

Lingnan University

## Digital Commons @ Lingnan University

---

Lingnan Theses and Dissertations

Theses and Dissertations

---

1-21-2021

### **Fictional characters and their discontents : a prolegomena to any future metaphysics of fictional entities**

Shamik CHAKRAVARTY

Follow this and additional works at: <https://commons.ln.edu.hk/otd>



Part of the [Philosophy Commons](#)

---

#### **Recommended Citation**

Chakravarty, S. (2021). Fictional characters and their discontents: A prolegomena to any future metaphysics of fictional entities (Doctor's thesis, Lingnan University, Hong Kong). Retrieved from <https://commons.ln.edu.hk/otd/121/>

This Thesis is brought to you for free and open access by the Theses and Dissertations at Digital Commons @ Lingnan University. It has been accepted for inclusion in Lingnan Theses and Dissertations by an authorized administrator of Digital Commons @ Lingnan University.

## **Terms of Use**

The copyright of this thesis is owned by its author. Any reproduction, adaptation, distribution or dissemination of this thesis without express authorization is strictly prohibited.

All rights reserved.

FICTIONAL CHARACTERS AND THEIR DISCONTENTS:  
A PROLEGOMENA TO ANY FUTURE METAPHYSICS OF  
FICTIONAL ENTITIES

CHAKRAVARTY SHAMIK

PHD

LINGNAN UNIVERSITY

2020

FICTIONAL CHARACTERS AND THEIR DISCONTENTS:  
A PROLEGOMENA TO ANY FUTURE METAPHYSICS OF  
FICTIONAL ENTITIES

by  
CHAKRAVARTY Shamik

A thesis  
submitted in partial fulfillment  
of the requirements for the Degree of  
Doctor of Philosophy in Philosophy

Lingnan University

2020

## ABSTRACT

### Fictional Characters and Their Discontents: A Prolegomena to Any Future Metaphysics of Fictional Entities

by

CHAKRAVARTY Shamik

Doctor of Philosophy

In recent metaphysics, the questions of whether fictional entities exist, what their nature is, and how to explain truths of statements such as “Sherlock Holmes lives in 221B Baker Street” and “Holmes was created by Arthur Conan Doyle” have been subject to much debate. The main aim of my thesis is to wrestle with key proponents of the abstractionist view that fictional entities are abstract objects that exist (van Inwagen 1977, 2018, Thomasson 1999 and Salmon 1998) as well as Walton’s (1990) pretense view, which denies the existence of such entities. In the process, I propose modifications to these views to deal with problems they face and show how the modifications better account for the philosophical data.

Key abstractionists (van Inwagen 1977, Thomasson 1999) make a strict distinction between discourse within fiction, in which statements about literary characters cannot be literally true, and discourse about fiction, as it occurs in literary criticism, where statements about fictional characters can be literally true. Fictional objects are postulated to account for the truth of the latter. This runs into trouble because statements thought to be literally true are not literal. (Yagisawa 2001, Friend 2002) I provide a uniform analysis to account for the truth of statements involving fictional characters by appealing to a presupposition involving a metaphor in both contexts. The presupposition is that there is an  $x$  such that  $x$  is fictional;  $x$  is likened to a real person; and  $x$  is and ought to be treated/counted as a real person for all relevant intents and purposes.

More generally, I adopt Everett and Schroeder’s (2015) realist view that fictional characters are ideas constituted by mental representations. This, to me, better accounts for how fictional characters are created within the world’s causal nexus (unlike non-spatiotemporal entities in abstractionism), among other things. One key challenge they face is to explain how ideas can possess properties such as being a detective. I present a fine-grained version of their view, according to which the mental representations constituting fictional entities encode mind-dependent properties. Moreover, I explain how reference to such representations is possible, using Bencivenga’s (1983) Neo-Kantian view of reference and Karttunen’s (1976) view on discourse referents. Finally, I suggest that the identity of fictional characters is interest-relative. The constant, and sometimes radical, change of properties that fictional characters can undergo is taken to be a consequence of the fact that unified mental representations are bundles of simpler mental representations. As change occurs, simpler representations are replaced by others.

A key theme that runs through the thesis is that neither fictionality nor pretense is relevant to the semantics of fictional sentences—a claim bolstered by Matravers’ (2014) arguments.

Whether or not my account works, this claim, as well as the new philosophical data I bring up, are some of the challenges I pose to the heart of established views.

## DECLARATION

I declare that this is an original work based primarily on my own research, and I warrant that all citations of previous research, published or unpublished, have been duly acknowledged.

SIGNED

---

(CHAKRAVARTY Shamik)

Date: 14 May 2021





CERTIFICATE OF APPROVAL OF THESIS

FICTIONAL CHARACTERS AND THEIR DISCONTENTS:  
A PROLEGOMENA TO ANY FUTURE METAPHYSICS OF  
FICTIONAL ENTITIES

by  
CHAKRAVARTY Shamik

Doctor of Philosophy

Panel of Examiners :

 SIGNED	(Chairman)
(Prof. SEWELL Andrew John)	
 SIGNED	(External Member)
(Prof. KIM Seahwa)	
 SIGNED	(Internal Member)
(Prof. DE CLERCQ Rafael)	
 SIGNED	(Internal Member)
(Prof. ZHANG Jiji)	


Chief Supervisor :

Prof. DE CLERCQ Rafael

Co-supervisor :

Prof. SAUCHELLI Andrea

Approved for the Senate : /

 SIGNED
(Prof. MOK Ka Ho Joshua)
Chairman, Postgraduate Studies Committee

21 JAN 2021

Date



## CONTENTS

LIST OF FIGURES .....	vi
ACKNOWLEDGEMENTS.....	vii
1. INTRODUCTION.....	1
1. Meinongianism	
2. Fictional Possibilism	
3. Anti-Realism about Fictional Entities	
4. Fictional Abstractionism	
5. My Account	
5.1 Lakoff and Johnson’s Conceptual Theory of Metaphor	
6. Motivations for my Account	
6.1 Fictional Characters as Sets of Properties	
6.2 The Problem of Creation	
6.3 Problems of Identity	
6.3.1 Problems with Essential Properties of Fictional Characters	
6.3.2 Problems with A Necessary Condition for Identity	
6.3.3 Interest Relative Identity	
7. The Problem of Internal and External Perspectives	
8. Advantages over Meinongianism and Possibilism	
9. Conclusion	
2. VAN INWAGEN’S FICTIONAL CREATURES AND THEIR DISCONTENTS.....	30
1. Introduction	
2. Van Inwagen’s Account of Fictional Characters	
2.1 Van Inwagen’s Argument for the Existence of Fictional Characters	
2.2 Classification of Fictional Characters	
Van Inwagen’s Argument Against Paraphrase	
3. Yagisawa’s Objections to Van Inwagen	
4. Goodman’s Defense of Van Inwagen against Yagisawa	
4.1 Goodman’s First Argument	
4.2 Goodman’s Second Argument	
5. Jung and Pellet’s Objections to Van Inwagen	

5.1 Jung and Pellet's First Objection	
5.2 Jung and Pellet's Second Objection	
6. Solodkoff and Woodward's objection to Van Inwagen	
7. Hayaki's Objection to van Inwagen	
3. THOMASSON'S FICTIONAL ARTEFACTS AND THEIR DISCONTENTS. ....	56
1. A Brief Overview	
1.1 Fictional Characters as Dependent Abstracta	
1.2 Reference to Fictional Characters	
1.3 Identity Conditions for Fictional Characters	
1.4 The Problem of Negative Existentials	
2. Zemach's Objections	
2.1 Zemach's First Objection	
2.2 Zemach's Second Objection	
2.3 Zemach's Third Objection	
3. Friend's Objections	
4. Iacona and Voltolini's Objections	
4.1 Iacona and Voltolini's First Objection	
4.2 Iacona and Voltolini's Second Objection	
5. Reicher's Objection	
6. Howell's Objections	
6.1 The Problem with a Non-Spatiotemporal Fictional Character	
7. Everett's Objection	
7.1 Solodkoff's Modification of Thomasson	
7.2 My Objection to Solodkoff	
8. Van Inwagen's Objection	
4. SALMON'S SINGLE TIER THEORY AND ITS DISCONTENTS.....	100
1. Introduction	
1.1 Salmon's Account	
2. Sawyer's Objections	
3. Interlude: My Objection to Salmon	
4. Contessa's Objection	
5. Hayaki's Objection	
6. Predelli's Objection to Salmon and Modification	

6.1	Predelli's Solution	
6.2	My Objections to Predelli	
6.3	Why Are Properties Salient	
6.4	The Problem with Negative Existentials	
5.	WALTON'S PRETENSE THEORY AND ITS DISCONTENTS.....	136
1.	A Brief Overview	
2.	Friend's Objections to Walton	
2.1	The Problem of Intersubjective Identification	
2.1.1	Friend's Objection to Walton	
2.1.2	Friend's Account of Intersubjective Identification for Fictional Characters	
2.1.3	Garcia-Carpintero's Objections to Friend	
2.1.4	Garcia-Carpintero's Account	
2.2	My Objections to Friend and Garcia-Carpintero on Intersubjective Identification	
3.	Zemach's Objections to Walton	
4.	Richards's Objection to Walton	
5.	Pautz's Objections to Walton	
5.1	Excursus into my account	
6.	Matravers' Objections to Walton	
6.1	How Matravers' Account Fits in With Mine	
6.	THE METAPHYSICS OF FICTIONAL ENTITIES.....	191
1.	Fictional Characters as Ideas	
1.1	Master Argument for Fictional Characters	
1.2	Mind-Dependent Properties	
1.3	What are Mind-Dependent Properties?	
2.	Arguments for Fictional Content	
2.1	Explanation for van Inwagen's Mistake	
2.2	Explanation for Kripke not finding an "Internal Structure"	
2.3	Persistence Conditions for Fictional Characters	
2.3.1	Thomasson's Continued Dependence Condition	
3.	Interest Relative Identity	
3.1	Objections Against Bundles	
4.	Referring to Mental Representations	
4.1	Discourse Referents and Denotation	
4.2	Karttunen's Discourse Referents	

- 4.3 Negative Existentials and Valuation Features
- 4.4 Discourse Referents Clarified
- 4.5 Reference, Not Denotation
- 5. Idiolects and I-Language
  - 5.1 Refers (to) as an Intentional (sic.) Transitive Verb
- 6. Objections
  - 6.1 Objections on Aboutness and Reference
  - 6.2 A Fregean Objection
- 7. Everett's Problem of Indeterminacy
- 8. Conclusion

BIBLIOGRAPHY.....259

## List of Figures

Figure 1.	Reference Picture for sentences (8), (9), (10).....	231
-----------	---	-----

## **Acknowledgements**

No amount of gratitude is enough for my supervisors, Rafael De Clercq and Andrea Sauchelli as well as my wife Ipshita Mukherjee for this thesis. My supervisors not only provided valuable feedback, challenging discussions and probing questions, but also bore with my seemingly endless arguments. My wife Ipshita, who is from an entirely different Business major background, not only patiently sat through as a guinea pig when I tested her intuitions on thought experiments and gave feedback, she also took care of me and provided valuable emotional support through the thesis. Kyle York was the bouncing board for my arguments even as I barged into his room without notice. So I thank him. I am grateful to my friend Chris Atkinson who heard out my ideas and helped me. I thank Shampika Mukherjee, my mum-in-law, for her hospitality and warmth. I thank Jeffrey Ketland, Derek Baker, Dan Waxman, Dan Marshall, Simon Goldstein, Micheál Lacey, Ezechiel Thibaud, David Saldanha, Haktan Akcin and Andreas Matthias for giving me insights, feedback and providing valuable discussion. My mentors back in India, Bijoy Boruah, Arudra Burra, Kranti Saran and Scott Dixon did readings with me, read my work and rigorously forced me to clarify my ideas for which I am deeply grateful. Last but not the least, I thank Nilanjan Bhowmick, who patiently checked my arguments for rigour and heard me out when I needed him most.

## Chapter 1

### Introduction

In recent metaphysics, the issue of fictional entities has undergone extensive debate. Are there any fictional entities? This question is motivated by another question: What makes a statement about a fictional character true? Apart from the ontological question, there is also the broader metaphysical question: What is the nature of these entities? (Thomasson 1999)

For those who accept the existence of fictional entities (fictional realists), there are three broad solutions to the ontological and metaphysical questions: Meinongianism, Fictional Possibilism and Fictional Abstractionism. Hanley (2012) sets out the data to be explained for any theory of fictional entities, in which they have to account for the following type of statements (as given by Hanley 2012: 360-361 who draws the examples from the “Sherlock Holmes” page on *Wikipedia*):

(1) Critical statements

Example: Holmes is a rounded character<sup>1</sup>.

(2) Fictional statements

Example: Holmes can often be quite dispassionate and cold; however, when hot on the trail of a mystery, Holmes can display a remarkable passion, despite his usual languor.

(3) Conative attitude statements

Example: That’s why I admire Holmes. But alas, I can only hope to be like Watson, one who follow his exploits as an observer. Both Watson and I want to be like him, but we can only be his friend.

(4) “Carryover”<sup>2</sup>

Example: Sherlock Holmes’ abilities as both a good fighter and as an excellent logician have been a boon to other authors who have lifted his name, or details of his exploits, for their plots.

---

<sup>1</sup> My example, not Hanley’s.

<sup>2</sup> Carryover statements are different from mixed statements since one fictional character is imported from one work of fiction to another.

(5) “Mixed” statements

Example: In 2002, Holmes was inducted as an honorary fellow of the Royal Society of Chemistry.

(6) Creativity

Example: Holmes is the creation of Scottish-born author and physician, Sir Arthur Conan Doyle.

Apart from these that Hanley outlines, one also has to explain the truth of negative existentials like:

(7) Sherlock Holmes does not exist.

In what follows in the introduction to my thesis, I will outline the main views that try to account for these statements as well as put forward key objections to them in the literature. After doing so, I will adopt Anthony Everett and Timothy Schroeder’s (2015) view that fictional characters are ideas and propose a more fine grained version of the view, which solves a key problem it faces. The account that fictional characters are ideas is itself a modification of artefactual theories like Amie Thomasson’s (1999), which claim that fictional objects are created abstract artefacts like marriage, money, literary works and the like. I also provide a glimpse of my master argument for why fictional characters exist and elaborate on Lakoff and Johnson’s (1980, 2003) “Conceptual Theory of Metaphor”, since my account is dependent on that theory. After this, I lay out motivations for proposing my account in the light of some key problems faced by other theories in the literature on fictional entities.

## **1. Meinongianism**

The main thesis of Meinongianism is to say that there are non-existent objects. This accounts for the truth of statements like (7) by taking them at face value as well as fictional statements like (2). Since the fictional object possesses the properties that the fiction attributes of it, statements like (2) true.

Meinong’s (1904/1971) motivation stems from his “principle of intentionality”: all our mental phenomena have an object that they are directed towards. For instance, when we imagine, we always imagine something. This would mean that there must exist that "something" for the mental state of imagining to be directed at. If one thinks about the



golden mountain, then the golden mountain must exist--that is, there must *be* a golden mountain as an object of the thought. Thus, beyond being (*sein*), there is the additional realm of nonexistent objects (or beyond being and nonbeing called *Ausserseinde* objects) which has objects like the golden mountain (fictional), the round square (impossible objects) etc. Meinong makes a distinction between the way in which something is given descriptively (“their being such-and-so” (Kroon & Voltolini 2011) (*sosein*), which is different and independent from its being (*sein*). This means that even though unicorns do not have being (*sein*), they have properties like *being a horse with one horn* etc. This is due to the Characterization Postulate as formulated by Richard Routley: “According to the Characterization Postulate, objects, whether they exist or not, actually have the properties which are used to characterize them, for example, where *f* is a characterizing feature, the item which *fs* indeed *fs*.” (Routley 1980: 46 as quoted in Voltolini 2006: 7) But if we ask the question as to how these objects possess the properties that are used to characterize them, then there is a problem. If we say that language or the imagination (if they are intentional objects as in Brentano) has this ascriptive power, then we are left to wonder how it can have such a power. Therefore, Meinong views these entities, which are, in a sense, beyond being and non-being, as Platonic entities which already possess these properties. This explains (7), that there is no contradiction when we say that “Sherlock Holmes does not exist”, because existence then would be predicated in the sense of *Sein*, whereas, Sherlock Holmes is *Aussersein*.

One problem for Meinong’s view is whether fictional entities possess properties that are different from those given to them in a story. The question arises to account for the data above, where statements other than fictional statements are possible. (Voltolini 2006: 23-24) It can be said of Hamlet that he is a fictional character, created by Shakespeare, that he does not exist in space and time etc. Furthermore, Thomasson (2009) points out an issue based on Russell’s (1905) objection against Meinong. Since stories attribute existence to Holmes, Holmes exists. At the same time, since on Meinongian views, there *are* fictional entities that do not exist, Holmes does not exist. So there is a contradiction with Holmes existing and not existing. Parsons (1980) and Zalta (1983) have both come up with solutions to this problem. Parsons distinguishes between two *kinds* of properties, nuclear and extranuclear properties, the former being possessed by a fictional character in the story like *being human*, *being a detective* etc. and the latter properties the fictional character has based on what the world actually is, like *nonexistence*, *being a fictional character* etc. So Hamlet does not really exist. Theorists like Zalta will deny the kinds distinction and say that there really are two

modes in which properties are predicable. Fictional characters thus *encode* properties like existence but *exemplify* nonexistence.

Sainsbury (2009) points out lacunae with Meinongian accounts as follows. Meinongians cannot account for our intuitions about the creative process, that fictional characters are created and thus brought into existence by an author. Moreover, after they are introduced, properties are added to them as they are developed. A Meinongian might reply that they are not brought into existence but into nonexistence. But if they accept the inference from “x does not exist” to “x is nonexistent”, they would countenance inferences relative to time like “Holmes did not exist in 1780” to “Holmes was nonexistent in 1780” (*ibid.*, 58). If this is possible, then Conan Doyle could not have brought Holmes into nonexistence before he was born. Moreover, even if they say that they were always nonexistent but their *Sosein* (which is minimal) gets enriched by the author, there will be a problem, what Sainsbury (2009) calls the “selection problem”: how does the author at the time of creation pick out the right non-existent object to add properties to?

Two options are:

- (i) These are individuated by all the nuclear properties they ever will have (those belong to the nature of the object rather than their metaphysical status).
- (ii) All the nuclear properties ascribed in Conan Doyle novels.

But the problem is that there is a conflict between fixed nuclear properties and the fact that more properties are added to a character as a story progresses. The character also does new things with time. (*ibid.*) The problem is that one cannot individuate any evolving or changing character with the help of fixed nuclear properties.

## **2. Fictional Possibilism**

According to Fictional Possibilism or nonactualism, which is mainly espoused by David Lewis (1978), fictional objects exist in a possible world in which a story is related as “known fact”. (*ibid.*, 40) The events, characters and objects that are spoken of in that world are those that are fictional and not concrete in the actual world but are actual in the possible world. This is in contrast to actualism, which says that the only objects that exist are those that exist in the actual world. No possible objects exist.

Hanley (2012: 363) points out that the advantage of possibilism over other realist theories of fiction is that it would imply that fictional objects are complete. That is, that an object is complete if, for every property F, it either has F or its negation -F. Usually realist theories face the problem of incompleteness, because objects only possess the properties that are

attributed of them in stories. The stories might not mention if Sherlock Holmes has a mole in his right shoulder or not. Given that the object lacks F and not-F, it is incomplete with regard to F and -F. But in possibilism, Sherlock Holmes will be a full blooded complete candidate in another possible world, its properties not dependent on the mercy of a story. Another advantage (*ibid.*) is that that mixed statements can be true, when we compare two characters from different fictions because they are comparisons between worlds.

One of the main problems in nonactualism is on the identity of *non-actual possibilia* which are fictional objects. The problem of ontological indeterminacy plagues this postulation; if in various possible worlds there are objects that match the description in the story, but also have other properties like being born from different parents, which one of these candidates is the character from the story? Kripke (1980) illustrates this by saying that there might have been many different possible people who do the same exploits as Holmes, but it is indeterminable which of them is actually Holmes. They might as well have been Darwin or Jack the Ripper who did the things which are ascribed of Holmes. It would then be arbitrary to pick any one of them and say that that is Holmes (Kripke 1972). This is similar to Sainsbury's Selection Problem. Lewis tries to deal with this by his counterpart theory by saying those worlds have the character in which the story is told as fact and other candidates with the same properties in which the story is not told as fact are merely counterparts of the character (Sainsbury 2009).

