker the understanding of the term, even if the speaker has not been acquainted with the object. In this way, as the reference of the term has been fixed by some speaker (it is important to note that the act of fixing the reference is itself a part of the scientific process), the meaning of the term thus depends on community consensus. In the case of Neptune, for example, an astronomer's participation in the scientific project suffices to give her an understanding of the term 'Neptune.'

The 'understanding as the ability to use" view may contradict some intuitions, however. For example, an objector may argue that speakers don't fully understand hypothetical names until the hypothesis is confirmed and there have been speakers who are perceptually (or otherwise) acquainted with the object. After all, there seems to be a difference between using the name 'Neptune' before the hypothesis was confirmed, and using it after. Perhaps the speakers who use the name after confirmation have a 'clearer idea,' or a clearer understanding of the name 'Neptune.' The confirmation of a hypothesis elevates our epistemic status with regards to the objects of hypothesis, after all.

[13] Similar to Lakatos or Kuhn's understanding of scientific practice.

I think this would be a misguided argument. It comes from a conflation of two senses of understanding: that of a name and an object. Understanding a name is a matter of semantics, basically to understand the semantic contribution of a name in a proposition. This, I have been arguing, is the ability to use the name. If one can appropriately use a name in a sentence, then one knows what it is that one is asserting about the world. The semantic understanding demands minimal effort from speakers. The other sense of understanding here is that of an object. This is what happens after confirmation. When we confirm a hypothesis, we deepen our knowledge of the objects of the hypothesis. We may observe them (perceptual acquaintance) or we may come to know more facts about them (epistemic acquaintance). So, our knowledge of the object changes. But note that this is an epistemological change, and not a semantic one. Understanding a name is not nearly as demanding as 'understanding' an object which is the point of scientific inquiry.

Moreover, it is important to attribute understanding to speakers before confirmation for another reason: if a scientific community is looking to confirm or disconfirm a theory, members of the community must know what it would mean for the theory to be true or false. As a result, if an astronomer is aiming to observe Neptune, they must understand the term 'Neptune,' at least as the object of their inquiry. Note that this does not require the *knowledge* of any facts about Neptune, nor does it require the knowledge of its existence. But if speakers were unable to understand terms like 'Neptune' before the relevant theories are confirmed, then scientists would not have an understanding of the theory they are trying to confirm – and hence have no way of confirming or disconfirming them. Therefore, unless we ascribe understanding to speakers, we are led into a contradiction about what makes scientific inquiry possible.

So, speakers must have an understanding of the term, even if they know very little about the object itself. This, I think, is a good motivation to adopt a conception of semantic understanding

that is minimally demanding. As long as one participates in the practice of scientific theorizing (or is causally connected to someone who is), one has all the understanding one needs in order to successfully use the name. It's important that to grasp the meaning of a referential term demands very little — if our condition for understanding was too demanding, it would be impossible to get any scientific project off the ground. Other conditions, like perceptual acquaintance or some epistemic ability to tell the object apart from others would impede scientific practice.

Chapter 2: Reference and Thought in Fiction

In this chapter, I will discuss fictional characters and how we refer to them. First, I will review some competing theories for the metaphysics of fictional objects. Then, I will present my version of fictional realism, which is that fictional characters are abstract artifacts that come into existence when there is a community with a linguistic practice which makes use of the characters. In other words, I will tie the existence and creation of ficta to communities of readers. This emphasis on community serves to highlight the social role of abstract artifacts.

Next, I will discuss two vagueness problems that arise from this account. The first vagueness problem I will discuss regards the imprecise boundaries of coming into existence: if ficta begin to exist after the establishment of linguistic practices, how can we pinpoint a determinate boundary to their existence? I will argue that this is not a problem for ficta alone, but for all artifacts, including physical ones, like buildings, sculptures, and even beef stew. The second problem is pointed out by Anthony Everett (2005), which regards characters that are intentionally written to exist indeterminately, or be indeterminately identical. I will respond by saying that such cases are where authors fail to create characters. Finally, I will argue that speakers gain the ability to refer to fictional characters by participating in the social and linguistic practices by which ficta are created and maintained.

Background and Alternative Views

Fictional names refer to fictional entities. But just what kind of things are these? It is clear that, for example, Raskolnikov is not a real person. Examples like this are what motivated philosophers to hold that fictional names are empty. Neither are fictional entities merely possible but real people as David Lewis (1978) once argued. According to possibilism, each work of fiction

describes a possible world, and the characters from the work are people (or other creatures) that exist in those worlds. There are many reasons why possibilism about fictional objects is implausible, but here's a simple one: fictional characters are always incomplete, that is, they are not (and arguably cannot be) described fully. This is so because fictional characters are metaphysically incomplete, that is, the characters can be 'completely described' in multiple different ways. Non-actual possible people, on the other hand, are metaphysically complete. For a real person, it is determinate whether that person has a particular property. This is not the case with fictional characters, since an author may be noncommittal about certain features of a character. As such, fictional characters and non-actual possible people have distinct statuses with respect to metaphysical completeness.