Sainsbury says that possibilism also fails to account for the intuition that fictional objects are created. Another problem that plagues possibilism is that of impossible objects. A story can talk about an impossible object like a round square, but given that in possibilism a fictional object is a possible object, it cannot account for the truth of statements in fiction about impossible objects.

### **3. Anti-Realism about Fictional Entities**

One point in favour of the realist is that they take our ordinary statements (at least extra-fictional statements) about fictional entities for granted, and provide a metaphysics according to which they are straightforwardly true. Anti-realists about fictional entities, however, deny their existence through paraphrase, further analysis of names in fiction as definite descriptions (Russell, 1905) or by construing fictional statements as part of games of pretense. The most prominent anti-realists about fictional entities are Bertrand Russell (1905), Kendall Walton (1985, 1990), Anthony Everett (2005, 2013) and Mark Sainsbury (2009).

In the paraphrase analysis, for instance, the fictional statement “Holmes lives at 10 Baker Street” will typically be prefixed by a fictional operator “*According to the fiction*, Holmes lives at 10 Baker Street”, which puts the fictional statement in an intensional context, not requiring an extension. Kendall Walton’s (1985, 1990) pretense theory claims that fictional statements are true only in a game of pretense and otherwise false, and reference to characters is pretend reference. However, his is not a paraphrase, like “In the fiction, Sherlock Holmes lives on 221B Baker Street”.

Anti-realists do seem to account for fictional statements (Hanley’s criterion 2)). But problems with anti-realist theories arise with explaining how statements like “Holmes was created by Arthur Conan Doyle”, that is external claims, are true (Thomasson, 2009). Thus they do not satisfactorily account for Hanley’s 6<sup>th</sup> criterion, creativity.

#### **4. Fictional Abstractionism**

Abstractionist theories claim that fictional entities are abstract objects that exist. Within abstractionist theories, there are two important subsets, one that claims that they are eternal necessary entities (Wolterstorff 1980, van Inwagen<sup>3</sup> 1977, 2018) and those that claim that they are contingent created abstract artefacts. I shall call the latter artefactual theories (Searle 1975, Salmon 1998, Thomasson 1999, Kripke 2013), which treat fictional objects as artefacts just like marriage, money and other socially created entities. These avoid the selection problem and argue that their theory explains the point that fictional entities are created in time, unlike Meinongianism which postulates eternal entities. Artefactualists like Thomasson argue that if rhyme schemes, novels etc. exist, then there is no reason to discount the existence of fictional characters.

Within realism, there are two forms based on individuation criteria (the distinction also extends to Meinongian views): internal realism and external realism (Friend, 2007). Internal realism would say that properties of a character are those individuated from a perspective internal to fiction, that is properties like *being a detective* and *being a Prince of Denmark*. For instance, Wolterstorff (1980) sees fictional characters as person kinds, Currie (1990) as roles and Lamarque (2010) as character types. According to these views, characters are eternal sets of properties.

---

<sup>3</sup> It is open in van Inwagen (1977) whether he is an artefactualist or not but he seems to be more of an artefactualist.

External realism is associated with artefactualist views. In external realism, fictional objects are to be individuated by means of external properties, for instance, those given by the language of literary criticism. Van Inwagen argues that fictional characters are “theoretical entities of literary criticism” (van Inwagen 1977: 302) because, although statements typically from fiction, like “Hamlet is the Prince of Denmark” or “Hamlet exists” are *prima facie* false, statements from literary criticism such as “Hamlet is created by Shakespeare” are true. His argument is that from (*ibid.*, 302):

(8) There are characters in some 19<sup>th</sup> Century novels who are presented with a greater wealth of detail than is any character in any eighteenth century novel.

one can infer that there are characters in some 19<sup>th</sup> century novels. If this is true, one can quantify over characters and say that therefore, there are (fictional) characters. Now in van Inwagen’s theory, for instance, Mrs. Gamp from Dickens’ *Martin Chuzzlewit* is individuated by external properties like “*being a character in a novel, being a theoretical entity of literary criticism, having been created by Dickens, being a main satiric villainess*” and “*being a fair representation of a hired attendant on the poor in 1873.*” (van Inwagen 1977: 305)

Kripke (2013) believes that ordinary language commits us to actual fictional objects. Unlike Meinongian objects, they are created artefacts like nations. Thomasson (1999), inspired by this account, believes that fictional characters as abstract objects make best sense of our literary practices.

Kripke's argument is as follows. “Moloch” might not have been the pagan god he is construed to be, but might mean Yahweh or is a sort of sacrifice. From this, it can be inferred that there was no pagan god as Moloch, and that it was an error to assume that such a god existed. When Kripke mentioned that there was no such pagan god but “Moloch” referred to a sacrifice, Professor Frankfurt responded by saying, “Of course there was not such a god. You don’t believe in pagan deities, do you.” (*ibid.*, 71)

Frankfurt’s joke reveals an ambiguity. In case of Frankfurt, this is a question regarding the world: “Was there a pagan god as Moloch?” The answer to which is, no. But, “Is there such a legendary character?” (*ibid.*, 71) The answer would be *yes* or *no* subject to empirical historical investigation. If the answer is *yes*, then the fictional character exists.

Also, Kripke asks, “was there a fictional or legendary character who married his grandmother?” (*ibid.*, 71) If indeed there was such a legendary character, then the statement “There was a fictional/legendary character who married their grandmother “ is true because it is based on a fiction or legend that was written/spoken. In case the fictional work exists, then so does the fictional character.

The strongest advantage of artefactualism over other theories is to account for the intuition that they are created, that is Hanley’s 6<sup>th</sup> datum. It also seems to account for critical statements, mixed and carryover statements. I shall lay out problems with abstractionism below.

## **5. My Account**

My account is a fine grained and modified account of Everett and Schroeder’s (2015) realist theory, that fictional characters exist and are ideas. The reason for their view is that ideas have the same properties attributed to them that fictional realists take to be attributed to fictional characters. For instance, ideas are created, cease to be, they can be original or dull and inspired from real life events or people or even other ideas. Like fictional characters being unidimensional or multidimensional, ideas are also detailed or simple. Moreover, they have an advantage over abstractionism, that ideas, being constituted by mental representations, are part of the causal spatiotemporal nexus of the world whereas abstract objects aren’t. So it would be more plausible for a realist view to adopt the view that fictional characters are ideas rather than abstract objects.

My modification is made in the light of a key objection to their view that they acknowledge, namely that there is a problem with how ideas can possess properties like *being a detective* and *being round*, since only people can literally be detectives and only objects can literally be round. I elaborate that if we do come up with a fine grained theory of mental representations and mind-dependent properties, we can resolve that issue. That in turn is a modification of artefactualism and tries to overcome its shortcomings.

The key claims are as follows:

a) The truth of all statements involving fictional entities involves presuppositions with a metaphorical relation “is likened to”. For instance, statements involving fictional characters like Holmes, who are specified as persons in the fiction will presuppose:

(MET<sub>FIC</sub>)\*\* There is an x such that x is fictional and x is likened to a real object and x is and ought to be treated/counted as a real object for all relevant intents and purposes.

(MET<sub>FIC</sub>)\*\* also entails that fictional characters exist. My master argument (which I elaborate on in Chapter 2 and 6) for the existence of fictional characters involves (MET<sub>FIC</sub>)\*\*, modifies van Inwagen's argument and attempts to overcome problems faced by his account. It must be noted that "a real object" in "x is likened to a real object" is not a particular object, but a concept of an object. For instance, if Jesus said, "The Kingdom of God is likened unto a lamp", "a lamp" would not be any particular lamp, but the concept of a lamp. More specifically, statements involving Holmes will presuppose:

(MET<sub>FIC</sub>)<sup>SH</sup> Sherlock Holmes is fictional and is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

I also propose that the differences between discourse within fiction and metafictional discourse, like the language of literary criticism where statements like "Sherlock Holmes is a fictional character" are due to the primacy of the intentional stance when it comes to discourse within fiction and the design stance when it comes to metafictional discourse. According to Dennett (1983), the design stance is a heuristic strategy in which it is assumed that an object is designed for a certain purpose and it will act accordingly. For instance, the behaviour of a computer can be predicted by analysing its source code. It is because we focus on what a fictional character's design is as a fictional character, for instance, that they are designed as rounded or flat characters, we tend to see them qua their design in literary criticism. The intentional stance, on the other hand, is a strategy to predict and explain intentional states of systems (living as well as non-living) that include beliefs and desires. The system is treated as a rational agent and the behaviour is predicted accordingly. For a fictional character, I focus on the flesh and bloodedness, for instance, I imagine situations in which they are embodied since I need to predict their beliefs and desires in this embodied state. I elaborate on this in Chapter 2 and Chapter 3.

b) The claim is from Everett and Schroeder that fictional objects are ideas that are constituted by mental representations. This best accounts for the truth of the statements above. My divergence from their view is with respect to the content of mental representations when it comes to fictional characters. The following are the tenets of my modification:

Tenet 1. The content of mental representations with respect to fictional characters is given by bundles of properties of mental representations. The properties are mind-dependent properties such as *being a thought of being red*, *being an appearance of being red*, *being a conception of thought of being a detective*, *being an experience of something's looking square*.

Tenet 2. A fictional object can be multiply realized as constituted by different mental representations, relative to time *t*, functional mind *m* and contextual convention *c*, a fictional character is constituted by particular token of a bundle of mental representations that form a unified representation.

Tenet 3. A bundle of representations is an instance of representations linked by neural binding. The bundle itself forms a unified representation. Neural bindings are brain processes and insofar as mental states are types of events that are multiply realizable in a functional mind, they are types of events. Each instance of a neural binding is a token event and so is a bundle of mental representations a token event in the brain of someone thinking of the character.

Thagard (2019) defines mental representations as

...not things or properties but rather processes consisting of patterns of firing that result from the synaptic connections among neurons. These learned or innate connections enable neural groups to produce useful patterns of firing in response to inputs from the environment or other neural groups. (Thagard 2019: 24)

Thagard says that simpler features (I shall use “features” and “properties” interchangeably) like sensory (like *what cats look and sound like*), motor (like *how it feels to pick up a cat*), emotional features (like *how much you like cats*) and verbal features (like cats being a kind of mammal) are bound through neural binding to form more complex mental representations which are unified. These features are information encoded by mental representations (each mental representation encodes a feature) and are mind-dependent. In Chapter 6, I argue that when it comes to fictional characters, these are the best candidates for the content of mental representations, especially since Everett and Schroeder's view faces the objection as to how ideas can literally have the property *being red*, *being a detective etc*. Since these are properties of mental representations, I prefer to express these features as properties like *being a thought of being red*, *being an appearance of being red*, *being a conception of thought of being a detective*, *being an experience of something's looking square*. In Chapter 6, I shall explain more about how these mental representations come to be about properties like *being red*, *being square*, *being flesh and blood etc*. It should suffice here to note that I consider these as tropes, like other views that see mental properties as tropes. (Maurin 2018) I expand on this in Chapter 6. I further explain how because a mental representation in these cases is “likened to” a concrete object or person or creature, we think of the mental representation as a (fictional) object or person or creature.



Tenet 2 explains how an author first thinks of a fictional character and brings together mental representations that encode properties (it is just that the information encoded in neural patterns are these properties, which is why I say that mental representations encode them) in doing so. At that time  $t_1$  when he first thinks of it, the fictional character is constituted by that bundle of mental representations, which further forms a unified representation. As he keeps adding properties that are encoded by the representations, while writing the novel, from  $t_1$  to  $t_n$  when he finishes writing the novel, at each time, the content of the mental representation is given by different bundle of properties from  $t_1$  to  $t_n$ , he keeps experimenting with different combinations of properties. The bundle can even be completely different from the one at the beginning, since the author can decide to overhaul all properties. Again, this shall be elaborated upon in sections 5.2, 6.1 and 6.2.

Before I move on to the motivations for my theory, it will be in place to outline the theory of metaphor that I adopt.

### **5.1 Lakoff and Johnson's Conceptual Theory of Metaphor**

Suppose there are fictional animals or other kinds of fictional entities specified in a work of fiction. That will involve metaphors in which the fictional entity will be "likened to" a relevant concrete kind. For instance, for Bambi, the fictional deer will be likened to a real deer (in a loose manner of speaking. I'd prefer to say a mental representation rather than Bambi the fictional deer, since even saying Bambi the fictional deer will presuppose (MET<sub>FIC</sub>)\*\*):

(MET<sub>FIC</sub>)<sup>BA</sup> Bambi is fictional and is likened to a real deer and is and ought to be treated/counted as such, for all relevant intents and purposes.

If the kind is indeterminate or unspecified in the fiction, or is exotic, it will involve some or other comparison to a concrete object derived from perceptual systems. The point is that our default cognitive systems will use more familiar and known concepts to understand exotic fictional kinds by means of metaphor. This should not be surprising in the least, since we do in fact understand concepts like time which are abstract by means of more familiar ones that are derived from perceptual experience in this manner. For instance TIME IS A MOVING OBJECT is used as a metaphor in statements like "I can't face the future", "The face of things to come . . ." Lakoff and Johnson (2003: 43) in which "time receives a front-back orientation facing in the direction of motion, just as any moving object would" even as the future has its face towards us as we move towards it. This would also explain why the vast majority of fictional entities are indeed modelled on concrete objects. TIME IS A MOVING

OBJECT is also what Lakoff and Johnson call an ontological metaphor, in which entities without clear boundaries like street corners and abstractions like time and love as well as the mind are conceived as objects with boundaries and surfaces so that we can refer to, quantify, talk and reason about them and this aids in understanding them. As Lakoff and Johnson say “Ontological metaphors are among the most basic devices we have for comprehending our experience.” (*ibid.*, 219)

The theory of metaphor I thus appeal to is George Lakoff and Mark Johnson’s Conceptual Theory of Metaphor. The theory’s main claim is that metaphors primarily have to do with systematic patterns of conceptual association and only secondarily are linguistic, gestural or visual metaphors (Lakoff 2014). Furthermore, as Lakoff and Johnson say “The essence of metaphor is understanding and experiencing one kind of thing in terms of another.” (Lakoff & Johnson 1980: 5 as quoted by Grady 2010: 188). For instance, the LOVE IS A JOURNEY metaphor maps (in the mathematical sense) concepts from a particular domain of our experience, that is, love, to another domain of journeys. For instance, difficulties are seen as obstacles and goals are destinations:

Our relationship has hit a dead-end street... Look how far we’ve come. We can’t turn back now. It’s been a long, bumpy road. We’re at a crossroads. We may have to go our separate ways. The relationship isn’t going anywhere. We’re spinning our wheels. Our relationship is off the track. This marriage is on the rocks. We may have to bail out of this relationship. (Lakoff and Johnson 1993: 206)

Metaphor in cases like these are not primarily a matter of language because each expression like “dead-end street”, “can’t turn back now” “on the rocks” cannot be seen in isolation but in the light of the single conceptual metaphor governing them with systematic mappings across the domains. Furthermore, there are patterns of inferences licensed by the metaphor. For instance, based on LOVE IS A JOURNEY the following courses of action can be taken (*ibid.*, 208):

They can try to get it moving again, either by fixing it or getting it past the DIFFICULTY.  
They can remain in the nonfunctional RELATIONSHIP, and give up on ACHIEVING THEIR LIFE GOALS.

They can abandon the RELATIONSHIP

But more importantly, for Lakoff and Johnson, we systematically understand more abstract concepts in terms of metaphors derived from our bodily experience. They say “Metaphor allows us to understand a relatively abstract or inherently unstructured subject matter in terms of a more concrete, or at least more highly structured subject matter.” (*ibid.*, 244–245)

Lakoff and Johnson give a wealth of examples. One of them is the TIME IS A MOVING OBJECT metaphor we saw above. Another is the MORE IS UP metaphor in statements like “Inflation is rising” and “the crime rates are going up” (Lakoff and Johnson 2003: 23), which derive from our experiences where we relate an “increase in quantity” with an “increase in height”. (*ibid.*, 254)

It must be observed that conceptual metaphors are not “dead metaphors”. Live metaphors are those that are cognitively real and their metaphoric force is active in speakers’ minds. Dead metaphors are those that were once active, but people are no longer aware of its metaphoric status. For instance, “coming to a head” is no longer understood as involving the “buildup of a pus in a pimple” (Pinker 2007: 238). One reason why conceptual metaphors are not dead is the way we saw two domains being related systematically in the LOVE IS A JOURNEY metaphor. The systematic manner in which they are related and the recurrence of patterns of these types of metaphors is more suggestive of there being an understanding of the relation between the source domain (Journey) and the target domain (love) than a result of mere accidental use or relating words one by one based on analogy. (Grady 2010: 195) It is also telling that not only can we immediately make inferences based on the LOVE IS A JOURNEY metaphor but also we can easily understand new sentences like “we’re driving on the freeway of love” because the existing mapping in the live metaphor is merely extended to this. (Lakoff and Johnson 1999: 57)

It is also pertinent to note here, that unlike traditional theories of metaphor, the Conceptual Theory of Metaphor rejects similarity as a basis for conceptual metaphor. One argument Lakoff and Johnson (1999) put forward is an instance of a metaphor where there no similarity. The KNOWING IS SEEING metaphor in the statements “I see what you mean”, “That’s a murky argument”, “Let’s shed some light on the subject” (*ibid.*, 129) consists of there being knowledge in the source domain, which is seeing, as well as the target domain of knowing. However, since literal seeing is not possible in the case of seeing what one means, the similarity between seeing what one means and knowing what one means is not literal. A second argument by them is that the bare concepts of love and journeys are not similar and love begins to be seen as a journey only due to the metaphor. A third argument is the argument from asymmetry. The systematic pattern of source-target domain mapping that we see in conceptual metaphor is asymmetric. On the other hand, similarity is a symmetric relation and there would be no such thing as source or target domain in conceptual metaphor if the relation was symmetric. The argument by Lakoff and Johnson is that if conceptual metaphor was built on similarity, then both source and target could be expressed equally in

terms of each other. However, the asymmetry becomes evident when we see that in LOVE IS A JOURNEY, while words related to journeys are used to describe facets of love but those related to love, like “relationships” are not used to refer to cars, for instance, in new cases, conventional cases as well as what we saw with inferences above. We also saw this in the KNOWING IS SEEING case. We also do not characterise objects using time in the ways we use concepts from the domain of objects to characterise time.

The theory of metaphor I am using will become particularly pertinent in warding off possible objections to my arguments against taking phrases like “fictional object” or “pretend object” literally<sup>4</sup> (see especially Chapter 2 and Chapter 3). It is likely that a metaphor in which the source is derived from experience and mapped on to the target domain is at work here since even when we think of fictional characters like Sherlock Holmes, we cannot avoid conceiving of them as “containers with an inside and an outside” (Lakoff and Johnson 2003: 58) and with boundaries, the way we experience our bodies as containers (see quote below). If this is so, then given our cognitive apparatus, in case it is specified in a work of fiction that a fictional character is a vague object, we still can’t help but conceive of it in terms of an entity with a boundary with an inside and an outside. In my account, such a fictional character will be constituted by a bundle of mental representations, and one of the representations encode a property like *being a thought of being vague*. The bundle of properties encoded by the representations are in turn likened to a person. Lakoff and Johnson say the following on the origins of these basic concepts by which we make sense of more abstract and less structured concepts:

For example, the concepts OBJECT, SUBSTANCE, and CONTAINER emerge directly. We experience ourselves as entities, separate from the rest of the world—as containers with an inside and an outside. We also experience things external to us as entities—often also as containers with insides and outsides. We experience ourselves as being made up of substances—e.g., flesh and bone—and external objects as being made up of various kinds of substances—wood, stone, metal, etc. We experience many things, through sight and touch, as having distinct boundaries, and, when things have no distinct boundaries, we often project boundaries upon them—conceptualizing them as entities and often as containers (for example, forests, clearings, clouds, etc.). (ibid., 2003: 58)

## 6. Motivations for my Account

---

<sup>4</sup> Of course, even though I say they cannot be taken literally, it doesn’t follow that fictional characters don’t exist, in just the way that we saw in the metaphors above that the existence of conceptual metaphor does not entail that time or love does not exist.

The motivation for my account comes from having to deal with previous theories of fictional characters and the data that is thrown up, especially after objections to them. In what follows, I shall put forward the motivations as to why I adopt my account. While part of the motivations are in response to issues faced by Meinongianism and Possibilism, the key motivations lie in attempting to deal with issues of the identity and change of fictional characters and key objections to artefactualism. In laying out my motivations, I attempt to show how modifying artefactualism about fictional characters leads me to the view that the most likely explanation for the issues that artefactualism faces is that fictional characters are in fact ideas (Everett and Schroeder's view) that are constituted by bundles of mental representations that encode properties. It is in the light of this that I put forward my fine grained account of mental representations and their content when it comes to fictional characters.