Even if we disregard this, there remains an issue of referential ambiguity. For each metaphysically incomplete fictional character, there corresponds an infinite number of non-actual possible people, because the 'gaps' in the incomplete characters can be filled in infinitely many ways. Fictional characters have some definite properties in virtue of the texts in which they appear, but there are some properties for which there is no fact of the matter whether they have them or not. Take Sherlock Holmes. He has the definite property of being a detective, but it is not definite whether he has ten toes. We can reasonably assume that he does, but our attribution of this property is not determined by the text in the way that the property of being a detective is. Then, there are many merely possible individuals who have all the definite attributes of Sherlock Holmes, but who have eleven toes, or twelve, thirteen, etc. Hence, if fictional characters are non-actual possible people, there are an infinite number of people who are Sherlock Holmes, an infinite number of Doctor Watsons, and so on for every fictional character. A consequence of this view is that fictional names have an infinite number of referents, especially given that David Lewis was a realist about

possible worlds. This runs against common sense intuitions about naming and reference. Considerations such as this make possibilism an implausible view.

An alternative to possibilism is that fictional characters are actual, but abstract objects. Abstract entity theorists divide into two types: neo-Meinongians and artifactualists. Neo-Meinongians see abstract entities as timelessly existing (Parsons 1980). One main issue for neo-Meinongians is the creation problem: since abstract entities are non-spatiotemporal and non-causal, and since creation is a causal process, the idea that one can create an abstract entity is incoherent, or so the problem goes.

The idea that creation must be a spatio-temporal causal act has been recently criticized. Irmak (2021) argues that ontological dependence is a better way to understand creation than a causal process. Moreover, there are other problems with neo-Meinongianism. This view holds that abstract entities are not created, but they exist timelessly (or subsist if we want to make that distinction). So fictional characters are not created by their authors. The role of the author, under neo-Meinongianism, is thus not to create fictional characters (though they may create fictional works), but to discover them. Doyle's writing of the Sherlock Holmes stories picked out Holmes among infinitely many other conceivable fictional characters, but it did not bring him into existence. This view is not incoherent; however, it seems committed to a counterintuitive conception of art. We think of art as a process of creativity, not discovery. Perhaps a Meinongian approach is appropriate for certain kinds of abstract objects (like numbers or sets), but fictional characters intuitively seem like they are created by authors or storytellers.

Artifactualism

Artifactualists, on the other hand, hold that fictional characters are abstract entities that depend on real entities for their existence (entities such as people, but also language, books, etc.)

We create such entities by certain social practices (Kripke 2011). These include storytelling, novel-writing, and the like. When Dostoevsky wrote *Crime and Punishment*, he created Raskolnikov, among other characters. Raskolnikov is not a physical object, nor is it a purely mental one. It is not a particular mental representation of the book's protagonist. If that were the case, there would be issues regarding privacy and reference — if the referent of 'Raskolnikov' is a mental representation, like an image in a reader's mind, then each speaker is referring to a different object. In that case, we can call into question whether the name is intelligible at all. The abstract entities of artifactualists don't have this problem since they are public objects. So artifactualism ensures that fictional names are intelligible. In fact, a major advantage of artifactualism is that it allows for a straightforward semantics of fictional names. According to artifactualism, ficta exist, so fictional names simply refer to existing (though abstract) objects.

Abstract entities depend on a few things for their existence. First, they depend on their creators. Second, they depend on copies of the relevant works and people who have read these works or are otherwise acquainted with its characters. So, the existence of Raskolnikov is a dependent existence; if we eradicate all copies of *Crime and Punishment*, destroy all available information about it, and make everyone forget about it, Raskolnikov would no longer exist.

To highlight some features of abstract artifacts, as opposed to 'eternal' objects, think of another abstract entity that is an artifact of social practices: states. The United States is an abstract entity. Note that we can also use 'the United States' to denote a place, which is not an abstract entity. The governing and legislating body, however, is an abstract entity that did not exist prior to certain social and political practices.[14] I will make two observations on the state example which will carry over to fictional characters. First, notice how the United States was created by public

[14] Social practice in a wide sense here. Persons coming together to create institutions and laws to live by is considered here a social practice.

agreement (though clearly the English had an initial disagreement over it). Thus, it would be incoherent to say that the United States, as an abstract entity, always existed, but the Founding Fathers 'discovered' it. At least from the perspective of our ordinary historical discourse, the United States was brought into existence as a result of political and social activity.

Second, related to the first observation, the existence of the United States is dependent upon human beings and social practices. Its existence is conditional; if humankind was wiped off the face of the earth, the United States would no longer exist. This condition is even more demanding since the mere existence of humans is not sufficient to secure the existence of the United States. As a social and political process, the United States may go out of existence. History is full of states that no longer exist; for whatever political or social reason, the 'social contract,' or the social and political practices that sustain the existence of a state, can be dissolved. If that were to occur, we would say that the United States no longer exists. So, there are at least some abstract objects that can go in and out of existence, even if some abstract objects exist timelessly (like numbers, for example). According to artifactualism, fictional characters are like the United States in these important ways. They depend on the existence of agents, as well as the social practices of these agents.

Existence Conditions and Creation

I will argue that fictional characters begin to exist when there is a community of readers with a linguistic practice that employs reference to the character. There are other possible accounts for creating ficta. The first is that a fictional character is created at the time of its mental conception. According to this view a fictional entity begins to exist as soon as the author conceives the character. This mental conception can be the author thinking to herself something like "The main character's name will be Jane and she is a lawyer." The second view is that a character is created

when a sentence is inscribed, typed, or uttered that contains a reference to that character. For non-verbal fiction, this may be the first instance when a character appears visible, on screen or on paper. I think these two accounts are inadequate, though mental conception and writing (or uttering, etc.) play an important role in the process of creating fictional characters.

To see why the two views from above are inadequate, and why *community* plays an important role in creation, consider the following example. Say I create a neural network that, if we provide enough short stories, can generate short stories that follow the general structure of the genre. Assume that the network is advanced enough to produce coherent English sentences and narratives. We also provide the network sets of names for characters, places, etc. Each time the network generates a story, it randomly picks the names and appropriately uses them in sentences. For the analysis I will provide, also assume that, despite the advanced capabilities of this network, it is not, in any sense of the word, a conscious entity[15].

Now let's suppose I generate a text and save it on my computer without having read it. So, there is a text which would be deemed a short story by any competent speaker, but at this point no one has set eyes on it. Setting aside the question of whether the story was created[16], would we say that the characters of the story are created? I don't think we would, since nobody is acquainted with the story. Abstract objects are mind-dependent, and if there are no agents that are conscious of the characters, they cannot exist. Remember that neither the neural network, nor my computer that contains the text, are conscious beings. So, the mere existence of a text is not sufficient to 'initiate' the existence of a fictional character. An analogous case to this would be if I, a conscious person, wrote a novel, destroyed it without showing it to an agent, and then completely forgot

[15] The inspiration for the thought experiment came from Ackerman (2021), although I'm using it to make a different, if not completely opposite, point.

[16] It looks to me like the story is created here, even though nobody has ever read it yet because the text exists.

about it, or died. Since nobody's conscious of any fictional characters from my hidden novel, they do not exist.

An interesting question here is whether the fictional characters came into existence and then ceased to exist after my amnesia, or they never came to be in the first place. I don't think there is a one-size-fits-all kind of answer here, but we can sketch a general picture. Once again, let's look at the definition of abstract objects. We defined abstract entities as mind-dependent entities that are created as a result of certain social practices (Thomasson 2009). Storytelling is one of them. Since abstract entities come to be as a result of practices, it makes sense to say that abstract entities are created only after a certain social practice is established. For example, say I write a novel, where the protagonist is named 'John Smith.' The writing processes of novelists differ, but say I first conceived of the idea in my head. I likely changed my mind about it many times before and during my writing process. Eventually, I completed the novel. Since at this point the novel is complete, it may be tempting to say that it is now that John Smith comes to exist. I do not find this a satisfactory answer. A major motivation for introducing abstracta into our ontology is their role they play in social practices (think the role states play in our lives). In other words, abstracta have purpose. At the time of my completion of the novel, the fictional object does not have a purpose to fulfill. There is no established practice with regards to the use of the name. I, the author, am the only person who uses the name. It is even doubtful that I ever use the name, at least in the book, referentially. At least some of my first uses initiated the process of creation of the object, as opposed to referring to it. Furthermore, as there is so far no audience of the book, there is no community with norms regarding the use of the name[17].

[17] Except, perhaps, for the trivial community that includes me and my future and past selves.

I think fictional characters begin to exist later, when the story is distributed, and there is a readership of the story, and especially when there is discourse surrounding the story. When stories reach their audiences, there is discourse about the book and its characters. Readers talk about the characters in different ways: they may recite the story to others, or make critical judgments about the characters, or compare one character to another, etc. When these kinds of discourses arise, the community of readers begin to employ names of the characters in the sentences that they use. As linguistic practices emerge, speakers start to refer to characters. The uses of names from the story by a community of speakers establishes the social practice, and that is when the fictional characters begin to exist.

Note that it is difficult to point out to a specific moment at which fictional characters begin to exist, at least in a way that is not arbitrary. The boundaries of existence for fictional characters are vague. They also vary from case to case. Some novels are instant hits, some are often stuck in small circles before they reach a wider audience[18]. But at some point, fictional characters begin to exist when there is a community of speakers with established use practices for their names. This view entails a few important consequences. First, the earliest uses of a fictional name may be non-referring. When the author is first telling or writing the story, the name does not refer. Rather, the use of the name is a part of the creation process of the object. Second, this means that even the non-authorial uses in the beginning fail to refer, at least until there is something we can plausibly call an established practice.