### **6.1 Fictional Characters as Sets of Properties**

The most direct motivation is to deal with the theory that fictional characters are sets of properties. (Carter 1980) However, I reject this conception, and explain that the reason why the objections are legitimate and the intuitions behind the objections can be explained is because fictional characters are actually constituted by bundles of mental representations, where each representation encodes a mind-dependent property. A bundle of mental representations is an instance of mental representations, linked by neural binding. Neural bindings, which I explain in section 6.1, are brain processes and insofar as mental states are types of events that are multiply realizable in a functional mind, they are types of events. Each instance of a neural binding is a token event and so is the occurrence of a unified mental representation, which in turn is a bundle of representations, a token event in the mind.

The first problem with the view that fictional characters are sets of properties is that the extensions of sets are fixed but the properties of fictional characters keep changing. (van Cleve 1995) Across novels, fictional characters can change properties, get even contradictory ones and yet be the same fictional character. This is in line with one of the objections that we saw levelled against Meinongianism by Sainsbury.

Another argument against fictional entities as sets of properties is the same argument against Currie's (1990) theory of names within fiction. For Currie, each sentence within a story does not express an independent proposition and the content of a story is expressed by a single Ramsey sentence. The scope of the existential quantifiers of the Ramsey sentence is the entire story (*ibid.*, 155) and its values are fictional names as well as the fictional author.

Thus, fictional names are disguised descriptions and the meaning of each name cannot be seen in isolation but is given by the set of descriptions given in the story. Lamarque (1993) argues against this that an implausible result of this view is that people reading a story do not understand sentences they read until they reach the end of the work. The same goes for considering fictional entities as sets of properties. It is even more implausible to say they are sets of properties given in a story, since we might not know or remember every property described in the story (or every predicated of a character) and yet know or remember the meaning of a fictional name given in the fiction before finishing it.

Thus, the datum from the first argument is that fictional characters keep changing properties and even get contradictory ones. The datum from the second argument is that we do understand sentences as we encounter them. The proposed explanation by me attempts to give a naturalistic and parsimonious explanation for this (without invoking non-spatiotemporal abstract objects). It is because a fictional character is constituted by a bundle of mental representations at time *t* (as a unified representation), which encode a bundle of properties, at another time, we can think of a different bundle of mental representations that constitutes the character. This allows for a change of properties as we keep thinking of different properties of the character. Furthermore, a representation that encodes *being a thought of being round* and another one that encodes *being a thought of being square* can be bundled together in a unified representation. So we can account for a character with contradictory properties. As we understand each sentence, we think of a particular bundle of representations that keeps changing as predicates are added through a work and/or as we keep thinking of it with different properties. This accounts for how we understand each sentence as we encounter it.

## **6.2 The Problem of Creation**

Robert Howell (2002) observes, in the case of Thomasson's artefactualist account, that she has not provided an account of how imagining a concrete person can lead to the creation of an abstract object, but it is also puzzling how an abstract object can be created. In a similar vein, Stacie Friend (2007) points out that artefactualism cannot account for how fictional entities are created. It seems that there is a point at which they do not exist, and when created, they come into being. But one does wonder about the intermediate process. It is almost as if they come about without any "raw" materials" and *ex nihilo*.

However, my account provides an account of the creation of fictional characters. In a creative writing class, there are lessons in which participants first describe and think of possible names (which is part of the descriptions). Once they finish with this process,

properties and a name have been associated with each other, a whole character emerges. This process is more evident in the case of a game like *Dungeons and Dragons*. In the game, if you need to generate antagonists for people to fight with, there are charts with which one can construct the villains. The rolling of the dice generates them based on various combinations. For instance, one can find basic personalities like *being easy going, sarcastic* etc and combine them with other attributes<sup>5</sup>. Once they are combined, the combination of attributes then emerges as a creature. The same process is seen in the video game *No Man's Sky*, where characters are generated as a combination of attributes. What is common to all these cases is the combination of properties after which a character emerges. Thomasson might be able to incorporate this into her account and thus refute Friend's objection. However, the problem still would arise as to why a character should emerge out of the combination of properties, that is, at what point is a combination of properties a combination and why does it suddenly turn into a character as if by magic the way Pinocchio turns from a wooden boy to a real one? Moreover, we do not think of characters as mere bundles of properties, so why do our thoughts become thoughts about a character rather than a bundle despite the fact that the properties and a name have merely been associated with each other? My account has a twofold explanation for why this combination of properties occurs. First, there are mental representations, each of which encodes a mind-dependent property like *being a thought of being easy going, being a thought of being sarcastic* etc. What accounts for combinations is that mental representations encoding these mind-dependent properties are bundled through neural binding to form a unified representation. This allows for the combination of various properties to create fictional characters and fictional characters are created when the bundle of mental representations is tokened in an author's brain as a unified representation. The mental representation is likened to a real person, also by means of neural binding circuits in the brain (Lakoff 2014:5), which associate properties in the brain to form conceptual metaphors.

Furthermore, my account has the advantage of not facing the *ex nihilo* objection by Friend. Fictional objects are very much created by combining mind-dependent properties by combining mental representations, properties we have abstracted from our experiences (Barsalou 1999), and "when [we] reflect on this content...they are bringing to mind kinds and qualities that are likely to be familiar through exemplifications in the real world" (Lamarque 1996: 36). Our mental representations also combine amodal properties of ideas that are innate. Of course, since these properties already exist, it is to be clarified what is created. Moreover, "bringing to mind" is different from bringing into existence. So

---

<sup>5</sup> I owe this example to Derek Baker.

properties are not just brought to mind, but what is in fact created is a unified mental representation, that does come into existence (a new mental state), that encodes a bundle of properties. Furthermore, the unified mental representation is neither a representation of something in the external world nor is an additional abstract object required to explain what it is a representation of. As Chomsky (NC) says in his interview with Peter Ludlow (PL):

PL: Right, but now one might ask what on earth is a representation if it is not a representation of something.

NC: That's a mistake that comes from a philosophical tradition. The way the term 'represent' is used in the philosophical tradition, it's a relation between an internal object and an external object. It's not the way it's used either in ordinary speech or in the sciences. So, when a perceptual psychologist, say, talks about an internal representation of a cube or something, there doesn't have to be any cube there. They're talking about something that's going on in the head. In fact, what they may be studying, and usually are studying, is the relation between things like tachistoscopic presentations and internal events. There's no cube. But, nevertheless, they talk about it as an internal representation. The concept, internal representation . . . The concept, internal representation, is used in the sciences—and I think that's true in ordinary speech too—in ways which don't involve a relation between an internal thing and an external thing. That technical concept derives from a particular interpretation of the theory of ideas, which said, well, ideas represent something out there. (Ludlow 2011: 182)

In the light of what Chomsky says, it must be clarified at the outset that I find the term “mental representation” very confused, since it automatically makes us think of representations like pictures or maps, as necessarily involving a relation between the representation and another object that it represents. I merely think that the brain processes happen to be called “representations”, which in a sense they are, since they have that function. But it's better to think of them as simulations (the neuroscientist Benjamin Bergen (2012) says that we simulate reality when we interpret fictions and the same sensorimotor systems activated when we perceive are activated). It just seems odd to look for one to one correspondences for simulations. Properties of simulations resemble or correspond to what they are simulations of in the external world in some respects, but they go beyond and explore various possibilities. Our ability to simulate reality in our imaginations also explains why we create objects with new properties (because we thought of them beforehand) and also why we think up new properties that objects can have (like artworks or statues) and yet do not create anything. It helps to explain this in terms of intrinsic properties of the simulations too. In Chapter 6, I explain that what explains “x is about y”, where x is a unified complex mental representation and y is a fictional character is explained in terms of more basic brain processes which include the “likened to relation” which explains why it is about a flesh and blood person, for instance. In my interpretation of Chomsky, he seems to



be trying to say that philosophers tend to take the “x is about y” relation as primitively about x and a mind independent entity<sup>6</sup>.

If we just consider simple representations which encode properties that go on to make complex unified representations, I claim that they not only represent and are about things and properties in the world due to a causal link with them but also themselves. For instance, our thought about red is partly individuated by its causal link to the external world (which we have in our immediate egocentric space or in the past). However, our minds also hold colour constant to variances in reflectance and wavelengths. Similarly, even with shape, if we think about the cube Chomsky talks about above, that is a result of computations our mind makes, which partly explains why our minds take very little information from the retina and constructs the world in our minds as radically different from what it is, with clear shapes and boundaries. Thus, there appear *sui generis* properties which are also about themselves, especially since we can independently introspect about the shapes, sizes, qualitative feels, colours etc, that our minds present to us. Thus, the representation can also be about itself, since it has the dispositional property that one can have higher order thoughts about the representation. It must be noted here that *being a thought of being red* is itself property encoded by a first-order thought and children can have these first order thoughts too, even though higher order thinking appears later. I explain and clarify this further in Chapter 6. But more importantly, since the simpler representations can represent themselves, the unified representation that emerges out of the binding process through the combination of properties as well as what I said about the “likened to” relation can also represent itself without there necessarily being an external object to which the unified representation is related, like with fictional characters. One factor that our representation of fictional characters can be explained in terms of is this feature, where we represent our representations to ourselves by virtue of their intrinsic properties (See Chapter 4 and 6).

---

<sup>6</sup> I think that Susan Haack echoes Chomsky’s sentiment in a way, and it seems to me that the reason why philosophers want to relate representations to a mind independent particular is because they mistakenly take “real” to mean “mind independent”. As Haack (2020) says:  
Disputes between ‘realists’ and ‘anti-realists,’ all the rage in Britain in the 1980s, seemed to force us to choose: either an understanding of ‘real’ as ‘mind-independent,’ and an acknowledgment of something like the ‘fixed totality of mind-independent objects’ that Putnam’s metaphysical realist imagined, or else a conception of the world as something we somehow bring into being ourselves. Transcending this dubious dichotomy, my Innocent Realism begins with the thought that there are many things (laws, kinds, our mental states and processes, etc.) which, though certainly real, aren’t existent particulars; and that what ‘real’ means is neither ‘independent of us,’ nor ‘independent of our minds,’ but something more like ‘independent of what you or I or anyone believes about it.’ (Haack 2020: 2-3)

It must also be noted that the “creation” of this unified mental representation despite the prior existence of separate properties is explained by nothing more than what the binding problem in cognitive neuroscience seeks to explain, namely

...in which one sort of visual feature, such as an object’s shape, must be correctly associated with another feature, such as its location, to provide a unified representation of that object. Such explicit association, or “binding,” becomes especially important when more than one visual object is present, in order to avoid incorrect combinations of features belonging to different objects, otherwise known as “illusory conjunctions” (Triesman and Schmidt, 1982)...In addition, evidence from neuroanatomy and neurophysiology indicates that processing streams in the visual system are segregated, so that feature dimensions such as color, motion, location and object identity are processed in separate brain regions. (Roskies 1999: 7)

Moreover, as Roskies notes in her introduction to the binding problem in a special issue of *Neuron*, that the binding problem is not merely limited to visual binding but different types of binding. There are other problems of perceptual binding in other modalities as well as what she calls “cognitive binding”: which includes

relating a concept to a percept, such as the linking of visual representation of an apple to all the semantic knowledge stored about it (it is edible, how it tastes, used in pies etc.), cross-modal identification, such as being able to identify an item that has previously only been seen by how it feels memory reconstruction the linking of previously encoded information to form a structured and unified representation. (ibid., 7)

### **6.3 Problems of Identity**

#### **6.3.1 Problems with Essential Properties of Fictional Characters**

The prime motivation for postulating fictional characters as constituted by bundles of mental representations that encode mind-dependent properties comes from responses to views like Wolterstorff’s (1980) that hold that fictional characters have core essential properties.

Another motivation comes from lacunae in Thomasson (1999)’s abstractionism. As we shall see, the identity of a fictional entity is interest-relative because of the “relevant intents and purposes” clause in  $(MET_{FIC})^{**}$ . The relevant intents and purposes can involve conventions as well as contextual knowledge.

An objection van Inwagen (2003) brings to Wolterstorff is as follows. Wolterstorff holds the view that fictional characters are kinds of persons. Each kind has properties and each property “incorporated by a character is essentially incorporated by that character.” (van Inwagen 2003: 152) If Alice essentially incorporates the property of *being asked the riddle* “*Why is a raven like a writing desk?*” (ibid., 152), then Lewis Carroll could never have substituted this property in another draft with the property of *being asked the riddle* “*Why is*

*a raven like a rolling-pin?* (*ibid.*, 152) and still retained the same character. Another character would have been created. Moreover, characters in incomplete stories would vanish each time the manuscript is revised by the author.

Thomasson too finds a problem with identifying “core properties” of fictional characters in Wolterstorff’s account. The objection is that in the play *Doctor Faustus* by Marlowe, the Faust character who struck a pact with the devil is called “Faust”, etc. But suppose one writes a play in which he does not have such a pact and is just a quiet scholar. And he might not be called “Faust” but “Phaust” in another play similar to Faust. If the Faust character has core properties, then “Phaust” and “Faust” can never be the same character under any circumstance.

One can still revise Wolterstorff’s theory by allowing for the continuity as well as the difference between the two texts in Thomasson’s case and the contingent nature of the incorporated property in van Inwagen’s case. One way might be to say that all the core or incorporated properties are merely accidental, none of them are essential at all. But if they all are accidental, then what might account for the reidentification of the character despite differences, like in the Alice case?

Suppose that if these properties are together and all accidental, then a way to place them together would be to say that they are bundled. They are bundled because they are encoded by a bundle of mental representations. The reidentification in the Alice case happens because Carroll decides to count a different bundle of properties as the same character and in this context, due to a convention of authorial authority, his deciding to do so makes it the same character. One can also explain the changes by means of contextual rules, rules of genre and conventions in general. In the “Phaust” example, the latter work will be constrained by the genre, let’s say parody. This will explain the continuity in terms of the inversion of properties (to bring in oppositions here -- overambitious vs quiet, pact with the devil vs no-pact with the devil as well as the change in for ‘f’ letter vs ‘ph’, something which is done in parody). Therefore, the characters are the same, constrained by a set of rules, one would recognise a new bundle of properties as the same character. However, it seems strange that these bundles are either concrete properties in the external world or in Platonic heaven, especially since there are issues of compresence or primitive relations (that is, relations that tie properties to so that they constitute an object) relating property bundles on both accounts. The main problem with such analyses is that these relations are also exemplified by properties, so a regress begins as to what binds these properties. (Orillia & Swoyer 2016) This is one reason why I propose that given that a naturalistic explanation for

this sort of phenomenon can be provided in terms of mind-dependent properties, it is plausible to think that the content of mental representations that constitute fictional characters are bundles of mind-dependent properties (each property encoded by each representation) and reidentification of fictional characters is interest relative, partly based on conventions and rules<sup>7</sup>. There is also a further explanation that the properties are bound with each other because the mental representations that encode them are bound through neural binding.

It must be noted that the identity of the bundles of mental representations themselves is not interest relative, since each bundle that has specific properties is identical with itself at a specific time. The point is clearer when we ask two questions with respect to what makes the fictional character what it is. In one sense, when we ask the question as to what makes it what it is, we are asking, what are the constituents of the world that constitute it? The answer is, that at any time *t*, a fictional character is constituted by a bundle of mental representations, each of which encodes a mind-dependent property. Each bundle of mental representations is identical with itself at time *t*. In the other sense, we ask another question: what makes a fictional character that particular fictional character and not another fictional character? In this sense, the identity is interest relative in that it is not an intrinsic property that defines the identity of the fictional (it may guide and constrain us, however), it is our interests that determine whether a particular bundle of properties is a particular character or not. In other words, fictional characters are like games of chess in that the conventions (what Searle (2006) calls constitutive rules like “*x* is counted as *y* in context *c*”) determine their identity.

### **6.3.2 Problems with A Necessary Condition for Identity**

In the above discussion, the motivation for coming up with interest relative identity and bundles of properties for fictional characters was because necessary and sufficient conditions for core properties of fictional characters failed. On the other hand, to avoid problems with core properties, Thomasson brings up a necessary condition for identity across works *K* and *L* respectively:

The author of *L* must be completely acquainted with *x* of *K* and intend to import *x* into *L* as *y* (Thomasson 1999: 67)

This condition can be incorporated in the conventions I mention, but merely as one of the conditions for reidentification of a character. Consider the Menard case from Borges.

---

<sup>7</sup> Lamarque (2010) also has an interest relative notion of the identification of fictional characters.

Suppose Pierre Menard writes *Don Quixote* word for word at the same time as Cervantes and both are unaware of each other doing so. One common intuition is that they are two different characters. The counterexample I give to that intuition is as follows. Suppose many years later, Friar Smith discovers both manuscripts but does not know who the authors are. Friar Smith now goes to write a sequel, having read both novels, thinking he is importing the same Quixote. In this case, it is arbitrary to say the character in the sequel is identical with either Menard's or Cervantes, so one can therefore say that it is likely that he has imported one and the same Quixote (My takeaway from this that it is due to the Friar Smith context, including the sequel written that the identity is the same. Interest relative identity comes to play its role, so based on knowledge of context, he construes it to be the same. Knowledge of different origin might account for interest relative identity of Menard's and Cervantes' Quixotes being different. That too depends on context and rules determined by contexts).

Now unfortunately, Thomasson does not go any further to explain why in the Friar Smith case, similarity of properties can make Friar Smith construe the Menard and Cervantes Quixote as one and the same. All she does is to set out one of the constraints as to why Friar Smith makes his character identical to both the Quixotes and can import them. She further mentions that the reason for there being just a necessary but not sufficient condition is an unintentional reference shift. In her example, suppose a rock or dog in a story is named "Sherlock Holmes" just with an old name for a new character without any reference to Sherlock Holmes, then this is an absolutely new character. In this, she misses out the idea that the reason why it does not seem to be the same reference is because relevant properties are absent. But there are similar properties too. Possessing the same name and the emotional associations with the Holmes character it will give rise to (since names do give rise to emotional associations) make it an instance of the same character type (the Holmes type). This is explained by the fact that there is a common property (*being a thought of "being called "Sherlock Holmes" for instance*), even if the dog case is very marginally an instance of the Holmes type. But this is not explained by Thomasson because for her, it just happens to be a new character. Thus, in this context, given the right kind of conventions regarding character types, the fictional dog named "Sherlock Holmes" from Thomasson's example can be "counted" as an instance of the same character type due to similar properties and the reference shift can be explained because it is counted as a different character. The difference though, between Wolterstorff and my account is that in here, identification of properties is a sufficient but not a necessary condition for identity of characters, that too only in particular contexts. This explains why, in the Friar Smith Menard case, the same properties are counted as the same character and identity of properties (as I prefer to think of mind-

dependent properties as tropes, two tokens are identical based on exact resemblances) can be sufficient for a bundle of properties to be counted as the same character.

Thomasson also writes that in 1957, a man named Fred Jones who does not know about Fielding and Richardson's Pamela (which are the same character, since Fielding's Pamela is a parody of Richardson's) describes a character called Pamela who has some similar properties (like a squire trying to seduce her). Fred Jones' Pamela is not the same character as Fielding's and Richardson's: because Thomasson's acquaintance condition is absent. But consider this. Many years later, again, Friar Smith, who has a predilection for manuscripts, discovers all three of them, and again thinks they are by the same author. Thinking they are the same, and thinking there's a continuity, he creates a sequel of Richardson's and Fielding's and Fred Jones' Pamela. There also no way to say there is no continuity, so later on a smart alec cannot point out, "Look, Fred Jones' and the older Pamela are different." This is also a counterexample to Thomasson's acquaintance condition: Fred Jones' Pamela turns out to be identical even if he is not acquainted with any previous work with Pamela in it nor does he have any intention to import it.

The lesson I want to bring out of this discussion is as follows. In both the cases, that is, the Menard and the Fred Jones cases I brought up, without knowledge of authorship, the fictional characters turn out as the same. Furthermore, it looks like whereas one might have a clearer intuition that two characters are different before my counterexamples, they would at least be tempted to wonder whether they are the same, if not consider them as the same.

In my view, what explains this change of intuition is that the interest relative identity is explained by the possession of what Fred Landman (1986) calls "information states". For instance, in the Fred Jones case, one possesses relevant knowledge of authors as well as knows that Fred Jones' Pamela is different because he does not fulfil the knowledge condition. So the identity with respect to possession of this knowledge is different from cases where there was no relevant knowledge of the authors. This also goes to show that in default conditions like the Menard case, in the absence of further contextual knowledge like knowledge of authorship, the Friar Smith case points to all three Quixotes being the same character. In the Fred Jones case, if there is sufficient similarity between representative or key properties of the characters, then it looks like he will think of them as the same. This will then form the basis of the sequel as well as the latter intuition that they are the same character. Even in Thomasson's case of the dog named Sherlock, even marginal similarity of a salient property allows us to observe a mild instantiation of the character type Holmes.