Here one may object to my view, since the truth value of a sentence like "Sherlock Holmes is a detective" should not really depend on when it is uttered. However, my view implies that the

[18] It is also conceivable that an author writes a novel, never publishes it, and prevents anyone else from reading it, and yet uses the names of characters with the intention to refer (maybe she is talking to herself). This is a bizarre case, but not impossible, and it may be reasonable to accept that in this case there is a norm that governs the use of the names, and the social practice is established.

sentence would likely be either false or nonsensical if it were uttered around the time the stories were first published. This may seem counterintuitive, but I think these first few uses contribute to the establishment of the practice, and consequently to the creation of the object.

Think of Wittgenstein's example of 'slab' as a paradigmatic linguistic practice (Wittgenstein 1953). For that language game to work, the person who wants the slabs first needs to instruct the other person to respond to the utterance 'slab!' in the appropriate way. So, the first few uses of 'slab!' are instructional, and they serve the purpose of establishing the norm that governs the speakers' behavior. We may call into question whether these initial uses have the same meaning as the latter ones. Their use, at least, is different. The case of the fictional names, and the sentences that contain them, is similar. The initial uses of the name lack a referent, so the sentences that contain them do not express a proposition. Moreover, in the fiction case, the first use of the name is likely when the author is writing the story, which also contributes to the creation of the character, rather than referring to it.

This brings us to another important point, which is that we should understand the creation of fictional characters not as sudden events that take place as a writer wills for some character to exist, but rather as processes that start with the author's conception of a character and is completed when there are enough people talking about it to establish a use practice. It is difficult to draw precise boundaries to this process. But once there is a linguistic practice in place, the fictional character exists. I will treat this vagueness problem in the next section.

The Temporal Vagueness Problem

A rightful worry regards the vague boundaries of existence, since it may not always be clear when a linguistic practice is established. The problem is that accepting a process of creation for ficta commits us to a certain period in which the existence of the character is indeterminate,

since we cannot designate a particular time at which a linguistic practice is established. That means ficta exist indeterminately at least for a while, which is admittedly counterintuitive. Moreover, if ficta can be destroyed, which I think is possible when linguistic practices gradually fade away, there is another period of vague existence pre-destruction. A theory that is committed to vague existence should provide good reasons to accept it. I will not do that here. However, I think this problem is not peculiar to fictional characters, or even abstract objects, but artifacts in general.

To be clear, the problem at hand isn't that ficta *in general* exist vaguely. Rather, if my view of creation is correct, then there is a certain period of time during which vague existence is an issue. Before Doyle started writing the Sherlock Holmes stories, Holmes determinately did not exist. Today, the character determinately exists. However, in between these two points in time, there is a period, say t_m-t_n, where the existence of the character is indeterminate, because it is unclear whether the process of creation is complete. Before t_m, Holmes determinately did not exist; after t_n, he determinately does. The issue is the period in between.

As mentioned above, however, no other artifact is immune to this problem. Think of a beef stew. There is a process of making stew, sometimes a rather long one. Since stew is the end product of this process, we do not say that the stew exists when, say, I sliced the vegetables. Nor would we say that the stew exists after we put it all together and just let it cook. The uncooked combination of ingredients is not yet the end product we crave for. Similar to ficta, the stew also does not have a specific point at which it is cooked. Even if we stipulate a cooking time, say two hours, this is an arbitrary boundary. The stew would not be significantly different at one hour and fifty-nine minutes and forty-five seconds. There is simply no specific point at which the process is complete, which entails that there may be a period of time where the stew exists indeterminately. Likewise,

when we think of artifacts that are decaying (like old buildings), there may be periods of indeterminate existence.

In practice, we deal with this indeterminacy rather smoothly in our ordinary discourse. If my friend asks whether the stew is done, I may simply answer "kind of!" I doubt my friend would raise metaphysical concerns about my answer, despite the vagueness of my response. The problem of vague existence with regards to creation, then, is a problem for all sorts of artifacts, abstract and concrete. Since few of us are skeptical about the existence of stews, buildings, sculptures, and other artifacts, we do not have a reason to reject the existence of fictional characters.

Some, like Evnine (2016) and van Inwagen (1990), find it plausible that abstract artifacts may have vague existence. I don't think fictional realists need to go that far. An alternative solution to this problem is to diagnose properly the source of vagueness. According to Friedell (2017), if we accept a plenitudinous ontology, we can avoid committing vague existence. His main idea is that, following Fine (1982), for each property an object has, there is a qua-object that is composed of that object. "For instance, […] Biden-qua-president and Biden-qua-spouse are two qua-objects, each composed by (but distinct from) Biden" (Friedell, 2021). In the case of artifacts, we may adopt this strategy to say that while the 'objects' exist, it is indeterminate whether 'object-qua-stew' or 'object-qua-Sherlock Holmes' exists. Given that the qua-objects supervene on the properties, this view is only committed to vague properties, not vague existence.