### **6.3.3 Interest Relative Identity**

There is hardly anything new about this proposal. Chomsky writes on the role of human interests in determining the identity of referents:

Even the status of (nameable) thing, perhaps the most elementary concept we have, depends crucially on such intricate matters as acts of human will, again something understood without relevant experience, determined by intrinsic properties of the language faculty and others. A collection of sticks in the ground could be a (discontinuous) thing - say, a picket fence, a barrier, a work of art. But the same sticks in the ground are not a thing if left there by a forest fire. (Chomsky 2000: 127)

Substances reveal the same kinds of special mental design. Take the term "water," in the sense proposed by Hilary Putnam: as coextensive with "H<sub>2</sub>O give or take certain impurities" (Putnam 1992, citing his now classic paper, Putnam 1975). Even in such a usage, with its questionable invocation of natural science, we find that whether something is water depends on special human interests and concerns, again in ways understood without relevant experience; the term "impurities" covers some difficult terrain. Suppose cup<sub>1</sub> is filled from the tap. It is a cup of water, but if a tea bag is dipped into it, that is no longer the case. It is now a cup of tea, something different. Suppose cup<sub>2</sub> is filled from a tap connected to a reservoir in which tea has been dumped (say, as a new kind of purifier). What is in cup<sub>2</sub> is water, not tea, even if a chemist could not distinguish it from the present contents of cup. The cups contain the same thing from one point of view, different things from another; but in either case cup<sub>2</sub> contains only water and cup<sub>1</sub>, only tea. In cup<sub>2</sub>, the tea is an "impurity" in Putnam's sense, in cup, it is not, and we do not have water at all (except in the sense that milk is mostly water, or a person for that matter). If cup<sub>3</sub> contains pure H<sub>2</sub>O into which a tea bag has been dipped, it is tea, not water, though it could have a higher concentration of H<sub>2</sub>O molecules than what comes from the tap or is drawn from a river. (*ibid.*, 127-128)

In the quotes above, not only do we find the interest relative identity I have been talking about, but it is our information states about human conventions, invoked by our interests that determine the meaning and identity of referents (of course, our information states are also constrained and determined by our cognitive mechanisms and knowledge of the world). In the case of water, while it is true that the actual structure of the water molecule constrains what we refer to, what constitutes what we refer to by "water" is determined by our interests and purposes. While both cups have are made of the same physical constituents, cup<sub>1</sub> is tea and cup<sub>2</sub> is water. If our interest is scientific, and we ask the question what is it about the world that makes liquids in both cups the same, we will talk about chemical composition. In ordinary contexts, however, they will be tea and water respectively. Similarly, in the case of the collection of sticks, from one viewpoint it is a collection of sticks, from another a barrier. Analogously, when we ask the question of what fictional entities are constituted by with respect to constituents in the world, the answer is that they are constituted by bundles of mental representations and that each particular bundle token is numerically identical with itself. But more pertinently, two tokens of the same types of property bundles encoded by mental representations would be identical (not numerically) since they exactly resemble each other. However, even if there is such identity, like we saw in the Menard case above,

where Friar Smith's, Menard's and Cervantes' Quixote consisted of the same type of properties, our interests and knowledge of context will determine whether they are the same fictional character or not. This explains why Friar Smith thinks they are the same without knowledge of authorship. But with respect to Fred Jones' Pamela, in the absence of Thomasson's acquaintance condition, we saw that what made the same type of bundle of properties count as two different characters (just like the case of tea and water above) was a certain convention for character identification.

Thus, the general point I make here is that in the absence of constraining contextual knowledge, it looks like the default condition is to count two separate but identical bundle tokens as the same character or to say that the character type is identical. This explains why theorists postulate core properties of characters. However, since it does not follow that there is an essence of fictional characters, the interest relative identity is dependent on contextual knowledge that we possess in our information states. Furthermore, that identification is interest relative and that under certain conditions exact resemblance of properties is sufficient to allow for identity of characters and under other conditions, that it is not sufficient, also accounts for why Thomasson is only able to come up only with a necessary condition for identity across stories but not a sufficient condition. However, my account explicitly would say there is no set of necessary and sufficient conditions for the identity of fictional characters.

#### **6.4 The Problem of Internal and External Perspectives**

Friend (2007) poses a serious problem for abstractionist theories of fictional characters who also happen to be external realists. Since statements like "I pity Anna Karenina" and "Holmes is more brilliant than Poirot" (*ibid.*, 151) cannot be prefixed by the "according to the fiction" operator, external realists have to take them as literally true. Therefore, the fictional characters have to be individuated from the external perspective. The problem that Friend points out with this is that from the external perspective, I cannot literally pity an abstract object. It is only from a perspective internal to fiction that I can see her as a person who can suffer and who is worthy of my pity. In the latter sentence too, neither Holmes nor Poirot can be literally brilliant. The external realist theory cannot accommodate such mixed statements and therefore the strong distinction between external and internal perspectives collapses.

However, my account allows for a uniform discourse. There is no strict distinction between internal and external perspectives in my account. All statements about fictional characters, (where *x* is a fictional character) presuppose:



(MET<sub>FIC</sub>)\*\* There is an x such that x is fictional and x is likened to a real object and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

Statements about characters who are persons in the fiction more specifically presuppose:

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

My account also explains the difference between the salience of the more abstract properties in statements that tend to be found in literary critical statements and those properties that belong to persons by invoking the primacy of Dennett's design stance in the former and the intentional stance in the latter. In literary criticism, we do use the intentional stance to understand and analyze belief-desire psychology. In order to even make these predictions properly, we cannot but think of the characters as embodied. One has to first assume that a mental representation is likened to a flesh and blood human being because I will have to imagine conditions to which they respond in an embodied manner and have beliefs and desires accordingly. But the reason why properties at a higher level of abstraction become salient in these contexts is because of the primacy of the design stance. On the other hand, contexts like "I pity Anna Karenina" as well as those in more person-like characteristics of fictional characters are prominent, the intentional stance would have primacy.

### **6.5 Advantages over Meinongianism and Possibilism**

The account I propose does not face the selection problem, since there is no non-existing object but only mental representations which encode mind dependent properties, and the properties are merely selected from pre-existing properties. As shown above, my view also accounts for the objection that we saw Sainsbury bring against Meinongianism, that it does not account for the creation of fictional characters and addition of properties. Additionally, interest relative identity also accounts for the creative process where the author is trying to think of a character and at various times changes his mind (with respect to its properties) as to what the character should be while retaining the identity of the character, thus accounting for the truth of statements such as (9).

(9) Sherlock Holmes might not have become a detective.

The problem in nonactualism is how to identify non-actual possibilia which are fictional objects. As mentioned earlier, my account denies the presence of such possibilia the way Lewis would have them, at least for fiction. What is possible is the set of ways a fictional object might have been, which implies that different mental representations encoding

distinct properties could have been chosen and combined in different ways to create different bundles of mental representations. This then accounts for the truth of a statement like:

(10) I initially conceived of Sherlock Holmes as a troll with a perpetual grin who solves crimes.

## **7. Conclusion**

I have laid out the general outline and motivation for why I not only take up Everett and Schroeder's account that fictional characters are ideas but also why I incorporate a fine grained view of mind-dependent properties that are encoded by mental representations that solve some of the issues faced by other theories. In the course of my thesis, I shall outline key views on fictional entities, namely Peter van Inwagen's, Amie Thomasson's, and Nathan Salmon's abstractionist theories as well as Kendall Walton's Pretense Theory. In doing so, I will not only elaborate on objections to them that can be found in the literature, but I will also attempt to modify their views in the course of the dialectic and further show how my account has advantages over them. The pattern of my dissertation, the pattern of each particular chapter as well as the way arguments are laid out somewhat form that of a Hegelian dialectic, where to a thesis proposed by a philosopher, I present an anti-thesis, which is a response by another philosopher or my own response. The anti-theses are what I call the "discontents" to the theses. What I do is then is to wrestle with both the thesis and the antithesis and give out a synthesis, whether it is by imagining a possible response by other philosophers or proposing my own synthesis. In the final chapter, which is the attempted synthesis, I will lay Everett and Schroeder's argument for why a realist theory would do better if it takes fictional characters to be ideas and not abstract objects. In the light of this, I will elaborate my view of mind-dependent properties and bundles of mental representations.

Consistent with that, I will also lay out Laurie Karttunen's (1976) view of discourse referents and a Neo-Kantian view given by Ermanno Bencivenga (1983) in which we refer to entities in our conceptual spaces. I will argue that this best makes sense of our cognitive mechanisms and how this is the best theory of reference for a theory like mine that postulates fictional characters as constituted by mental representations. I will end by attempting to resolve a few key objections my account faces and put out a key advantage of my view.

In sum, my project is as follows. Amie Thomasson (1999) starts out with the question that if we postulated fictional entities, what would they be, given our beliefs and literary practices? I reformulate that question and ask “If fictional characters existed, what would they be in the light of our cognitive mechanisms?” In doing so, I squarely situate myself in the tradition of naturalized metaphysics, which is not only empirically informed but also eschews entities postulated outside the causal nexus of the world.

But last, but not the least, it needs to be mentioned why the subtitle of the thesis is “A Prolegomena to Any Future Metaphysics of Fictional Entities”. The chief philosophical thrust of my thesis is to bring a plethora of cases and arguments, some speculative, that I hope will be grappled with by other theories, so they can account for them. Moreover, for those who find some of the key things I say plausible, I hope they can refine them. For those who think I am wrong, I hope they are challenged, since I am in fact challenging some of the deeply held views in the literature. As for some of claims I am making the grounds of empirical theories, I hope they can be buttressed by further empirical work or refuted by it.

## Chapter 2

### Van Inwagen's Fictional Creatures and Their Discontents

#### 1. Introduction

Fictional abstractionism is the view that fictional characters like Sherlock Holmes exist and that they are abstract objects. Peter van Inwagen, one of the key proponents of this view, in his influential paper "Creatures of Fiction" (1977), outlines his theory of fictional entities. Although van Inwagen's position has been standardly viewed as a creationist view, in which fictional objects are seen as abstract objects created by us, the argument in van Inwagen (1977) leaves it open between them being eternal uncreated necessary objects and created objects (although it seems to veer more towards creationism). However, van Inwagen (2018) clearly rejects the creationist view and accepts the former.

Van Inwagen adopts Quine's (1948) strategy for ontological commitment, arguing that since there are true literal statements that cannot be further paraphrased, and since objectual quantification over fictional entities is required in order to account for the truth of these statements, then fictional entities must exist. In the course of his arguments, van Inwagen makes a strict distinction between discourse *within* fiction (which I shall call "intrafictional statements"), in which statements about literary characters cannot be literally true and discourse *about* fiction (which I shall call "metafictional statements), like that of literary criticism, in which statements about fictional characters can be true. Fictional characters are postulated to account for the truth of the latter. This analysis runs into trouble when it is seen that statements that were thought to be literally true are not really literal. (Yagisawa 2001)

In what follows, I outline van Inwagen's argument and lay out the details of his view as well as objections that have been raised against him. In light of these objections, especially Yagisawa's, I present a modified version of van Inwagen's view and show that my modified view is better equipped to deal with objections than his original view.

First, I will provide a uniform analysis to account for the apparent truth of fictional and metafictional statements. A uniform analysis would mean that even though almost all metafictional and intrafictional statements about literary characters are not literally true, appealing to the presupposition (METFIC)\* in both contexts accounts for their truth.

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

The metaphorical relation “x is likened to y” is an ontological metaphor (Lakoff and Johnson 1980), the sort used in THE MIND IS A BRITTLE OBJECT (presupposed by statements like “the experience shattered him) and COMPANIES ARE PERSONS, where entities are conceived in terms of concrete objects so that we can talk, quantify over, and reason about them.

Second, if the nature of the abstract fictional objects postulated by van Inwagen is underdetermined, then I propose that Everett and Schroeder’s (2015) view that fictional characters are ideas is a plausible and more naturalistic alternative. I also touch upon my fine grained view of Everett and Schroeder’s account that ideas are constituted by bundles of mental representations and their content is given by mind-dependent properties.

## **2. Peter Van Inwagen’s Account of Fictional Characters**

### **2.1 Van Inwagen’s Argument for the Existence of Fictional Characters**

In this section, I lay out van Inwagen’s (1977) argument for the existence of fictional characters based on the reasoning that the only way to account for the truth of certain statements of literary criticism is by quantifying over fictional characters. Van Inwagen puts forward the following sentences for consideration (*ibid.*, 301):

(1) She was a fat old woman, this Mrs. Gamp, with a husky voice and a moist eye, which she had a remarkable power of turning up, and only showing the white of it. (Martin Chuzzlewit, XIX)

(2) Mrs. Sarah Gamp was, four-and-twenty years ago, a fair representation of the hired attendant on the poor in sickness. (From Dickens's preface to an 1867 edition of Martin Chuzzlewit)

(3) Mrs. Gamp ... is the most fully developed of the masculine anti-women visible in all Dickens's novels. (Sylvia Bank Manning, Dickens as Satirist [New Haven, 1971] P- 79).

His basic argument for the existence of a particular fictional character such as Mrs. Gamp is as follows:

(P1) Metafictional sentences like (2) and (3) cannot be paraphrased further to eliminate purported reference to characters.

(P2) There are characters in novels.

(P3) If there are characters in novels, there is something such as Mrs. Gamp.

(C1) From (P2) and (P3), there is something such as Mrs. Gamp.

(P5) If (P1) and (C1) are true, then (2) and (3) are about Mrs. Gamp.

(P6) (P1) and (C1) are true.

(C2) Therefore (2) and (3) are about Mrs. Gamp.

Van Inwagen seeks to further justify (P2) to answer the question: Why are sentences (2) and (3) about Mrs. Gamp? Thus, if fictional characters exist, in general, there is an explanation for saying that Mrs. Gamp is one of those characters. (*ibid.*, 301) His justification for (P2), which would show that fictional characters exist, is as follows.

Van Inwagen puts forward the following sentences for consideration (*ibid.*, 302):

(4) There are characters in some 19th-century novels who are presented with a greater wealth of physical detail than is any character in any 18th-century novel.

(5) Some characters in novels are closely modeled on actual people, while others are wholly products of the literary imagination, and it is usually impossible to tell which characters fall into which of these categories by textual analysis alone.

(6) Since 19th-century English novelists were, for the most part, conventional Englishmen, we might expect most novels of the period to contain stereotyped comic Frenchmen or Italians; but very few such characters exist.

The truth condition for (4) is: (4) is true iff there are characters in some 19<sup>th</sup> Century novels who are presented with a greater wealth of detail than is any character in any 18<sup>th</sup> Century novel. Therefore (4) is true only if there are characters in novels. Van Inwagen translates (4) into the quantifier variable idiom as follows (*ibid.* 302):

(4)\*  $(\exists x)(x \text{ is a character in a 19th-century novel} \ \& \ (\forall y) [y \text{ is a character in an 18th-century novel} \ \supset \ x \text{ is presented with a greater wealth of physical detail than is } y])$ .

This entails (*ibid.* 302):

(7)  $(\exists x)(x \text{ is a character in a 19th-century novel})$

Thus, there exist such things as characters in 19<sup>th</sup> Century Novels. It follows from this that there exist characters in novels.

## 2.1 Classification of Fictional Characters

In his classification of fictional characters, van Inwagen puts them under the ontological category of creatures of fiction. Apart from Mrs. Gamp, other fictional entities like the Forest of Arden also belong to this category. Creatures of fiction, in turn, belong to a larger category of theoretical entities of literary criticism. Other examples of theoretical entities of literary criticism include plots, novels, poems, meters, rhyme schemes, and the like. Van Inwagen does not mention anything about their nature as such but describes them by analogy with theoretical entities of physics, like electrons. Specialized fields have sentences with theoretical vocabulary that contains the “conceptual machinery” of those disciplines. Thus, physics has true sentences and for those sentences to be true, their theoretical terms must have extensions. The extensions of these theoretical terms are the theoretical entities of physics.

Similarly, sentences of literary criticism like (4), (5), and (6) are true and for them to be true, van Inwagen says that the extensions of their “theoretical general terms” (*ibid.*, 303) are the theoretical entities of literary criticism. Furthermore, he thinks that (4), (5), and (6) cannot be paraphrased further to those sentences that do not quantify over creatures of fiction.

## 2.2 Van Inwagen’s Argument Against Paraphrase

Van Inwagen’s support for the first premise in his argument for the existence of particular fictional characters like Mrs. Gamp, namely that “metafictional sentences like (2) and (3) cannot be paraphrased further to eliminate purported reference to characters” is as follows. Suppose someone attempts to paraphrase (4), that is, the sentence “There are characters in some 19<sup>th</sup>-century novels who are presented with a greater wealth of physical detail than is any character in any 18<sup>th</sup>-century novel” and accepts novels in their ontology but rejects creatures of fiction. So the paraphrase strategy will involve an open sentence “*x dwelphs y*” which is satisfied only by novels but not fictional characters. Thus van Inwagen says that the ordered pair <the class of 19<sup>th</sup> century novels, the class of 18<sup>th</sup> century novels> satisfies the open sentence “*x dwelphs y*” iff the proposition expressed by (4) is true. (*ibid.*, 303) The proposed paraphrase for (4) with only novels in the ontology would be (*ibid.*, 303):

(4)\*\* The class of 19<sup>th</sup> Century novels *dwelphs* the class of 18<sup>th</sup> Century novels.

Now (4) entails (8), since (8) is an English translation of a sentence (in quantifier variable idiom) that is a logical consequence of the quantifier variable idiom of (4), which is (4)\*(*ibid.*, 303):

(8) Every female character in any 18th century novel is such that there is some character in some 19<sup>th</sup> century novel who is presented with a greater wealth of physical detail than she is.

Furthermore, suppose someone who accepts novels in their ontology but rejects creatures of fiction were to paraphrase (8) in the same way (4) was paraphrased as (4)\*\* , there would be a new open sentence “x *praphs* y” to eliminate reference to fictional characters. Thus, as per van Inwagen, the ordered pair <the class of 19th century novels, the class of 18th century novels> satisfies the open sentence “x *praphs* y” iff the proposition expressed by (8) is true. (*ibid.*, 304) The paraphrase of (8) would be the following (*ibid.*, 304):

(8)\* The class of 18<sup>th</sup> Century novels *praphs* the class of 19<sup>th</sup> Century novels.

However, the paraphrases above do not express the entailment relation from (4) to (8) because there is no way of accounting for how (4)\*\* entails (8)\*. Therefore, van Inwagen lays out the following condition that any paraphrase strategy must adhere to:

(PC) An adequate paraphrase must not be such as to leave us without an account of the logical consequences of (the propositions expressed by) the paraphrased sentences. (van Inwagen 1977: 304)

One objection to his view that van Inwagen anticipates is as follows (*ibid.*, 304):

(P1) If Mrs. Gamp existed in 1843, then she had the following properties: being old, being fond of gin, being named “Sarah Gamp”, having a friend called “Mrs. Prig”.

(P2) No one in 1843 had every single one of those properties.

(C) Therefore, Mrs. Gamp did not exist in 1843.

Van Inwagen rejects P1. He claims that Mrs. Gamp does not *have* or *exemplifies* the properties but *holds* them (van Inwagen 1983) and that the abstract object, which is the creature of fiction, bears another relation to the properties which he calls “ascription”, in contrast to “predication”. Thus, Mrs. Gamp bears the relation, that of ascription, to the above properties such that one can talk about a theoretical entity of literary criticism with the sentence “Mrs. Gamp is fond of gin” without predicating *being fond of gin*. In other words, one may ascribe fondness for gin to Mrs. Gamp without predicating it. Van Inwagen describes the ascription relation with the open sentence “x is ascribed to y in z” where “x is a property, y is a creature of fiction” (*ibid.*, 305) and z is a “place” where place is a work of fiction, a part of it, a sentence or a clause in it. The ascription relation is thus A(x, y, z). The



place *z* is added in order to account for cases like contradictory statements in the same work of fiction, where in one place it says “Shiela Smith was fat” and another that says “Shiela Smith was not fat”. (*ibid.*, 305)

“Mrs. Gamp is fat” expresses either of the propositions: “A (fatness, Mrs. Gamp, *Martin Chuzzlewit*) or A(fatness, Mrs. Gamp, Chp XIX of *Martin Chuzzlewit*) or A(fatness, Mrs. Gamp, the only occurrence of (1) in *Martin Chuzzlewit*).” (*ibid.*, 305) It is important to note that when Dickens was writing (1), that is “She was a fat old woman, this Mrs. Gamp, with a husky voice and a moist eye, which she had a remarkable power of turning up, and only showing the white of it”, he was not ascribing a property to a theoretical entity of literary criticism. This is because in writing (1), he did not express any proposition. If he had ascribed a property, he would have expressed a proposition. Nevertheless, having noted this, the truth condition for “Mrs. Gamp is fat” is:

“Mrs. Gamp is fat” is true iff A (fatness, Mrs. Gamp, *Martin Chuzzlewit*) holds.