I think this view is inadequate, though it is a step in the right direction. There are two main problems with it. First, there is a worry that this may not be applicable to abstracta, even if it works for concrete artifacts. Korman (2014, 2015) points out the fact that while for concrete artifacts, there are (concrete) intermediate objects, there are no such intermediate abstract objects. Let's look at the stew example once again. When I'm cooking the stew, though it may be vague whether the

stew is ready, it is not vague whether there is ‘something’ that is about to become the stew. These things may include potatoes, carrots, beef chunks, the stock, etc. The existence of these objects, and their combination, is not vague at all. So when we talk about ‘object-qua-stew,’ there are physical entities that constitute the ‘object.’ There is no analogue to this for ficta: what can the ‘object’ be in ‘object-qua-Sherlock Holmes,’ if the object is not yet Sherlock Holmes? The intermediate stage seems to be something obscure, and one may rightly worry about this.

The second problem is that, while the plenitudinous ontology view looks like it’s locating the vagueness in properties and not existence, the ‘object-qua-property,’ insofar as we accept it as a distinct object in our ontology, still suffers from vague existence. So Friedell’s proposal does not solve, but only postpones the vague existence problem. What I think this line of thinking gets right is its underlying idea that the vagueness issue is one about properties. In the stew case, it is not vague whether ‘something’ exists, but it is vague whether that something is a stew. This solution, like I acknowledged above, is not readily available to abstract entities, but I think the right solution will follow the same idea. The important question here is to give an account of what the ‘something’ is for abstract entities in the making, but I do not have the space to treat this question adequately here.

Vague Existence and Identity

Everett (2005) discusses two vagueness issues for ficta: vague existence and vague identity. I will start by discussing the former. Everett argues that there are fictional characters that, according to the texts they appear in, and their authors, exist indeterminately. His example is Tatyana Tolstaya’s novel *The Slynx*. He argues that, in the end, it is left open whether the novel contained a Slynx. Since it is indeterminate whether there is a Slynx in the novel, a fictional realist must commit to the indeterminate existence of the fictional character. I will argue that *The Slynx*

is a limiting case for fictional character creation, and that a pretense theory is better suited to handle that case. Although I will concede to Everett's argument that much, I will argue that this is actually consistent with fictional realism, because my account of artifactualism makes use of pretense in establishing linguistic practices.

To begin, let's recreate *The Slynx* case. In order to avoid getting caught up in the details of Tolstaya's novel or its rival interpretations, I wrote a short story to make the same point:

> *The Slunx*: It is indeterminate whether the Slunx exists. Jim and John have incompatible beliefs about the Slunx's existence. Jim has evidence that it exists, and believes that it does. John disagrees. He thinks the Slunx does not exist, and he also has good evidence for his belief. As it happens, both of their beliefs are true, though not determinately true.

In *The Slunx,* the existence of the Slunx is indeterminate. This is given to us by the story. Jim and John have opposite beliefs, and they are both true, again, according to the story. So the world of *The Slunx* is one that the law of no contradiction does not hold, since it is both the case that the Slunx exists or the Slunx does not exist. We can write variations of this story where, instead of violating the law of no contradiction, we violate the law of the excluded middle. For the purposes of my argument, it does not matter how we are deriving the indeterminate existence.

Everett's second challenge appeals to vague identity. Although this is a distinct problem, my response to both of Everett's challenges will follow the same line. Everett argues that fictional realists must commit to vaguely (or indeterminately) identical objects, as well as vague existence.

He points out that authors may write characters that are, by authorial design, vaguely identical. To demonstrate this point, he writes a short story:

> Frackworld: No one was absolutely sure whether Frick and Frack were really the same person or not. Some said that they were definitely two different people. True, they looked very much alike, but they had been seen in different places at the same time. Others claimed that such cases were merely an elaborate hoax and that Frick had been seen changing his clothes and wig to, as it were, become Frack. All that I can say for certain is that there were some very odd similarities between Frick and Frack but also some striking differences.

In this story, it is indeterminate whether Frick and Frack are identical. There is no fact of the matter, at least according to the story or Everett, regarding the identity of the two characters. Fictional realists, Everett argues, are committed to the indeterminately identical objects Frick and Frack. He then appeals to Evans' (1978) argument against vaguely identical objects and concludes that fictional realists are committed to contradictions.

One possible objection against Everett's argument is by arguing that the indeterminacy in question is not metaphysical, but epistemic. One can argue that it is not the ontological status of the Slunx that is indeterminate, but our knowledge. We are uncertain whether it exists. Similarly, we are uncertain whether Frick and Frack are identical. I don't think this response is satisfying. In *The Slunx*, it seems clear that the problem isn't that the audience isn't given enough information to determine the fact of the matter. Rather, it looks like there is no fact of the matter whether the Slunx exists or not. One could argue, albeit in an extremely *ad hoc* way, that there still is a fact of the matter that is known to the author, but that the author is hiding it from the audience. This still falls short of an explanation. As the author of *The Slunx*, I can assure the reader that I am non-

committal about the existence of the Slunx. There is no sense in which I "know" the truth of the matter. A similar case can be made for *Frackworld*, where the identity in question is vague according to the story. In cases like these, I think Everett is right that the fictional realist is dealing with a problem of ontological indeterminacy, and not an epistemological one.

I think a better response to Everett's challenge is to concede, but with a caveat. I think examples like the Slynx, and the Slunx, are cases where authors *fail* to create fictional characters. Discourse that refers to these characters only pretend to refer to the objects. This is not, however, a complete defeat for fictional realism, like Everett suggests. Remember that my account of fictional character creation relies on a process by which social and linguistic practices are established. During this process, we make use of pretense. We do so by pretending like we refer to existing objects, or pretending like we make truthful assertions. These acts of pretense help establish the practice. In cases of successful creation, there is a (vague) point at which we stop pretending, when the relevant practices are put in place. We can call this the "fake it till you make it" view: for a while, the reader community 'fakes it,' or pretends to refer to fictional characters[19], until there is a linguistic practice established, at which point we 'made it'. In the Slynx/Slunx cases, because the authors failed to create the characters, we never move past the 'fake it' phase, and in these cases, a Waltonian theory of make-believe is a better way to understand talk of these indeterminate characters. In such a case, we merely pretend to commit to indeterminate existence, although no such object exists.

There must be, then, some conditions the author needs to satisfy so as to succeed in creating a fictional character. Here I will give a sketch of what some of these conditions may be. First, the Slynx/Slunx cases show us that the author must successfully undertake some creative act (writing,

[19] This pretense is not limited to referring, but encapsulates all sorts of discourse, attitudes, behaviors, etc.

uttering, etc.) that posits the existence of a character, determinately. In *The Slunx*, I failed to do this, since I intended to write about a character whose existence is indeterminate. In cases where a character's existence is not indeterminate, the author intends[20] to and successfully writes a character who determinately exists in the work. This applies to other inconsistencies as well.

Here, I think it is necessary to make a distinction between weak inconsistencies and strong inconsistencies. A weak inconsistency is when a character has inconsistent properties. Say, a character is attributed different eye colors at different parts of a novel. Since this is an inconsistency with regards to particular properties, accepting the existence of the character does not necessarily commit us to a logical contradiction. Such inconsistencies are easy to explain away. To account for this inconsistency, we may say that the character's eye color changed throughout the book, or it is vague whether the character's eye color is either of the colors, etc. Or we simply defer to the author or be non-committal about this part of the work. We may even say that the author failed to attribute the property, but still successfully created the character.

A strong inconsistency regards existence, identity, truth, etc. Strong inconsistencies cannot be explained away unlike their weak counterparts, and if we were to commit to the existence of strongly inconsistent characters, we would be committing to logical contradictions. The Slynx/Slunx cases are contradictions about existence. *Frackworld* is an example of a contradiction about identity. In the former case, if we were to accept the Slunx into our ontology, then we would be committing to the truth of "The Slunx exists and the Slunx does not exist," which is an apparent contradiction. The Frackworld case commits us to a contradiction about identity. As such, when fictional works contain strongly inconsistent characters, there is a failure in character creation.

[20] It is actually doubtful whether authors explicitly have this intention. Since most authors do not want to create vague characters, it is most often the authors that write about vague characters that need to explicitly intend to do so.

Other such failure conditions may include some completeness criteria, that is, the work must be complete for an author to be successful. Of course, there are incomplete works that are published and reach audiences, but if I start writing a novel, and decide halfway to burn everything I wrote and abandon my project completely, my characters fail to come into existence. Another way in which an author may fail is if her novel is extremely unpopular. If there is no wide enough readership, and the people who read so disliked the book that they never talk about it (and even want to forget they ever read the book), then we can say that the character fails to come into existence, since no relevant linguistic practice is established. This is by no means a complete list of authorial failures, but my aim is to show that there are cases where one can fail to create ficta.

In cases of authorial failure to create characters, we make use of make-believe to continue our discourse. For *The Slynx*, for example, the sentence "The Slynx exists" is false, when we take it literally. This is because Tolstaya failed to create it[21]. We can still talk about the Slynx, but we do so by pretending that there is such a character that exists indeterminately. Same would be the case for the Slunx, if it were widely read. Similarly, when we talk about Frick and Frack, we employ make-believe, since Everett's characters contain a strong inconsistency regarding identity, and consequently have failed to come into existence. Fictional realists, then, are not committed to vague existence, but they only pretend to be committed when there are characters like the Slynx.

Referring to Fictional Entities

I will now turn to the question of reference with regards to fictional characters. An advantage of artifactualism, and fictional realism in general, is that accepting ficta in our ontology allows for a straightforward semantics. Since fictional characters exist, names like 'Sherlock

[21] This is, of course, not a judgment on Tolstaya's work. The failure here is not an artistic, but a metaphysical one.

Holmes' and 'Gandalf' are not empty, non-referring names. Instead, they refer to actual objects — the abstract objects created by readers.

However, given the non-physical nature of abstract objects, we are once again faced with a problem about reference without acquaintance. Abstract objects are such that we can never be perceptually acquainted with them. Contrast this to the hypothetical cases. Although speakers started referring to Neptune prior to observation, the astronomical community eventually observed it, meaning some of them were perceptually acquainted with it. The abstract object case is starkly different from cases like Neptune in this respect: no speaker is, and ever will be, perceptually acquainted with Sherlock Holmes. This is not a problem, though. My account of reference in the hypothesis case did not rely on the possibility of acquaintance, and it works just as well when there is no such possibility. In fact, in order to account for the full range of scientific theorizing, it must work: think of quarks, for example, which are unobservable — if the possibility of acquaintance was a prerequisite for reference, then we could never meaningfully talk about quarks.