Van Inwagen considers the ascription relation to be primitive. He writes: “A(fatness, Mrs. Gamp, *Martin Chuzzlewit*) is nothing more than the way I choose to express what we normally express by the above sentence (“Mrs. Gamp is fat”...), and that is all the explanation I am able to give of the use of “A(x, y, z).” (van Inwagen 1977: 306) One reason that he gives in support of the ascription relation being primitive is by considering one possible explanation of the ascription relation. The explanation is as follows: “A (fatness, Mrs. Gamp, *Martin Chuzzlewit*) means that “Mrs. Gamp is fat” iff *Martin Chuzzlewit* were not a novel but a record of events, then there would be a woman called “Mrs. Gamp” and she would fat.” (*ibid.*, 306) Since, the ascription relation is equivalent to “A(fatness, the main satiric villainess of *Martin Chuzzlewit*, *Martin Chuzzlewit*)” (*ibid.*, 306), then the above conditional is equivalent to: “If *Martin Chuzzlewit* were not a novel but a true record of events, then there would be a woman denoted by “the main satiric villainess of *Martin Chuzzlewit*” and she would be fat.” (*ibid.*, 306)

However, this is problematic since the phrase “the main satiric villainess of *Martin Chuzzlewit*” wouldn’t denote any woman described by it.

Van Inwagen’s explanation of why this problem arises with the above account of ascription is as follows. Ordinarily, objects have proper names by virtue of being named in a baptism-like act, in which they are named through ostension or by the fixing of the reference by

means of a description<sup>8</sup> (Kripke 1972/1980). But unlike them, there is no such baptism for “the main satiric villainess of Martin Chuzzlewit”. The advantage of Van Inwagen’s theory is that it accounts for the reference of the definite description because there is, in fact, a unique object that satisfies it (in the way the Eiffel Tower satisfies the description “the tallest structure in 1905”, which is how the description succeeds in denoting). Van Inwagen claims that for there to be a theory of successful reference to fictional characters, there have to be definite descriptions which serve as the “primary means of reference to those objects” and proper names “as a secondary (though more common) means of reference.” (*ibid.*, 307) Thus, “the only reason “Mrs. Gamp” denotes a certain creature of fiction is that this creature of fiction satisfies the open sentence “A (being named ‘Mrs. Gamp,’ x, Martin Chuzzlewit).” (*ibid.*, 307) The problem for the explanation for ascription arises and is circular precisely because, unlike a real proper name which has a clear cut explanation for how it refers, the explanation for how a “fictional proper name” like Mrs. Gamp refers is that it refers only because it already involves an ascription relation (the satisfaction of the open sentence A (being named “Mrs. Gamp”, x, Martin Chuzzlewit). (*ibid.*, 307)

### 3. Yagisawa’s Objections to Van Inwagen

Yagisawa (2001: 63) summarises Van Inwagen’s argument thus (where  $\alpha$  is a fictional singular term,  $\Phi$  is “a sentence of literary criticism with the relevant existential import” (*ibid.*, 163<sup>9</sup>):

(P1) It is a truth of literary criticism that  $\varphi$ .

(P2) That  $\varphi$  entails that  $\alpha$  exists.

(C) Therefore,  $\alpha$  exists.

Yagisawa calls this argument invalid since at most, the premises entail (C’) if one assumes closure under entailment of the truths of literary criticism (*ibid.* 163):

(C’) It is a truth of literary criticism that  $\alpha$  exists.

To make van Inwagen’s argument a valid one, Yagisawa appeals to van Inwagen’s taking the sentences of literary criticism at “face value” and the consequent analogy between accepting the theoretical entities of physics and that of literary criticism. If we take

---

<sup>8</sup> Kripke (1972/1980: 79) gives the example of Neptune which is fixed by description. So instance, Le Verrier fixes the reference of Neptune as “the planet that causes such and such disturbances in the orbits of certain planets” before he even saw the planet by means of a telescope.

<sup>9</sup> The argument is given in the same page.

sentences of literary criticism at “face value” then, if it is a truth of literary criticism that (4) (“There are characters in some 19th-century novels who are presented with a greater wealth of physical detail than is any character in any 18th-century novel”), then it is true that (4). This allows us to remove the sentential operator, “it is a truth of literary criticism that” (*ibid.*, 163). Thus, the Yagisawa points out that it is the additional lemma (P1’) makes the above argument valid.

(P1’) (It is true that)  $\varphi$

But even then, Yagisawa does not accept the move from (P1) to (P1’). This is because, according to him, the main purpose of literary criticism is not to describe the world but to enhance aesthetic appreciation and experience. Propositional truths like “Dickens lived in the 19<sup>th</sup> Century and wrote in English” (*ibid.*, 164) might happen to be expressed in literary criticism, but they are merely incidental to it, as their truth is determined by a discipline that describes the actual world, like history. Moreover, if literary critical statements are taken at “face value”, then a sentence like (3) (“Mrs. Gamp ... is the most fully developed of the masculine anti-women visible in all Dickens's novels”) will entail (*ibid.*, 164):

(3’) Mrs. Gamp is a masculine anti woman.

But (3’) is obviously not true, since Mrs. Gamp is an abstract object which cannot be a masculine anti woman. Van Inwagen’s case is further weakened since he allows feminine pronouns to refer to the abstract object Mrs. Gamp when he takes sentences like “Mrs. Gamp is the most developed of the fat old gin-drinking females in all English novels. She is well known among the readers of 19<sup>th</sup> Century English novels.” If we take the above sentences at face value and the abstract object referred to by Mrs. Gamp is neither female nor fat nor gin drinking, then evidently, the sentences turn out to be false. Yagisawa writes

Thus, here is a dilemma for Van Inwagen: Either we can refer to Mrs. Gamp by ‘she’, ‘the fat old’ and the like or we cannot. If we can, then literary criticism should not be taken at face value as the language of criticism is not metaphysically straightforward but bound to the host story in such a way that what might appear to be asserted in literary criticism is heavily parasitic on what is true in the story. If we cannot, then most pieces of literary criticism fail to refer to the intended characters. (Yagisawa 2001: 165)

The first horn of the dilemma is when “we can refer to Mrs. Gamp by ‘she’ ‘the fat old’” (*ibid.* 165) etc. Given Yagisawa’s argument above, we cannot take sentences with these terms at face value. The quantification too seems to be over Mrs. Gamp, a fictional individual, if we are to read the sentence at face value. But this is not the result Van Inwagen is looking for. The quantification at face value on van Inwagen’s account would be over an abstract object, not a fictional individual (that is, the fictional Mrs. Gamp qua Mrs.

Gamp and not qua abstract object). Thus, according to Yagisawa, given that objectual quantification read at face value leads to the wrong result, it is better to understand the quantifier in fictional cases as substitutional, in which “an apparent quantification over fictional individuals is to be interpreted as real quantification over (apparent) singular terms for fictional individuals in the host story.” (*ibid.*, 165) Therefore, since we cannot refer to Mrs. Gamp by “she”, “fat” etc, the second horn of the dilemma holds and reference to fictional characters fails in most metafictional statements.

One objection against this that Yagisawa anticipates might be that in sentences like “The crowd lunged forward” (*ibid.*, 166) there are no apparent fictional singular terms for every person in the crowd. However, the quantification need not be over terms mentioned explicitly in sentences in the story. Just in the way a story can imply that Mrs. Watson has a liver, it can imply that according to the story, every individual member of the crowd does exist, and thus the domain of quantification, will have descriptive terms like “the first member of the crowd to lunge”, “the tallest member of the crowd” (*ibid.*, 166) and the like.

#### **4. Goodman’s Defense of Van Inwagen against Yagisawa**

##### **4.1 Goodman’s First Argument**

Jeffrey Goodman (2004) argues against Yagisawa’s strategy of denying the inference from (P1) to (P1’).

(P1) It is a truth of literary criticism that  $\varphi$

(P1’) Therefore (it is true that)  $\varphi$

So Goodman says that Van Inwagen’s argument is not that if  $\varphi$  is any truth of literary criticism, then  $\varphi$  is true simpliciter, but that only a subclass of truths of literary criticism come out true simpliciter, that is, those which are metafictional sentences.

Therefore, Goodman’s reconstruction of Van Inwagen’s argument is (*ibid.* 137):

(P1’’) There are some sentences of kind  $\varphi$  that are truths of literary criticism.

(C2) Therefore, some sentences of kind  $\varphi$  are true.

(P2’) All true sentences of kind  $\varphi$  entail that entities of kind  $x$  exist.

(C’’) Entities of kind  $x$  exist.

On this construal, Yagisawa’s criticism that (3) taken at face value will lead to Mrs. Gamp being taken as a masculine anti woman and not an abstract object does not hold. This is because (3) is not included in the sub-class of metafictional sentences and is not thus true

simpliciter. One can thus infer from (P1'') to (C2) because sentences like (3) are just unfortunate examples of apparent metafictional sentences that van Inwagen invokes. An example of a real metafictional sentence would be (*ibid.*, 138):

(9) James Tiberius Kirk is a fictional individual

So (9) taken at face value will quantify over an abstract object, given that it is true simpliciter.

However, Goodman's counterexample here (that is (9)), seems to miss out on Yagisawa's general point. Yagisawa's general point seems to be that applying predicates like fat, masculine anti woman etc. to an abstract entity is a category mistake. The reason for this would be a correct application principle like:

(CA) A predicate that correctly ascribes a property that can belong only to a concrete individual can be literally predicated only of a concrete individual.

Given the violation of (CA), the literal application of the predicates "is a fictional individual" and "is a fictional person" is itself a category mistake, since the predicates "is an individual" (at least for a literary critic and not for a philosopher like Goodman for whom ordinary language has gone on holiday) and "is a person" are only correctly applied to something non fictional. Being an individual (based on the dictionary definition of individual in an ordinary context as used by a literary critic which is "a single human being as distinct from a group" (Oxford Dictionary of English (2010)) which entails being a person is and can be predicated correctly only of a person.

(9\*) James Tiberius Kirk is an individual.

(9\*) is true iff James Tiberius Kirk is a person.

Furthermore, let's consider an analogy with companies. If we look at a sentence like "Corporate entities (like companies) are persons", at most, we can only read them metaphorically as saying something like "Corporate entities are to be treated (or are treated) as persons" or "Corporate entities behave like persons". An analogous sentence to (9) can be seen in a textbook definition of a company:

"A company is an artificial person created by law having a separate entity with a perpetual succession and common seal." (by a Professor Haney as quoted by Singh (2015))

Thus, when we read the sentence "A company is an artificial person", it can at most be understood metaphorically, meaning

(MET<sub>COMP</sub>) A company is artificial but is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

In (MET<sub>COMP</sub>), relevant intents and purposes, for instance, can be so that it can enter into contracts and held liable for its actions.

One might object that the expression “artificial person” is analogous to the expression “a rubber duck”, from which one cannot infer “there is an artificial person” in the way one cannot infer that “there is a rubber duck”. But the claim is not that there is an artificial person. From “there is a rubber duck” one can in fact infer that “there is an entity that is likened to a duck”, where there may be some respects in which they may be regarded as having similar properties. Similarly, from “A company is an artificial person” the inference is not there is an artificial person but that there is an artificial entity that is likened to a person.

Analogously, we might read (9) as

(MET<sub>FIC</sub>) James Tiberius Kirk is fictional but is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

But does this metaphorical use and its analogy with (9) undercut Goodman’s objection completely? On one hand, it does appear that (9) cannot be taken at face value or literally. The analogy shows that the predication is metaphorical in nature and not literal. To even apply the predicate “is a fictional individual” literally to an abstract object would be a category mistake. Even “James Tiberius is a fictional character” cannot be taken literally of an abstract object because the meaning of “fictional creature” (“fictional person” in the case of Holmes) will be presupposed<sup>10</sup> by the meaning of “fictional character”. In other words, again, the meaning of “creature” for a literary critic (as well as any ordinary person for whom language has not gone on holiday) will always be literally something concrete. Thus, since the literal application of predicates like “is a fictional creature” and “is a fictional person” involve category mistakes, the literal application of the predicate “is a fictional character” will also involve a category mistake.

Moreover, (9) is the only counterexample that Goodman offers and the only type of sentence that can be accepted as literal, with “James Tiberius Kirk” substituted for any fictional character or thing. In the light of this, perhaps the only possible type of metafictional

---

<sup>10</sup> There should be nothing mysterious about such presuppositions. Consider how the metaphor MIND IS A THING is presupposed by the metaphor MIND IS A BRITTLE OBJECT which in turn is presupposed by statements which can be true like “He was mentally shattered”.

sentence that can be taken literally and at face value is the sentence (where James Tiberius Kirk can be substituted for any character):

(9\*\*) James Tiberius Kirk is an abstract object.

However, if we accept this, then Goodman and van Inwagen would beg the question regarding ontological commitment, since (9\*\*) can only be inferred from statements like (9) or “James Tiberius Kirk is a fictional character”, which have already been shown above to involve category mistakes. So, without presupposing those statements, it will look that they are merely reasserting their ontology, which is what they set out to prove in the first place.

Still, a modification in van Inwagen’s theory can be used as a reply to Yagisawa.

(MET<sub>COMP</sub>) is in fact true and it is not metaphorical that there is a company, and there can be quantification over that. It is only that the predication involves a metaphor. Similarly, one can quantify over James Tiberius Kirk because from (MET<sub>FIC</sub>) one can infer

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

In the light of the above, the modification that will have to be made to the above theory would be that the only possible metafictional sentence that can be taken literally and at face value is:

(9)\*\*\* James Tiberius is fictional.

In this light, it needs to be observed how (9)\*\*\* is similar enough to (9\*\*) so that Goodman and van Inwagen do not need to beg the question.

As we can see, unlike (9\*\*), the predicate “is fictional” can presuppose the existence of an abstract object (it will presuppose the existence of an idea rather than an abstract object on the view I adopt of Everett and Schroeder’s (2015) realist account that fictional characters are ideas. Nevertheless, establishing (MET<sub>FIC</sub>)\* is sufficient for now to save a realist theory, whether one that takes ideas or abstract objects to be fictional characters). The ontological commitment is seen in the quantification in (MET<sub>FIC</sub>)\* and the predicate merely indicates its origin in a work or works of fiction (I use “fictional” only in the minimal sense as “originating in a particular work of fiction or works of fiction”). Taking “is fictional” in the minimal sense prevents one from begging the question and allows for a statement which even an anti-realist would agree that it is made literally in metafictional discourse.

Moreover, “James Tiberius Kirk is fictional” is not equivalent to “James Tiberius Kirk does not exist” because in a true/false exam on Star Trek, “James Tiberius Kirk is fictional” will

be true (versus “James Kirk is mythical” and “James Tiberius Kirk does not exist” which are false) and not for the reason that he does not exist in the real world.

As to the modification in van Inwagen’s theory, in the light of Yagisawa’s objection, and what I have said above, it can be observed that apart from the property *being fictional* that the abstract object has (which is not intrinsic to it, given that to prevent question begging, the predicate was applied only to indicate its origin in a work of fiction), the other properties that van Inwagen thought it to *have*, like *being a masculine anti woman* as well as the gender based pronouns presuppose metaphorical predication. One can only apply gender based pronouns like “she” and predicates like “is fat” only if an object is animate or that one already counts/treats an inanimate entity (which is an abstract object in van Inwagen’s theory) like a real person for all relevant intents and purposes, which in this case are fictional. From what we also observe, the only property it has literally is that of *being abstract* and all the other properties that van Inwagen claimed it to *have* and *hold* are metaphorically attributed.

Thus, my argument against van Inwagen weakens the *have/hold* description and weakens the need for ascription. This will be made clearer as this chapter progresses.

In the light of this, it is pertinent to say here that the presupposition (MET<sub>FIC</sub>)\* is a necessary condition for the truth and falsity of truth-evaluative statements about fictional characters, just in the way that (MET<sub>COMP</sub>) is a necessary condition for the truth and falsity of statements about companies. This is generally seen when we say that a truth-evaluative statement A presupposes B, if A is true, then it is true only if B is true. Similarly if A is false, it is false only if B is true. So A can be false or true only if B is true, and A is truth-evaluative only if B is true. For instance, to say of the King of France that he is bald and for the statement “The King of France is bald” to be true or false, the “The King of France exists” is presupposed. So the “King of France exists” is a necessary condition for the truth or falsity of the truth-evaluative statement “The King of France is bald”. Since there are true and false statements in fiction which are truth-evaluative, and they presuppose (MET<sub>FIC</sub>)\*, the presupposition has to be true. If (MET<sub>FIC</sub>)\* is literally true and says there exists an entity that is fictional and this is presupposed for all truth-evaluative statements about fictional characters, then we can conclude that fictional entities exist. Furthermore, these are likened to concrete persons in the case of fictional characters. Also, as we can see, (MET<sub>FIC</sub>)\* is literally true despite there being a metaphor involved and entails fictional entities.

#### **4.2 Goodman’s Second Argument**



Yagisawa's argument that no truth of literary criticism can be taken out of the scope of the operator "it is a truth of literary criticism that" and be made true simpliciter because literary criticism is not aimed at propositional truth is countered by Goodman as follows.

Firstly, Goodman argues that contrary to what Yagisawa says, van Inwagen does not equate literary criticism and physics as equally fruitful methods for discovering truths. Van Inwagen's claim is merely that insofar as there are at least some truths in literary criticism, the simplest semantics would give objectual quantification like physics over theoretical entities for those sentences to be true. This could be the case even though literary criticism would, for all we know, be an inferior way to discover truths.

Secondly, Goodman points out that there can be an activity whose primary goal is not truth seeking and yet there can be true assertions made in the course of the activity. One can utter "That anvil is headed to your foot" as a true statement to their friend towards whom an anvil is headed, even though the aim of the activity might be practical. Moreover, unlike "interpretative and evaluative" sentences in literary criticism whose goal would be to enhance appreciation, metafictional sentences of literary criticism like (9) that are true simpliciter are more like claims cultural historians make about literary truths.

## **5. Jung and Pellet's Objections to Van Inwagen**

### **5.1 Jung and Pellet's First Objection**

Eva-Maria Jung and François Pellet (2018), like Goodman, do not accept Yagisawa's critique of van Inwagen from the aim of literary criticism being different from science. However, they do show disanalogies between physics and literary criticism. One disanalogy is that the reason why theoretical entities are postulated in physics is to account for natural phenomena. The aim of physics is to look at the way the world works and find out what it is made up of fundamentally. The aim of literary criticism certainly does not require such postulates to explain "fictional phenomena", because its aim is to merely understand works of literature. Therefore, there is no reason why technical terms in literature should be treated any differently for ontological commitment from everyday talk about fictional characters.

The argument put forward by them is as follows:

Within van Inwagen's argumentation, the assumption that fictional characters are theoretical entities of literary criticism is a necessary premise. If fictional characters were not of that kind – like the Meinongian golden mountain – the Quinean ontological commitment strategy would lose its ground. But if this is true, then the criterion for the justification of assuming that fictional objects are existent is obviously a contingent and historically dependent criterion. We are justified in assuming that exactly these objects, which belong to the theoretical

discipline of literary criticism and its traditions, do exist. What would happen, for example, if the golden mountain became an important part of a novel and, thus, a matter of debate in literary criticism? And what would we say about the existence or non-existence of these fictional characters and objects in novels, which, for contingent reasons, have never been discussed by literary critics?" (*ibid.*, 119)

The point above seems to highlight the disanalogy between discourse in physics and literature. When it comes to physics, there is a clear dividing line between theoretical discourse (about electrons etc.) and our everyday talk (about chairs etc.). In literary criticism, however, the above questions about the contingency and historical dependence of van Inwagen's criterion lead us to wonder that it might easily have been the case that that everyday discourse about fictional characters (intrafictional discourse) might also have involved Meinongian entities. It thus seems that the line between theoretical discourse and everyday discourse about fictional characters is more difficult to draw.

Van Inwagen's (2018) reply to this is that the purpose of Quine's criterion of ontological commitment is not to "discover what exists simpliciter" but what "one's statements and theories imply the existence of." (*ibid.*, 229) Thus, if one says "There are characters in some 19th-century novels who are presented with a greater wealth of physical detail than is any character in any 18th-century novel" then this implies the existence of fictional characters. The point of using Quine's criterion is to say precisely that what the Meinongian thought was a nonexistent thing does exist. Van Inwagen writes:

My ontology of fiction is designed to enable Quineans to say "Yes Meinong, the words "the golden mountain" do have a referent but it is not a non-existent mountain, as you have supposed- it is not a non-existent thing that has the property 'being a (golden) mountain', it is rather an existent thing that holds the property 'being a (golden) mountain'. (van Inwagen 2018: 229))

Nevertheless, it seems to me that the line between theoretical and everyday discourse for fiction is more blurred than van Inwagen thinks it is. The evidence for this is found in metafiction, which is "fiction about fiction: novels and stories that call attention to their fictional status and their own compositional procedures." (Lodge, 1992: 206) Interestingly, Lodge also says: "Metafictional writers have a sneaky habit of incorporating potential criticism into their texts and thus "fictionalizing" it." (*ibid.*, 207) For instance, if Dickens was writing metafiction, he could have introduced Mrs. Gamp as "Mrs. Gamp is the most fully developed of the masculine anti-women visible in all" of his novels and then written that "She was a fat old woman, this Mrs. Gamp, with a husky voice and a moist eye, which she had a remarkable power of turning up, and only showing the white of it".

There would be three pertinent things to observe in that. One would be that the pronoun in the intrafictional sentence would be referring to the same Mrs. Gamp as the one in the sentence from literary criticism. The second would be that the literary critical sentence in the

fiction is of the same type as the one that would be used by a critic, and if there can be quantification over the creature of fiction when it is used by the critic, there is no reason why there cannot be quantification in the fiction. Thirdly, the sentence in the fiction is true, just like the one used by the literary critic. If there is quantification over the literary critical sentence within the fictional work, then there is no reason that the next typically intrafictional sentence should not quantify over a creature of fiction, as it is also talking about the same Mrs. Gamp as the previous one. Thus, for these three reasons, the line between theoretical and everyday discourse in fiction is blurred.<sup>11</sup>

Now suppose I do accept van Inwagen's possible reply. If I were Jung and Pellet, I would still press on and say that firstly, the nature of the entity quantified over van Inwagen postulates is underdetermined. It could be the abstract object that van Inwagen is talking about. But it could also be an idea. It would make no difference in quantification. The quantification occurs over an entity, but it is just that its nature happens to be different.

Secondly, let's just substitute *ideas* for *Meinongian golden mountain* in what I have written above as a summary of Jung and Pellet. "If fictional characters were not theoretical entities of literary criticism, but some other kind like *ideas*, then the criterion for justifying their existence as theoretical entities of literary criticism is a "contingent and historically dependent one". If that is so, then what if it was in fact an *idea* which was an entity that became "part of a novel" and also happened to be discussed by critics?"

Having accepted van Inwagen's point about the Quinean method, I would partly reject that the "criterion for justifying their existence as theoretical entities of literary criticism is a

---

<sup>11</sup> On the contrary, Pelletier (2003) uses the case of metafiction and the blurry distinction between discourse in fiction and discourse about fiction to support the anti-realist view. His argument is that since discourse about fiction seamlessly fits into fiction without causing any disruption to the reader while engaging with the world of the narrative, such intrusions of the external perspective are not a tertiary but a central property of fictional narratives. Even if they are not a central property, "at least that fictional narratives have the reflexive capacity of representing themselves as fictional – the capacity of speaking of creatures they represent as fictional characters or as created by authors –without stepping outside the pretense or breaking the rules of fiction." (*ibid.*, 199) Thus, contrary to a realist like van Inwagen, one does not have to stand outside the pretense to make metafictional statements and that fictional narratives involve both external and internal perspectives.

His argument, however, fails because the realist can always reply that the reason why that technique is used is precisely to disavow the pretense within the fiction, so there is an effect on the readers, sometimes to bring the reader into confidence, for humorous effect as in *Tristram Shandy*, to "acknowledge the artificiality of the conventions of realism even as they employ them; they disarm criticism by anticipating it" (Lodge *ibid.*, 207) and many other effects. The effects can be explained precisely because one in fact stands outside the pretense at that point. Moreover, merely because something is an intrinsic part of fictional narratives does not make them the same type of statements, since ordinary assumptions about the actual world and statements about it fill fictional narratives.

“contingent and historically dependent one”. What I would reject is the bit about them being theoretical entities of literary criticism, given that I have argued for the blurring of the line between theoretical and everyday literary discourse. But what I would accept is that quantifying over them in metafictional contexts does in fact reveal ontological commitment. However, that we cannot easily do so for “intrafictional sentences” is purely “contingent and historically dependent”, perhaps because of the way conventions of fiction have developed where the line between theoretical and everyday discourse happened to be more precisely drawn. Moreover, as we saw above, if the criterion for justifying the existence of fictional characters as theoretical entities of literary criticism is “contingent and historically dependent”, then perhaps for all you know, they might be ideas and it is ideas that are plausible candidates for being fictional characters.

My above argument is also consistent with my previous argument against Goodman, and the claim that further predication like masculine anti woman or fat on fictional objects like masculine anti woman or fat, presupposed an entity being treated as/counted as a real person for all relevant intents and purposes. If that is so, then whether it is in typically intrafictional or metafictional sentences, there must be an entity presupposed in both which is treated as a real person on which there is further predication. I have left open a plausible alternative to fictional characters being abstract objects which are theoretical entities of literary criticism, namely that fictional characters might actually be ideas. This will be buttressed to show in Chapter 6 from Everett and Schroeder’s (2015) argument, that ideas are good and plausible candidates to be fictional characters as opposed to abstract objects.

## **5.2 Jung and Pellet’s Second Objection**

Jung and Pellet have another objection to creationism regarding fictional objects. Fictional creatures are created and therefore dependent on human minds, whereas electrons and the like are part of the real world, needed to explain natural phenomena. However, the former are not part of the real world and its causal interactions, but they are “characters in fictional worlds”. The objection seems to stem from the assumption that if they are created, then they ought to have been part of the causal nexus, but they are not, which poses a problem.

This problem is avoided by the later van Inwagen (2018) who gives up creationism and holds that fictional objects are in fact abstract objects which are necessarily existent, independent of human minds, language and the causal nexus. They are also non spatiotemporal.

However, even this revised idea leads to questions like how an abstract object can be a stereotype of a “poorly competent” Victorian nurse. How is it that they are detected or

discovered, particularly given that van Inwagen's argument from quantification gives no such justification for the picture of a relationship between authors and fictional characters?

Van Inwagen's (*ibid.*) response is to give an explanation for how a novel comes to be spatiotemporal. He claims that the novel *War and Peace* exists in all possible worlds (and so does a creature of fiction). What Tolstoy merely did was to write it, publish it and "cause it to be accessible to readers". This is not dissimilar to there existing  $10^{123}$  possible chess games, but only one of them was actually played by Bobby Fischer and Boris Spassky in 1972 in Reykjavik. Van Inwagen says that "these two gentlemen did not reach into some Platonic 'chess heaven' and pull out an abstract object into the world of flux and impermanence. The fact that this abstract object, this game, was actually played was the outcome of a very complex historical process." (*ibid.*, 228)

To the question of how an abstract object can be a stereotype, his reply is merely that of course, stereotypes are abstract objects because they cannot be created or touched.

My objection to van Inwagen's reply that abstract objects are necessarily existent is this. Consider (4), which is contingently true. There might not have been "characters in some 19th-century novels who are presented with a greater wealth of physical detail than is any character in any 18th-century novel." And that is because no author might have written a novel with such a character in the 19<sup>th</sup> Century, perhaps due to a lack of imagination. Or consider "Mrs. Gamp is a masculine anti woman." Dickens might have made her more feminine or not written about her at all. Of course, van Inwagen would reply that it could have existed uninstantiated and Dickens did not "cause it to be accessible to readers." But it does seem implausible to say that the properties the necessary existent *has*, like *being a masculine anti woman* would be a necessary property of that object, given that it is the case that Mrs. Gamp might have been feminine. Van Inwagen might reply that it is a contingent property, but then he has to explain why it is a contingent property. My simple explanation is that it is because it was contingently written that such and such character had such and such property. A further reply might be that these properties are not intrinsic, which explains why they are contingent. But in my argument against Goodman, I have argued that the only properties that the abstract object literally has is that of *being fictional* (which is not an intrinsic property) and *being an abstract object*. All other predications are based on metaphors. In the light of this, what van Inwagen is left with is an eternal abstract object, like a bare particular, has no intrinsic property left except for *being an abstract object*. However, bare particulars and abstract objects that are like bare particulars are hard to accept in one's ontology. But if one rejects bare abstract objects, at least as theoretical entities of literary criticism, and accepts quantification over fictional characters for fiction

and considers them historically and dependent on minds, then there is no reason why a picture cannot emerge of them being bundles of properties, the sort which are part of the causal nexus, having rejected the bare abstract substratum as theoretically unnecessary. Furthermore, since we are talking about fictional objects, it is indeed odd to say that these bundles of properties are identical with a concrete object in the external world, as since they are dependent on minds, one can plausibly say that they properties are likely to be in the mind.

Thus, the picture that emerges becomes akin to that of an empiricist who believes that objects are mind-dependent bundles of properties, having rejected bare substrata. However, there is a problem with this too. If a fictional character  $x$  is identical with a particular bundle of properties  $y$  at time  $t_1$  and the same character  $x$  is identical with another different bundle  $z$  at time  $t_2$ , then there will be a problem with transitivity, since both the bundles are not identical with each other ( $y \neq z$ ). Therefore  $x \neq x$ . I have a parsimonious and naturalistic explanation for why these problems arise and how the problem of transitivity can be avoided. I have proposed in Chapter 1 that fictional characters are ideas (Everett and Schroeder's (2015) view) that are *constituted* by bundles of mental representations in which each representation encodes mind-dependent properties that Thagard (2019) talks about, like sensory properties (*what cats look and sound like*), motor properties (*how it feels to pick up a cat*), emotional properties like (*how much you like cats*) and properties of verbal representations like cats being a kind of mammal. So it's not that fictional characters are identical to bundles of properties, but are constituted by mental representations that encode bundles of mind-dependent properties. This avoids the problem of transitivity. There also seems to be apparent quantification over bundles of properties without a substratum most likely because the properties encoded by bundles of mental representations which constitute fictional characters make it appear so. That is, one might end up thinking it is these bundles of properties in an eternal and Platonic realm because they're reifying the properties encoded by their mental representations, mistaking contents of their minds for entities out there, the sort of mistake the empiricists and Kant tried to correct in the history of philosophy. Furthermore, the explanation for why van Inwagen ends up with a substance that holds properties together is because he misses out on an ontological metaphor, which makes the object of thought appear as a substance. This picture becomes clearer in Chapter 6.

## **6. Solodkoff and Woodward's objection to Van Inwagen**

Tatjana von Solodkoff and Richard Woodward (2017) question the need for the having/holding distinction. The having/holding distinction was invoked by van Inwagen to

explain why Hermione could be an abstract object and yet be clever even though abstract objects cannot be clever. Thus, the abstract object does not *have* the property but *holds* it. Consider the following:

10 (a) Hermione is clever.

(b) Hermione holds the property of being clever.

(c) According to *Harry Potter*, Hermione is clever.

Solodkoff and Woodward argue that both 10 (a) and 10 (b) are analyzable by 10 (c)<sup>12</sup>, unlike van Inwagen, who thinks that holding is primitive. One reason why van Inwagen thinks the distinction is primitive is because he finds analyses of holding inadequate (this can be seen in the problems with analyses of ascription). One of the analyses that we saw above that van Inwagen rejected was the one which said that the properties that Hermione holds are the properties which an actual person like Hermione would have or, if the story were told as fact, the properties of Hermione in that possible world. However, Solodkoff and Woodward note that this is merely an objection to a Lewisian analysis of truth in fiction (where “p” is true iff “p” expresses a truth if the story is told as fact) which is in turn an analysis of the fictional operator rather than an analysis of *holding*.

Solodkoff and Woodward point out that van Inwagen’s motivation for rejecting the operator analysis is that even though there is a three place ascription relation, it is not that Rowling would be ascribing cleverness to an abstract object. The reason for that is that she is not expressing a proposition, and as seen above, she is not referring to anything by using “Hermione”. There is thus no relation between the fictional work *Harry Potter* and a proposition <Hermione, being clever> to allow for the operator analysis.

The alternative proposal by Solodkoff and Woodward is that a work of fiction prescribes us to imagine that propositions like <Voldemort, being evil> are the case, which in turn is what makes the proposition expressed by “According to the fiction, Voldemort is evil” true. For this, they appeal to Walton’s (1990) idea that works of fiction “prescribe imaginings” that something is the case. We are authorized by the work *Harry Potter* to imagine that Voldemort is evil. This picture of fictionality causes problems for van Inwagen, as Solodkoff and Woodward point out: “The question that he faces is simple: if it is not fictional that Hermione is clever (because *Harry Potter* does not prescribe us to imagine that

---

<sup>12</sup> I would assume Solodkoff and Woodward as saying “According to the *Harry Potter* stories” where the stories are contained in the series by J.K. Rowling.

proposition) then what is true according to Harry Potter? Which proposition are we prescribed to imagine?" (*ibid.*, 412)

A possible answer, according to Solodkoff and Woodward, is Sarah Sawyer's (2002) which says that a work of fiction "prescribes metalinguistic imaginings", in that it prescribes us to imagine that "there is a true proposition expressed by the sentence "Hermione is clever" and that there is a clever person to whom the name "Hermione" refers." As per Solodkoff and Woodward, this is in line with van Inwagen for whom even though Dickens was not predicating anything of an abstract object while writing *The Pickwick Papers*, Dickens was "crafting a linguistic object that people pretend was a record of the doings of people called "Mrs. Bardell" and "Mr. Pickwick." (van Inwagen (1983) as quoted in Solodkoff and Woodward 2017: 412) The prescription then is not to imagine an actual proposition as such, but that even though there is no such proposition to be imagined, it is merely imagined that there is a proposition expressed by the sentence "Mrs. Gamp is female".

The problem with this view that Solodkoff and Woodward (*ibid.*, 413) point out is that it will lead to the idea that works of fiction in different languages, because they involve different sentences (even if they mean the same), will prescribe different imaginings for the sentences "Voldemort ist ein Zeuberer" in German and "Voldemort is a wizard" in English. The same would be the case for fictional names where the metafictional imagining cannot explain how sentences with different names of the same fictional character can be used to talk about (or refer to) the same character (how "Tom Riddle" and "Voldemort" can be used to talk about "he who must not be named").

The other objection that van Inwagen faces from Solodkoff and Woodward is on his contention that the distinction between *having* and *holding* is due to an ambiguity in "is": "what appears to be the apparatus of predication in fictional discourse is ambiguous. Sometimes, it expresses actual predication and sometimes it expresses a different relation [...] to which I give the name "holding." (van Inwagen 2003: 148 as quoted by Solodkoff and Woodward 2017: 420) Thus, apart from the "is" of identity, and that of predication, van Inwagen introduces an "is" of ascription.

However, one problem that Solodkoff and Woodward see is that the linguistic evidence does not suggest any such ambiguity. For example, due to two different senses of "is", the following can be considered ambiguous.

(11) The President is democratic and Obama.

But consider:



(12) Frodo is the ringbearer and protagonist of Tolkien's trilogy.

Van Inwagen would postulate an ambiguity with the "is" of ascription in the first part of the conjunct in However, unlike (11), (12) does not seem zeugmatic ("zeugma" is "a figure of speech in which a word applies to two others in different senses (e.g. John and his driving license expired last week" (Oxford English Dictionary, 2010))

Similarly,

(13) (a) Fiona cried out for help; Charles, his eyes out.

(b) Frodo is the protagonist; Samwise, his companion.

(13) (a) is zeugmatic due to different senses of "to cry" but there seems to be no such parallel with (13) (b).

Even if we think that there is a relevant difference between the uses of "is" which might not be exactly as effective as the other ambiguities, the following case resolves that.

(14) (a) Ross settled on the park but Ross didn't settle on the park.

(b) Bond is famous but Bond isn't famous.

The former does not necessarily involve a contradiction if we see two different senses of "to settle". But the latter does have a contradiction, which is possible only because there is only one sense of "is".

This argument further bolsters my arguments above against Goodman and Jung and Pellet. If there is no tenable distinction between the "is" of predication and that of ascription, and there is no distinction between everyday talk (intrafictional discourse) and theoretical discourse in fiction, then something like (13) (b) is completely consistent with my picture and we can still quantify over fictional objects in both standardly intrafictional and metafictional contexts.

## 7. Hayaki's Objection to van Inwagen

Reina Hayaki (2009) argues against van Inwagen by showing a tension between the following sentences (*ibid.*, 142).

(15) It is true according to the Canon that Sherlock Holmes is a detective.

(16) It is true according to the Canon that Queen Victoria hired a private consulting detective, gave him an emerald tiepin, and offered him a knighthood which he refused.

According to Hayaki, van Inwagen will have to see a disanalogy between the two sentences. In (16) what the Canon says about Queen Victoria is false, because the Canon predicates properties to Queen Victoria that she does not possess. As per someone like van Inwagen, for (15) however the reason for the falsity of what the Canon says would be different. The reason why Sherlock Holmes is not a detective is because an abstract object cannot be a detective. Of course, prefixed by the “it is true according to the Canon” operator, the sentence “Sherlock Holmes is a detective” is true, otherwise false.

Hayaki then rehearses Yagisawa’s objection to van Inwagen’s theory. The abstract object Holmes is not a detective. However, the Canon misattributes being a detective to an abstract object. Moreover, not only does the Canon do that, but it also attributes concreteness, personhood etc. to the abstract object. The objection is that while the Canon is merely “(intentionally) inaccurate” (*ibid.*, 142) about entities like Queen Victoria, London etc., that is, it is just a massive set of lies about abstract objects.

To this, Hayaki says that van Inwagen’s reply would merely be that the Canon does not attribute any property whatsoever to the abstract object but that there merely is the primitive ascription relation between properties, creatures of fiction and places.

The problem with this, according to Hayaki, arises when real objects or people like Napoleon are not allowed in the extension of A <vanity, Napoleon, War and Peace>. This is where the analogy problem between (15) and (16) arises. Van Inwagen would say that they are disanalogous. But on all counts, it seems that even if one allows only for fictional characters, “this does not touch on the intuition that if Sherlock Holmes is just as real as Queen Victoria, then the Canon says things about both of them equally.” (*ibid.*, 143) The Canon not only ascribes personhood to both (correctly in the case of Queen Victoria and incorrectly for Sherlock Holmes) “both (as objects) can appear in, or be (mis) described in, or have properties (inaccurately) ascribed to them in various works of fiction, either canonical or otherwise.” (*ibid.*, 143) Given the analogy between the two in these respects, why not just accept Yagisawa’s objection, that the Canon falsely “ascribes” (in the “ordinary sense”) to Sherlock Holmes?

As per Hayaki, Van Inwagen would still persist and reply that the Canon does not at all attribute or ascribe any property to the fictional object because of his two tier approach to fictional names: fictional names within the fiction do not refer to anything at all, whereas metafictional theoretical terms of literary criticism do.

This, according to Hayaki, leads to odd consequences about truth-in-fiction.

Firstly, Hayaki says that Van Inwagen, might reply that (15) and (16) are not true and the truth-in-fiction operator does not work. However, there are implied truths in a work of fiction, like “Sherlock Holmes lacks a third nostril”. The problem would be that it would seem odd that  $\langle \text{lack of a third nostril, Holmes, the Canon} \rangle$  belongs to the ascription relation A. The other problem is that if truth-in fiction operators are proscribed, then there is no way to account for the truth of evident truths in fiction like (16). This is especially because A-paraphrases (according to the fiction, p) are rejected by van Inwagen for real people and places. That is, if we reject the fiction operator, we cannot have statements like “according to the Canon, Queen Victoria is a person” (*ibid.*, 144) as true.

Secondly, suppose we allow for the use of the truth-in-fiction operator. The problem according to Hayaki is this. The name “Sherlock Holmes” will not refer in the original canonical text as well as the context of “it is true according to the Canon that.” (*ibid.*, 144) Thus, there can be “no existential generalization outside the scope of the operator” (*ibid.*, 144):

(15')  $\exists x (x = \text{Sherlock Holmes} \ \& \ \text{it is true according to the Canon that } x \text{ is a detective})$

If that is so, then (15) and (16) will be disanalogous because existential generalization is allowed for (16) (*ibid.*, 144):

(16')  $\exists x (x = \text{Queen Victoria} \ \& \ \text{it is true according to the Canon that } x \text{ hired a detective (etc.)})$

This might be acceptable but there is still another problem. Hayaki argues that if one does not allow the inference from (15) to (15'), then a major theoretical advantage that creationism or abstractionism can have would be lost, namely, that characters can be “exported” from inside to outside a fiction to be used in pastiches and parodies, as well as other fictions, which make all of them about the same object. In other words, the advantage that abstractionism has over anti-realist theories is that works of fiction like Conan Doyle’s canon introduce an existent fictional character which “Holmes” refers to and explains why subsequent references to Holmes are about the same character. If Conan Doyle’s canon does not allow for reference to Holmes, then the abstractionist cannot explain why subsequent works are about the same character as elegantly and either has no explanation or has to come up with a further, perhaps strained account.