It is just as well with abstract objects. Our ability to refer to them does not hinge on any sort of perceptual acquaintance. Instead, it relies on our participation in linguistic practices. Remember that fictional characters ontologically depend on social practices. Practices such as storytelling, publishing, discussing, etc. are what makes it possible for fictional characters to come into existence, and maintain their existence. The other side of this coin is that, by being a part of these practices, speakers gain the ability to refer to fictional characters.

A result of my account of fictional character creation is that the naming practice precedes the creation of the object. At most, the character is named at the beginning of a process by which the character is created. This presents a problem about reference: before the object comes into

existence, what is the name referring to? Furthermore, given that my account accepts a certain period of vagueness, when is reference fixed in a fictional name?

I think there are two possible ways to answer this question. The first parallels my remarks about the truth values of sentences from above. If the initial utterances of sentences are not assertive, but rather plays a role in establishing the linguistic practice by which their truth is judged, then similarly, the initial uses of a fictional name are not referential either. We start using the name so as to establish the practice and create the character, and we thereby get the ability to refer. This is, again, a 'fake it till you make it' view, this time about reference.

Alternatively, we could appeal to the mental files theory again. Just like how descriptive reference-fixing 'opens' a mental file, we could say that introducing a name in a fictional context does the same. It does not matter that the object does not yet exist — we can form our temporarily false (or meaningless) beliefs in said mental file, and once we successfully create the character, the truth value of some of our beliefs undergo a change.

I think the first answer, the 'fake it till you make it' view is more plausible, given the metaphysical vagueness that surrounds fictional characters. The mental files theory runs into issues about reference shifts. If I had a belief that "Sherlock Holmes is a detective" which was, for a while, meaningless, and then I suddenly had a true belief that "Sherlock Holmes is a detective," I am not sure if we can maintain that the propositions expressed by the two sentences were the same to begin with. As such, I favor the 'fake it till you make it view.'

Finally, I would like to talk about the causal chains that can still allow reference to fictional characters. Although participating in the community practices is the primary way to secure reference to a fictional character, it is not the only way. Causal links still play an important role in enabling speakers to refer. Not everyone who refers to fictional characters are participants in the

social practices. To reiterate my former example, I have never read Anna Karenina. I don't often talk about Anna Karenina either. However, when I do, I successfully refer to her, since I am causally linked to people who have read Tolstoy's novel. Thomasson (1993) talks about chains of communication and chains of publication — these chains connect speakers to others who are actively engaged in the practices (what Evans and Thomasson call 'producers' of the name), thereby giving them the ability to refer.

Conclusion

I argued that fictional characters are abstract objects. They are created as a result of social practices, and they 'come into existence' only when certain social norms are established. This account of the ontology of fiction gives rise to (at least) two vague existence problems. I argued that the first problem, which pertains to creation and time, is one that all artifacts are susceptible to, whether they are concrete or abstract artifacts. The second problem, brought up by Everett, is about characters that indeterminately exist or are indeterminately identical according to fictional works. I made a distinction between strong and weak inconsistencies, and concluded that characters that contain strong inconsistencies, like indeterminate existence or identity, fail to come into existence. In order to refer to fictional characters, and have the ability to form singular thoughts about them, I argued that speakers need to participate in the relevant social practices, or be causally linked to certain participants.

Conclusion

I have argued that acquaintance is not necessary for reference or singular thought, using examples from science and fiction. In science, I have discussed cases like Neptune where the reference of terms is fixed not by a baptism, but by stipulation. Scientific hypotheses often require such naming practices, especially when a theory posits the existence of an object, or a kind, to explain a set of data. I argued that descriptive reference-fixing is a legitimate way to establish reference, since scientists meaningfully talk about their hypotheses, and that participating in scientific practice is sufficient to understand the meaning of such terms, since scientists understand what their theories mean. Both of these things are possible without acquaintance. Therefore, I conclude, reference, singular thought, and understanding are possible without acquaintance.

Similarly, I have argued that we refer to fictional entities. I gave an account of the ontology of fiction that accepts fictional characters as abstract objects that are created by community practices. I responded to two vagueness problems that arise from my account of fictional realism by discussing similar cases of ontological vagueness for physical objects, as well, concluding that such vagueness issues are not sufficient to deny existence to fictional characters — unless we deny existence to many other artifacts. Then, I suggested that in order to refer to fictional characters, speakers need to either participate in said community practices, or be causally linked to others who are.

The two cases, science and fiction, display remarkable similarities, as well as differences. They are both contexts where referring to objects without having been acquainted with them is essential for our linguistic practices. In both cases, the linguistic community plays an important role in allowing us to meaningfully talk about the objects of interest. In science, the community

consensus fixes the reference of terms, whereas in fiction the community practices create the object, which in turn makes our discourse meaningful.