However, my modification of van Inwagen’s account escapes all these objections by Hayaki.

First, my objection to Yagisawa's point that van Inwagen's view leads to the conclusion that the Canon is a massive set of lies about an abstract object. In my view, all fictional statements regarding characters like Holmes presuppose (MET<sub>FIC</sub>)\*.

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

The metaphor that is involved in the "likened relation" is an ontological metaphor (Lakoff and Johnson 1980), which are "ways of viewing events, activities, emotions, ideas, etc. as entities and substances" the basis for which are "our experiences with physical objects (especially our own bodies)" (*ibid.*, 25). For example, THE MIND IS AN ENTITY leads to the THE MIND IS A BRITTLE OBJECT metaphor, which is presupposed by statements like "the experience shattered him", "her ego is very fragile", "his mind snapped" and COMPANIES ARE PERSONS. In these metaphors, entities (even street corners like in "Let's meet at the street corner") are conceived in terms of concrete objects so that we can talk, quantify over, reason about them and in general make sense of events, ideas etc.

Now, in the way statements like "the experience shattered" him is true, so are statements about Holmes and other fictional characters. So contrary to Yagisawa's argument, there are no lies (about abstract objects in van Inwagen's case if he adopts this and about ideas in my case). The true statements in fiction and in literary criticism involving fictional characters are true because they involve the presupposition (MET<sub>FIC</sub>)\*.

Secondly, ordinary predication occurs of statements that presuppose these metaphors and there is no need to invoke any ascription relation with them, as ordinary predication suffices. So van Inwagen need not invoke ascription. Nor does he need to invoke a two tier theory of names, since all the statements involve the same name involved in the presupposition.

Thirdly, the "equal" ways in the Canon says true things, about both real people like Queen Victoria and fictional characters like Holmes and that statements (15) and (16) true in analogous ways, is that both are presupposed to be fictional ("originating in fiction"). The different ways in which they are false is because in the case of Queen Victoria, she does not possess the properties, and in the case of the abstract object (for van Inwagen but in my account, an idea), there is an ontological metaphor involved. Therefore, Holmes cannot literally be a detective. This is not implausible, since even in anti-realist theories like pretense theories, pretense would be invoked in the case of Holmes to explain why "Holmes is a detective" is literally false and thus, two different ways of determining falsehood will be present.

Finally, as mentioned, my account is a single tier theory where fictional names in intrafictional and metafictional contexts refer. Such operators become redundant for semantic purposes due to the presupposition and they appear only because there are presuppositions like (MET<sub>FIC</sub>)\* which contain predicates like “is fictional” and the clause “relevant intents and purposes” covers all fictional statements. This is because it is for the relevant intents and purposes of fiction or the larger Canon in which these statements like (15) and (16) are being made. Therefore, even without such operators doing any semantic work such as determining scope, one can account for the truth of statements like “Sherlock Holmes lacks a third nostril” because of the objectual metaphor involving a real person. Furthermore, for normal intents and purposes in narrative interpretation, ordinary assumptions about real people hold unless established by the narrative.

## Chapter 3

### Thomasson's Fictional Artefacts and Their Discontents

#### 1. A Brief Overview

Amie Thomasson espouses the artefactual theory of fiction, in which she holds that fictional characters are abstract artefacts such as money, works of literature and the like (Thomasson 1999, 2003a, 2003b, 2015). In what follows, I will focus particularly on her theory as she lays it out in Thomasson (1999) and explore some key objections raised to it. In the course of the dialectic, I shall look at some of my own discontent with the theory and attempt to modify parts of the theory in the light of all the discontent expressed here.

In Thomasson (1999), she begins her argument for the existence of fictional entities as abstract artefacts by asking what entities must be postulated to best make sense of our literary talk and practices. Once she postulates such entities, she argues that if one can accept a novel as an abstract object in one's ontology, then there is no reason to deny the existence of the fictional characters she postulates as abstract artefacts in the same ontology. In doing this, she moves away from approaches that try to fit in fictional objects in a pre-given ontology of possibilities, or abstract or non-existent objects.

In effect, while Thomasson asks the question about the kind of entities one must postulate to best make sense of our literary practices, she says that it makes sense that George Washington could not be a fan of Sherlock Holmes because he was not created before 1887. That is, Sherlock Holmes was created when Arthur Conan Doyle wrote about him, rather than discovered him from a range of eternal or non-existent or possible objects. Thus, in this respect, entities created (not discovered or selected) by mental and physical acts make best sense of our literary discourse and practices.

For Thomasson (1999), our intuitions concerning the origins of fictional characters also provide their identity conditions. Two similar characters would be identical only if they derived from works that have a common source and each author in the sequence is acquainted with the previous work, intending to import it into his. This also shows that the creation of a fictional character is by the particular "intentional act" of a single author or particular "intentional acts" of authors (*ibid.*, 35) who are part of a "story-telling tradition" (*ibid.*, 7) (like Odysseus might have been made by a number of authors collectively called

“Homer”). But, given that the creation is tied to specific acts in its origin, it is not the case that the origin can be at any time or by any author.

Moreover, in Thomasson’s account, characters can keep existing after they are created in literary works which can be preserved physically or in memory, as with oral traditions. As long as someone remembers the literary work, or a physical copy remains and that someone can interpret it (otherwise there will be mere marks on paper), a fictional character exists. Just in the way a character can be created, it can also be destroyed if the literary works and their interpreters are destroyed, unlike Platonistic abstracta, which cannot be destroyed. Thus, the best candidates to explain our practices are fictional characters created by an author or multiple authors through mental and linguistic acts, sustained in existence through memory, literary works and linguistic understanding. But unlike their creation where they are tied down to a particular author/authors who create them, fictional characters are sustained in existence if any work of literature in which they appear remains. Fictional characters cannot be tied down to a particular work in which they appear because that would not allow for them appearing in different editions or sequels. One would also be left thinking that each different edition or sequel would be about a different character.

Works of literature are, like fictional characters, rigidly tied to their origin in the particular author(s) who created them. The reason given by Thomasson for this is that identical sequences of words are not sufficient to make them the same work of literature. If *Animal Farm* was written in 1905, it would not be the same satire of Stalin that it was. The same sequences of words would also have different aesthetic and artistic properties. If Joyce had written *The Portrait of the Artist as a Young Man* after *Ulysses*, it would not have exhibited an original use of language.

The dependence of fictional characters on works of literature for their sustenance and their status as artefacts that can be created as well as destroyed is further seen in the fact that works of literature can be destroyed, unlike eternal Platonic abstracta. Our lament that stories and fictional worlds from great cultures like Greece have been destroyed as well as our efforts to preserve stories are evidence that literary works can be destroyed, just like other artefacts such as universities, umbrellas and the like. Furthermore, when literary works are destroyed then so are the fictional characters that are written of in them.

The picture thus given by Thomasson (1999) is that of fictional characters being dependent on illocutionary acts of an author or authors for their creation and on concrete entities like literary texts and comprehending humans for their continued existence. This reveals that

they are not part of a mysterious ontological realm separate from space and time but dependent on spatiotemporal entities and intentionality. The mystery of their creation is further dissolved once we find that other artefacts like contracts, money and marriage too are brought about into existence as a result of illocutionary acts in which they are represented as existing. In the case of a contract, for instance, it is brought into existence by the mere utterance of “I hereby promise to” (*ibid.*, 13).

### 1.1 Fictional Characters as Dependent Abstracta

Thomasson (1999) talks about the dependence of a fictional character on the creative act(s) of a particular author(s) as well as about its dependence on a literary work for its continued existence under the broad rubric of existential dependence. She defines existential dependence as:

Necessarily, if  $\alpha$  exists, then  $\beta$  exists. (*ibid.* 25)

Where  $\alpha$  depends on  $\beta$ ,  $\beta$  is a necessary condition for  $\alpha$  (*ibid.*, 25) and both  $\alpha$  and  $\beta$  are “individuals” and “individuals” covers “individual objects, properties, particular tropes and states of affairs.” (*ibid.*, 27)

To avoid the pitfall of the sort of counterfactual definition wherein a metaphysical relation is absent if  $\beta$  is a necessary existent, Thomasson says that the above schema is merely “the formal approximation of what is at bottom a metaphysical relation.” (*ibid.*, 25)

More pertinently, “the dependence of a fictional character on the intentional act(s)” of its author(s) for its creation is what she calls “rigid historical dependence”. (*ibid.*, 35) Its dependence on a literary work for its continued existence is “generic constant dependence”. A dependence is rigid when the dependence is on a specific individual as opposed to when a dependence is generic when it is on some or other individual of a given kind. There is historical dependence when for an entity to exist, its cause exists, but it can exist independently of its cause only *after* it has come into existence. This is in contrast with constant dependence in which the existence of an entity requires the existence of another entity throughout the time it exists.

An example of rigid historical dependence is that Queen Elizabeth’s existence depends on the particular gametes of her parent, George VI and Elizabeth. An instance of generic constant dependence is that if the US exists, then a citizen of the US exists, but not any citizen in particular.



Given that fictional characters are rigidly historically dependent on intentional acts of authors, it should be no surprise then that they are artefacts, since artefacts too are dependent on intentional acts. The consequence of fictional characters being generically constantly dependent on a literary work is that literary characters are dependent on whatever literary works depend upon. This is because, according to Thomasson, dependence is transitive (For instance, my being historically dependent on my mother who is historically dependent on her mother makes me historically dependent on my grandmother.) Thus, fictional characters are rigidly historically dependent on authors because literary works are also dependent on authors in the same way. Moreover, the generic constant dependence of a fictional character on some or other work that is about it makes it dependent on a copy and the existence of at least one person capable of comprehending it as well as possess relevant background knowledge to interpret it or a memory of it. This explains why, if all memories of a literary work are destroyed, or if the last copy is destroyed and the last individual capable of comprehending the particular work about the fictional character is dead, then the fictional character too is destroyed.

This picture also provides justification for the abstract and non-spatiotemporal nature of fictional characters. Thomasson contends that fictional characters cannot be spatiotemporal entities like other artefacts (for example, chairs and tables), since anyone who thought Sherlock Holmes could be found on 221B Baker Street would go to the address in London but not find him there. One might still argue that fictional characters are located in a copy or a memory of a literary work that is about it, but a fictional character survives despite the destruction of the copy or memory if there are other copies/memories extant. Furthermore, when we say a fictional character is “in” a literary work, what we do not literally mean that it is there “in” a spatiotemporal copy of it. The justification for the above picture, given by Thomasson, is the following argument<sup>13</sup> (*ibid.*, 37):

---

<sup>13</sup> Thomasson says:

If fictional characters were rigidly constantly dependent on some single spatiotemporal entity, we might have reason to locate them where that foundation is. Because they are not constantly dependent on any particular spatiotemporal entity, there is no reason to associate them with the spatiotemporal location of any of their supporting entities. (*ibid.*, 37)

To be fair, she does write “in the absence of a more plausible candidate” in the next line. I will try to bring out the plausible candidate. Even then, her not being able to find a candidate can be explained in terms of the fallacy I point out:

If fictional characters cannot be located either where they are said to be according to the story, or where copies of the literary works are located, in the absence of more plausible candidates it seems best to treat fictional characters simply as entities that lack a spatiotemporal location.

(P1) If there was rigid constant dependence of a fictional character on a particular spatiotemporal entity, there might have been reason for fictional characters to be located at their spatiotemporal foundation.

(P2) But there is no rigid constant dependence of a fictional character on a particular spatiotemporal entity.

(C) There is no reason for a fictional character to be located at their spatiotemporal location.

The problem I find with this argument is that it commits the fallacy of negating the antecedent. There might be no rigid constant dependence for a fictional character on a particular entity and yet there might be reason for fictional characters to be located at their spatiotemporal foundation. I will explore this reason later.

Thomasson's picture of dependencies of fictional characters also allows for a modal metaphysics which explains why a sentence like "If Arthur Conan Doyle had been busier with his medical practice, Sherlock Holmes might never have existed" (*ibid.*, 39) is true, something which other theories of fictional characters cannot account for.

The truth of the sentence is accounted for by a modal metaphysics in which, given the rigid historical dependence of a fictional character on its author, if Shakespeare did not exist in a world, then Hamlet would not exist in that world either. Furthermore, in any world in which Hamlet exists, so does Shakespeare and his act of creation at which time Hamlet begins to exist. A fictional character's generic constant dependence on a literary work implies that in any world in which a literary work about it exists (even if it is a different work), the fictional character exists, so long as it can be traced back to the creative act of its author. Thus, in a world in which Doyle would have been busier with his medical practice, he would not have committed the creative act that led to Holmes' creation. This accounts for the truth of the sentence above.

However, for anti-realists, Doyle's being bored of his medical practice does not really explain the creation of Holmes, nor the truth of the sentence, because the status of Holmes is the same: Holmes does not exist now, and if Doyle had been busier, he would not have existed earlier either. The possibilist and Meinongian are also at a disadvantage because Doyle's medical practice would have made no difference to whether Holmes existed,

---

Indeed this is just what we do in practice: Sophisticated readers treat fictional characters as lacking any spatiotemporal location, and thus as abstract in that sense. Thus fictional characters may, in brief, be characterized as a certain kind of abstract artifact. (*ibid.*, 37)

because Holmes was always there (as a concrete object in a possible world for the possibilist and as a non-existent object for the Meinongian) and will continue to do so. Holmes is merely discovered or picked out but not created in these accounts.

Additionally, Thomasson's theory can account for the truth of the sentence "There might have been other fictional characters than those that actually are." (*ibid.*, 40) Doyle might have created another fictional character in another possible world. On the other hand, the ontological distinction between possible and actual fictions is one that possibilism and Meinongianism do not make.

A further advantage of Thomasson's view is that of allowing for other contingent abstracta (as against views that think of all abstracta as necessary beings existing in all possible worlds) like musical entities and literary works which analogously exist only in worlds in which they are created by a particular author or particular authors.

## **1.2 Reference to Fictional Characters**

Thomasson accommodates her theory within the causal-historical account of names. (Kripke 1972/1980, Donnellan 1970) Typically, causal historical theories have problems with names referring to anything that cannot be causally related to the name, that is, anything non-spatiotemporal. To counter this, Thomasson adds that the name for a fictional character is related to a text that provides a "foundation" for a referent, given that the referent is ontologically dependent on it.

Her argument stems from a disagreement with Kripke (1972/1980) on fictional names in *Naming and Necessity*. Kripke's basic idea is that names refer only if there is a causal relation between the naming a person (in an act of baptism) and the use of a name via a chain of historical usage. Thus, even if we discover a detective who does all the things that are attributed to Holmes in the stories, the name would still not refer, because it just happens to be a coincidence that the person happened to do all the things that are attributed of the fictional character. The reason for this is that Doyle did not baptise anyone with that name. This is what Kripke calls his epistemological thesis. (*ibid.*, 156-157)

A fictional name also does not refer to a possible person who "*would have been* Sherlock Holmes, had he existed" (Kripke 1972/180: 158), because possible individuals doing Holmes-like actions cannot be uniquely picked out as Sherlock Holmes. Thus fictional names do not refer to any possible individual. This is Kripke's metaphysical thesis.

Thomasson questions this by saying that just because fictional names do not refer to a real person or to a person in a possible world, this does not entail that fictional names do not refer at all. She says:

To arrive at that conclusion one would have to unjustifiedly assume (1) that fictional works merely provide us with descriptions; (2) that, because there is no spatiotemporal thing to be baptized there can be no baptism process, and hence no rigid reference (of the sort names require); and (3) that if fictional names refer, they must refer to some actual or possible person.” (ibid., 47)

However, (3) is undercut by the reference not to an actual or possible person, but to an actual object.

Thomasson further goes on to undercut the key assumption, that is (2) by her own account of the causal link.

Although there can be no direct pointing at a fictional character on the other side of the room, the textual foundation of the character serves as the means whereby a quasi-indexical reference to the character can be made by means of which that very fictional object can be baptized by author or readers. Something counting as a baptismal ceremony can be performed by means of writing the words of the text or it can be merely recorded in the text, or (if the character is named later, for example by readers), it can remain unrecorded in the text.” (*ibid.*, 47)

For instance, in George Eliot’s *Silas Marner*, the name “Silas Marner” is introduced in the beginning with descriptions. The introduction of the name is the baptism (or at least “a public record of it”), like someone saying, “the character founded on these very words is to be called “Silas Marner.” (*ibid.*, 48) There can be a baptism at any time while writing or revising the text, even when the name is being changed in the final draft. That the baptism is “recorded...in the text corresponds to the requirement that the naming ceremony has to be somehow public.” (*ibid.*, 48) It needs to be noted that a character can also be named by the author before beginning to write about it. The chain of reference is maintained through the transmission of texts, akin to the chains in a speech community for regular names.

Thomasson’s case for there being a causal historical chain for fictional names is made stronger when parallels between them and regular names are seen while people communicate as a result of the chain. As a result of the difference between producers (those who are “acquainted with the individual names) and consumers (those who are not) (as made by Evans 1982), there can be people who hardly know about the character or are wrong about attributed properties but can still use the names. For instance, “Scrooge” is used by people proficiently even though most have not read *The Christmas Carol*.

The second parallel is with reference shifts. These happen for regular names when people mistakenly shift their intention to a different object from the one previously named. In fiction, for instance, this shift happened when in Mary Shelly’s novel, “Frankenstein” was the name of the scientist, but it later became the name of the monster. Moreover, just in the manner in which reference shifts for regular names happens when usage is distanced from its original use, in this case too, shifts can happen when consumers are distanced from the practices of readers. Thus, fictional shifts happen from a “fictional character to an imaginary character”.

Thomasson also adduces reasons outside of fiction that not only allow for causal chains but also “chains of dependence” for the transmission of reference. One is that people do in fact

refer to abstract entities with spatiotemporal objects as mediators. Thomasson's example is of a person pointing at a Studebaker and saying, "Now that's a great car". Even though the car may be damaged, in Thomasson's case, he is referring to a general type. Similarly, one can point at a painting like Corregio's *Jupiter* and say, "That's Jupiter", referring to the mythical god.

The second reason is that Thomasson's account of chains of dependence does not restrict causal historical chains to a specific ontology, like one only allowing for concrete particulars. One can refer to musical works, car types, or laws too using this framework.

### 1.3 Identity Conditions for Fictional Characters

One Quinean objection against fictional objects might be that they cannot provide identity conditions. However, Thomasson attempts to provide identity conditions for them so that one can infer that they exist.

Thomasson first attempts to show the inadequacy of Meinongian identity conditions. Parson's (1980) theory's identity conditions are that if "there is a unique object cor-related with every set of (nuclear) properties, so that fictional objects  $x$  and  $y$  are identical if and only if  $x$  and  $y$  have exactly the same nuclear properties." (*ibid.*, 56) Zalta's (1983) theory's identity conditions is that  $x=y$  iff they encode the same properties.

Thomasson's objection is that it is intuitive to think of the Pamela of Richardson's Pamela and Fielding's Pamela to be the same since the latter refers to the former and parodies her. However, it is not intuitive to think that in 1957, Fred Jones who is not aware of either of the works coincidentally writes about a Pamela who is very similar to Richardson's Pamela (she has the same name and a squire attempts to seduce her), writes about a character identical to Richardson's Pamela.

The problem thus with Meinongian identity conditions is that they will predict that Jones' Pamela is the same character as Richardson's and Fielding's, whereas intuitively, she is not. The problem with allowing for coincidentally identical characters is a side effect of the Meinongian propensity to ignore the origin of characters. Even in the case of intratextual identity conditions they get the wrong prediction because even if there is a slight variation in the same text in which a change to the description of a character is made, even a typing error, it will lead to more than one character in the same text, even though they are supposed to be identical.