However, the differences between the two are as important as their similarities. What is remarkable about fictional discourse is that it leads to the creation of abstract objects; the reference of fictional names is fixed through this creative process. So, the community causes a metaphysical change: it creates a new object. In science, although the community plays an important role in making our discourse meaningful, it doesn't thereby create anything new. In scientific practice, the processes of hypothesizing, observation, and confirmation cause a lot of shifts, but none of these are metaphysical. Instead, such shifts concern our knowledge of things in the world, so they are epistemological shifts.

In this thesis, I gave an account based on the causal theory of reference that explains the linguistic intricacies of scientific practice and fictional discourse while also paying attention to the close relationship between reference and thought. These two are by no means the only philosophically interesting cases that a theory of reference should account for. However, I think the specific mechanisms that allow for reference and singular thought will be particular to the immediate context, which I what I aimed to show in my project.

Bibliography

Ackerman, Greg. “Securing Singular Thought about Merely Hypothetical Entities.” *Source: Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 173, no. 8 (2016): 2193–2213. https://doi.org/10.1007/sl.

Braun, David. “Empty Names.” *Noûs* 27, no. 4 (1993): 449–69. https://doi.org/10.2307/2215787.

Donnellan, Keith S. 1977. The contingent a priori and rigid designators. *Midwest Studies in Philosophy* 2 (1):12-27.

Evans, Gareth. “Can There Be Vague Objects?” *Analysis* 38, no. 4 (October 1, 1978): 208–208. https://doi.org/10.1093/analys/38.4.208.

Evans, Gareth. 1982. *The Varieties of Reference*. Edited by John McDowell. Oxford: Clarendon Press.

Everett, Anthony. “Against Fictional Realism.” *Source: The Journal of Philosophy*. Vol. 102, 2005. https://www.jstor.org/stable/3655676.

Evnine, Simon. 2016. Making objects and events: A hylomorphic theory of artifacts, actions, and organisms. Oxford University Press.

Fine, Kit. 1982. The problem of non-existents –Internalism. Topoi, 1, 97–140.

Friedell, David. 2021. “Creating Abstract Objects,” *Philosophy Compass*. Vol. 16.

Hawthorne, John, and David Manley. "Epistemic acquaintance." In *The Reference Book*. Oxford: Oxford University Press, 2012. Oxford Scholarship Online, 2012. doi: 10.1093/acprof:oso/9780199693672.003.0003.

Irmak, Nurbay. 2021. “The problem of creation and abstract artifacts”. *Synthese* 198 (10):9695-9708.

Jeshion, Robin. 2002. “Acquiantanceless De Re Belief”. In Joseph Keim Campbell, Michael O.’Rourke & David Shier (eds.), *Meaning and Truth: Investigations in Philosophical Semantics.* Seven Bridges Press. pp. 53-74.

Korman, Daniel. 2014. “The vagueness argument against abstract artifacts.” *Philosophical Studies*. 167(1), 57–71.

Korman, Daniel 2015. Objects: Nothing out of the ordinary. Oxford University Press.

Kripke, Saul A. 1980. *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.

Kripke, Saul. 2011. “Vacuous Names and Fictional Entities.” In *Philosophical Troubles*. Oxford University Press.

Kripke, Saul. 2013. Reference and existence: John Locke lectures. Oxford University Press.

Lewis, David. “Truth in Fiction.” *American Philosophical Quarterly*. Vol. 15, 1978.

Thomasson, Amie. 2009. “Fictional Entities.” In *A Companion to Metaphysics*, edited by Jaegwon Kim, Ernst Sosa and Gary Rosenkrantz, 10-18. Malden: Wiley-Blackwell.

Thomasson, Amie. 1993. “The Reference of Fictional Names.” *Kriterion* Nr. 6: 3-12.

Parsons, Terrence. 1980. Nonexistent objects. Yale University Press.

Récanati François. 2013. *Mental Files*. Oxford: Oxford University Press.

Russell, Bertrand. 1985. *The Philosophy of Logical Atomism*. Open Court Classics. LaSalle, Illinois: Open Court.

Salmon, Nathan. “Nonexistence.” *Noûs* 32, no. 3 (1998): 277–319. http://www.jstor.org/stable/2671935.

Searle, John R. “Proper Names.” *Mind* 67, no. 266 (1958): 166–73. http://www.jstor.org/stable/2251108.

Urtizberea, Agustin Arrieta (2005). ‘Neptune’ between ‘Hesperus’ and ‘Vulcan’: On descriptive names and non-existence. *Acta Analytica* 20 (3):48-58.

van Inwagen, Peter. 1990. Material beings. Cornell University Press.

Wittgenstein, Ludwig, G. E. M. Anscombe, P. M. S. Hacker, and Joachim Schulte. *Philosophische Untersuchungen =: Philosophical Investigations*. Rev. 4th ed. Chichester, West Sussex, U.K. ; Malden, MA: Wiley-Blackwell, 2009.