Another view that Thomasson objects to is Wolterstorff's (1980) idea of a fictional character as a person kind. A fictional character is a kind that is a maximal component in a story or across stories. A kind A incorporates kind B if it cannot be kind A without also being kind B, like the kind *Indian* incorporates *being human*. In case there is a kind in a work and no other kind in the work incorporates that kind, then the kind is maximal within the work. Thus, Pierre Bezukhov is a maximal kind within *War and Peace* which incorporates the properties *being human* and *being the son of Count Cyric Bezukhov*. (van Inwagen 2003) A result of this is that Gogol's character Chichikov is a person kind Chichikov and anyone belongs to the kind if they have all the properties attributed to Chichikov in the work. The idea of the maximal kind within a story allows Marlowe's Faust and Goethe's Faust to be different. Yet, one can still write about the same character. Thus, there are properties essential to the kind that allow for identification of the character. For instance, there is also a character simpliciter Faust (a person kind) which is there in both Goethe's and Marlowe's Faust. It has core properties incorporated within the kind and is called "the Faust character" which then develops in both works.

Thomasson's objection is that there is a problem identifying these "core properties". The Faust character strikes a pact with the devil and is called "Faust". But one may have written a book in which he does not make such a pact is just a quiet scholar. He might also not be called "Faust" but "Phaust" in another play similar to Faust.

The other suggestion is by Maria Reicher (1995) who suggests that the total Holmes character is identified with the Holmes that encodes all the properties attributed in all the literary works written about Holmes and a particular character in each novel is a logical part of his total character. Thomasson's objection to this is that this conception would create problems for identifying a literary work as a member of the same sequence as other works. Reicher's problem is that her criterion cannot say why Hamlet, Tom Sawyer and Holmes are not part of the same total character and therefore would be unable to say that the three novels in which they appear are not part of the same sequence. One condition for saying that they are part of the same sequence might be that they are about the same total character. But that would be circular (Why are they about the same character? Because they are part of the same series.) If, however, one says that the works be about similar characters, the problem is that a sequel might have a character who has changed a lot since the previous work or that there can be similar characters in works that are not sequels.

In the same manner as in the case of reference, for Thomasson, the identity of fictional characters is tracked by the tracking of its foundations, that is, the concrete texts. For intratextual identity, the sufficient conditions for Thomasson are:

1. x and y appear in the same literary work
2. x and y are ascribed exactly the same properties in the literary work.  
(ibid., 63)

The second condition helps keep characters like Rosencrantz and Guildenstern different. As for the first condition, once one identifies a particular literary work, they figure out that they are the same character by means of 2.

Thomasson's necessary condition for intertextual identity is: (where L is the second sequel, K is the preceding work in the series, x and y are fictional characters in K and L respectively)

The author of L must be completely acquainted with x of K and intend to import x into L as y. (ibid., 67)

In the condition, complete acquaintance is when an author is a proficient user of x's name in K.

Thomasson acknowledges a problem in drawing out sufficient conditions for intertextual identity. The reason why the above condition is not a sufficient condition is because a reference shift might occur. An author might be completely acquainted with x, might use the name of x and yet create a different character by the same name. For instance, he might name a rock or dog "Sherlock Holmes" in the new story.

#### **1.4 The Problem of Negative Existentials**

The problem of negative existentials arises when we want to deny the existence of something. However, in order to do so, we say of it that it does not exist. It then becomes self-contradictory to say that it does not exist, for how can something (of which we say something) not exist?

One objection to Thomasson would be that since the characters according to her theory exist, the statement "Sherlock Holmes does not exist" will turn out to be false. However, this statement is true. Thus, Thomasson would not be able to account for its truth.

Nevertheless, she accounts for the truth of such statements like "there is no Lear" and "there are no unicorns" by making them equivalent to "there is no (real) person who is Lear" and



there are no (real) animals that are unicorns.” (*ibid.*, 112) They are equivalent, since we often make statements like “there’s no such person (he’s only fictional) and “there are no such animals (they are only mythical figures)” (*ibid.*, 112). A further reason is the presence of restricted quantification in our everyday communication. Thus, when we say that there is no such thing as Lear, it is not to say of an unrestricted domain of the world that Lear does not exist at all, but merely the domain of (real) persons.

Quite like van Inwagen, Thomasson distinguishes between fictional contexts and real contexts. Fictional contexts are contexts which are internal to the fiction that describe what happens in a story. These are contexts in which we often use pretense, the presence of which renders phrases like “according to the story” redundant. These contexts also include our descriptions of fictional goings on. However, fictional contexts do not exclude talk of real objects and events as depicted in the work of fiction. In real contexts, we take a perspective external to fiction, in the case of fiction, from the viewpoint of a literary critic. Real contexts also include predications about real objects, events and places, but not as how they are depicted in the fiction. The burden on a theory of fiction is to explain these two contexts and not the other way round, which is to accommodate fictional discourse to the theory itself. Pretense views like Walton’s and Meinongians unfortunately do the latter. Thomasson points out that Walton’s pretense theory, which is anti-realist, is guilty of producing ad hoc paraphrases for statements about fictional characters in real contexts (which will be elaborated on in Chapter 5). She also points out that Meinongians are guilty of introducing ad hoc distinctions in their theories to accommodate their ontology.

Instead of ad hoc paraphrases and ad hoc distinctions, Thomasson’s theory offers the advantage of uniformly literal (“genuine”) predications about real objects in real contexts. There is thus no difference between treatments of “Jimmy Carter is a person”, “Notre Dame is a cathedral”, “Hamlet is a fictional character” (*ibid.*, 106) and other metafictional statements as given by van Inwagen. These are “genuine predications applied to the object referred to.” (Thomasson 1999: 106) Statements in fictional contexts made about fictional characters are not true but when prefixed by the operator “according to the fiction”, they describe what is true in the story. Thomasson says: “What is true according to the story is, roughly, a combination of what is explicitly said in the story and what is suggested by the background knowledge and assumptions on which the story relies.” (*ibid.*, 107)

This account is also superior to van Inwagen’s, in that it does not encounter Hayaki’s criticism of not being able to import fictional entities into fictional contexts from real ones or from other works like parodies. Contradictory truths regarding the same character in

different works of fiction are also accounted for, since there can be contradictory truths regarding the character according to the fiction.

## **2. Zemach's Objections**

### **2.1 Zemach's First Objection**

In his first objection to Thomasson, Eddy Zemach (2003) argues that the reason why names in fiction cannot refer to fictional characters is that a story or play is about flesh and blood people, like Hamlet the Prince. However, in Thomasson's account, fictional names refer to abstract objects, characters which can be flat (two dimensional simple characters in fiction) or rounded (complex characters who undergo significant change). This would have the adverse effect of making stories about abstractions rather than people, which is not the case.

Thomasson's accommodation of this problem is in the distinction she makes between real and fictional contexts. This allows for literal readings of statements about Hamlet in real contexts, which are about abstract objects. That allows for statements like "Hamlet is a round character". On the other hand, the reason why fictional names in fictional contexts can be about flesh and flood people is because intrafictional statements are not to be taken literally due to the fictional operator. This is why Hamlet is not literally a Prince, but only one according to the fiction.

In my opinion, Zemach's reply to this is a very strong one. Zemach says that we often may not know whether we are reading is fiction or non-fiction. A biography might be embellished and the Bible slides between fiction and non-fiction. This makes the distinction between fiction and non-fiction more tenuous. Yet, we do understand the meanings of sentences in these texts when we read them, without a determination of the piece being fictional or non-fictional. "Yet for Thomasson that decision is momentous: whether a book is deemed fiction or nonfiction determines the meaning of every sentence in it. If a book's classification is revised no sentence in it will retain its meaning. Giving such power to librarians is unprecedented-and ill advised." (*ibid.*, 429)

However, Thomasson would have a reply that someone reading the fictional parts of the Bible might not really understand what a fictional name in a particular context means and merely think they understand it to mean a real person whereas it actually means a fictional character. In other words, they are open to correction once it is pointed out to them that what they're thinking about is not a real person but fictional. For instance, if they've taken a character from a biography to be actual but someone corrects them, saying, "Oh you missed

out on the embellishment by the author. It was a clearly a fictional context meant for embellishment. The author clearly *meant* it to be fictional.” In the case of the Bible, if they find out that their belief was false, they would correct it by saying that I thought "Adam" meant a real person but it was actually fictional. What accounts for this intuition that they incorrectly understand is that what determines meaning is that it is in fact a fictional context, and it is a fictional context because there was no baptism of an actual person while introducing the name, but only an abstract fictional object (which was also created).

But my reply to this possible response by Thomasson would be the following. Suppose that Conan Doyle was in a delusional state while writing Sherlock Holmes stories, thinking that he was telling known facts. This is not an utterly implausible situation, since Coleridge wrote “Kubla Khan” while he was intoxicated by opium. Now suppose Doyle died the moment he finished writing it, in the delusional state. People who found the extant manuscript thought it was about reality and the sentences were true. Or perhaps they thought it was fiction. In this case, there are no producers who know that it is fiction who can determine its meaning. Moreover, whether it will be seen as fiction or non-fiction or the account of a delusion depends on the people who classify it later after having understood it. Suppose Doyle leaves clues before dying that it is an account of a delusion. It is thus entirely contingent that someone may classify this as fiction or non-fiction or as the account of a delusion or even both. Suppose as a publisher, I want to sell it as fiction. I classify it as fiction and then someone later comes and finds out that it is an account of a delusion. It might pique readers’ interest to know that, and psychiatrists might even study it for that reason, but most people will still read the stories as fiction. To the psychiatrists, it is primarily an account of delusion. And yet, everyone is aware of the meaning of the manuscript, which is a precondition for classifying it. All of them are about a flesh and blood person who did such and such things. Whether it is fictional or non-fictional is irrelevant to understanding it and therefore to its semantics, and all that is required is a grasp of English. We, like Zemach says, do not need a librarian to determine its meaning.

This makes the distinction between fictional contexts and real contexts, insofar as they determine the meaning of sentences at the semantic level even more problematic. Perhaps we need to look at reasons other than the fictional context/real context difference as to why fictional names are about flesh and blood human beings rather than about abstract objects despite referring to abstract objects. In my chapter on van Inwagen, I attempted to show how the distinction between metafictional and fictional statements (parallel to real and fictional contexts) is not tenable, at least for ontological purposes. As I proceed, I will try to show further as to why it is not semantically tenable.

There is also another lesson to be drawn from this. Suppose that after Doyle dies, his sentences are construed to be fiction. Since there is no difference in the nature of the semantic object, its fictionality or that there is pretense nowhere figures in an account of meaning or even in the creation of the semantic object (assuming realism). What role then does pretense play here? It would be helpful to see what the difference is between the delusional context and the fictional context. The difference is merely that whereas in the former, there is belief (by Doyle) that Sherlock Holmes is a real person, in the case of the latter there is no such belief. Thus the role of pretense seems to be limited, in the second case to not believing that Holmes is a real person. This distinction will come of use later, as we shall see.

## **2.2 Zemach's Second Objection**

Zemach's second objection is that despite Thomasson's attempt, the sentence with a fictional operator will say of an abstract object that it did whatever it did. The paraphrase would be equivalent to "According to the fiction, an abstract object named "Claudius" killed his brother. But what we understand when we read fiction is that a Danish man named "Claudius" who killed his brother, not an abstract object who did it. This is as absurd as saying that the number 27 killed his brother. We also get the absurd situation, where "Brutus" is the name of a real person and (suppose) "Lucius" is that of an abstract entity, it seems absurd to think that the play asks us to "imagine that a human being (Brutus) carries on a conversation with an abstract entity." (*ibid.*, 430)

Thomasson's reply would be that Zemach misses her point that one can in fact pretend of the abstract object that it is a person, since the operator just signifies pretense. However, I have already pointed out in the Deluded Doyle case that this is untenable for the semantics and consequent aboutness with respect to Holmes being flesh and blood.

On the other hand, I have a proposal to solve this kind of objection, which has been recurring in the literature (Friend (2007), Sainsbury (2009), Sawyer (2002)). While there is nothing intrinsic to the content of an abstract object that in itself would make it about a flesh and blood person, an imaginative act that sets up a comparison between the abstract object (the same would apply to my account where reference is to a unified mental representation and we don't think of a mental representation as flesh and blood, but let's just go with abstract objects for now) and a person would. This is on the lines of a corporation being understood as a person for all intents and purposes as well as the sort of objectual metaphor like THE MIND IS A BRITTLE OBJECT presupposed in statements like "He underwent a

mental breakdown” or “I was shattered” (Lakoff and Johnson 1980). If we recall, in my section on van Inwagen, I analysed the statement “James Tiberius Kirk is a fictional individual” as (let it be Sherlock Holmes in this case)”

(MET<sub>FIC</sub>) Sherlock Holmes is fictional but is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

From this we can infer:

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

But given that I have construed fictionality as extraneous to meaning in fictional and non-fictional contexts, in both contexts, the presupposition behind each sentence about Sherlock Holmes in Doyle’s narrative would be:

(PRE) Sherlock Holmes is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

Note that this is a presupposition. However, it would seem very strange that this presupposition holds even in the delusional state. For one does not know that there is a comparison. However, one can argue that the comparison is part of a cognitive mechanism but not really known by the speaker. There are good reasons for this.

One reason would be the presence of an intentional stance and a design stance (postulated by Dennett 1987) in both cases. The design stance is the heuristic strategy of assuming that an object is designed for a purpose and that it will act according to its purpose. A concrete example would be: If I push a few buttons in an alarm clock, it will ring at a certain time. In an abstract case, I can predict the behaviour of a computer by analysing its source code. The physical stance is to use the method of the physical sciences where one’s predictions are based on the laws of physics and the knowledge of the physical constitution of things.

Dennett postulates the intentional stance (a subset of the design stance) as an innate heuristic strategy. He defines it as “the strategy of prediction and explanation that attributes beliefs, desires, and other “intentional” states to systems- living and non-living- and predicts future behaviour from what it would be rational for an agent to do.” (Dennett 1988: 1) In the intentional stance, an object or artefact is treated as if it is a rational agent. (Dennett 2014)

An example for the usefulness of the intentional stance would be that of playing a chess game with a computer. If we merely use the design stance (which is still more simple than the physical stance), we will have to analyse the source code to predict its behaviour, which will involve innumerable steps and will take more than one's lifetime. So, for simplicity, I have to assume that it:

1. "knows" the rules and "knows how" to play chess,
2. "wants" to win, and
3. will "see" these possibilities and opportunities for what they are, and act accordingly (that is, rationally). (Dennett, 2014: 77)

The best way to predict its moves would then be:

- a) First, list the legal moves available to the computer when its turn to play comes up (usually there will be several-dozen candidates).
- b) Now rank the legal moves from best (wisest, most rational) to worst (stupidest, most self-defeating).
- c) Finally, make your prediction: the computer will make the best move. (Dennett 2014: 81)

Similarly, we also use this folk psychological stance to predict ourselves and other human beings. Like the computer, we treat each other as intentional systems ("Anything that is usefully and voluminously predictable from the intentional stance is, by definition, an intentional system." (*ibid.* : 78)) This is so, because not only is it difficult and well nigh impossible to predict one's behaviour by figuring out the laws of nature and the physical constitution, but it is also impossible to know what is going on in someone's head. Thus, in new situations, it makes it easy to make sense of what one believes (say that if one's eyes are open, and they see a bus in front of them, they will believe it is a bus) and desires in a particular scenario.

What is more pertinent for our case of fictional characters is Dennett's description of what the intentional stance does. He says: "One attributes to the system the beliefs and desires it ought to have, *given its place in the world and its purpose*, and then predicts that it will act to further its goals in the light of its beliefs." (Dennett 1988: 496) The italics are important here, because they reveal what is to be presupposed by the intentional stance in the case of fictional characters. In the case of a computer, it is the software as if it is a rational agent. But it does not end there. The pertinent respects in which it will be treated as a rational agent also have to be determined, which are given by (1.), (2.) and (3.) above, which in turn are made pertinent "given its place in the world and its purpose".

Analogously, “given its place in the world and its purpose”, that of being a fictional detective who has adventures, it has to be presupposed of an abstract object (a mental representation in my case) that it is likened to a flesh and blood person, in order to even *attribute the right kind (that is, the most relevant) of beliefs and desires, the sort we find attributed in fiction*. Also, to predict the right sort of beliefs and desires, one cannot but think of embodied agents acting in situations which elicit the sort of beliefs and desires embodied agents have. Thus, in order to even interpret sentences in the deluded Doyle case, along with the intentional stance, whether it is fictional or not, (PRE) has to be presupposed.

(PRE) Sherlock Holmes is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

Thus we do not have to say of an abstract object that it committed such and such actions, qua abstract object or that it interacted with Brutus since further facts about representation have been accounted for. The abstract object literally did not do such things because of a comparison presupposed in (PRE). Unlike Thomasson, there is no appeal to the “according to the fiction” operator and the predication in the scope of the operator, which does not lead to predicating of it that it is such and such. Since I adopt Anthony Everett and Timothy Schroeder’s theory that fictional characters are ideas constituted by mental representations and not abstract objects, the same “likened to” relation would apply to a mental representation, since a mental representation too is not literally a flesh and blood person and does not do and cannot be the things typically attributed to Holmes in the fiction. Specifically, the mental representations that constitute ideas are likened to a real persons.

My account has a further advantage. Suppose Deluded Doyle’s sentences about Sherlock Holmes are construed to be fiction by posterity. The relevant purpose and the design is thus of a fictional object. In that case, like I mentioned in the case of van Inwagen, both intrafictional sentences and metafictional sentences will presuppose:

(MET<sub>FIC</sub>) Sherlock Holmes is fictional but is like a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

This hold for all relevant intents and purposes, even in statements like “Sherlock Holmes is not a real person”, “Sherlock Holmes was created by Doyle” etc. This would be elaborated upon below.

The advantage my account will have will be to explain why there can also be aboutness with respect to the abstract object (for Thomasson’s account, but mental representations for my

own account<sup>14</sup>) where we can call Holmes being a rounded character or a flat character because we have primarily taken the design stance rather than the intentional stance being primary (if x is a fictional character, it can be rounded or flat because it was designed that way.) This will also have to presuppose its property of flesh-and-bloodedness, since in a way, a rounded character is significant to us only because we value the understanding of complexity in persons, so that we can predict their desires and beliefs. For instance, in literary criticism, we might wonder what Anna Karenina would have done had she not had the affair with Vronsky or we are surprised by a character's actions and we say that such and such action was surprising because we were predicting their beliefs and desires. Thus, relevant comparisons of abstract objects (or mental representations) with real persons need to be made before there can even be a pronouncement whether a character is rounded or not. In other words, the results of the design stance, that is, the pronouncements of literary criticism will presuppose the results of the intentional stance.

If this looks like a problem consider this. Dennett's intentional stance is based on the hypersensitive agency detection device, selected for by natural selection to attribute agency even when it is not present. The anthropologist Stewart Guthrie's (1993, 2008) hypothesis is that given that our minds bet in perception as to the most likely source of our stimuli, and that we bet on the most salient possibility that is important for survival. So, for instance, it is better to bet that the source of the rustling sound or eye like glowing in an undergrowth belong to an agent, given that we stand less to lose than if they are now. Given that organisms, particularly other humans exert a broad and powerful influence on us, they are the most salient source of our stimuli. Also, humans can be tough to detect because they might deceive or camouflage, which increases uncertainty as to the source of our stimuli, we bet that they are the most likely source of stimuli. This mechanism is proposed as an innate mechanism that is the source of anthropomorphism and therefore animism, because of "our need to discover any agents in an uncertain environment" (this is universal, and he marshals a lot of other evidence too, like we have a "low threshold" for the attribution of agency, given that there is a tendency to see even a group of dots in motion as an agent).

Further on, McCauley (2011), argues that this origin of religion, in our hypersensitive agency detection device that applies theory of mind to cognition is part of "maturationally natural cognition", which is reasoning that is more automatic and unconscious, prone to

---

<sup>14</sup> There is a further nagging worry about how despite all of what I've written on the intentional stance and ontological metaphor, I'm still thinking of a mental representation rather than a flesh and blood person when I think of a fictional character. The worry will be addressed in full in Chapter 6 even as I separate facts about reference and facts on aboutness.



illusions rather than “practiced” cognition, which is reasoning which is part of our learning. Gopnik (2010) further traces the origin of fictional characters to imaginary companions in childhood which is grounded in theory of mind.

Therefore, given that Dennett’s intentional stance is a terminological version of “theory of mind” and he bases the intentional stance on the hypersensitive agency detection device, he too grounds the stance in explaining the origin of religion and the attribution of beliefs and desires to gods. If that is so, and Gopnik’s hypothesis is right, what he says below about gods applies to fictional characters too:

Language gave us the power to remind ourselves of things not currently present to our senses, to dwell on topics that would otherwise be elusive, and this brought into focus a virtual world of imagination, populated by the agents that mattered the most to us, both the living but absent and the dead who were gone but not forgotten. Released from the corrective pressure of further actual encounters in the real world, these virtual agents were free to evolve in our minds to amplify our yearnings or our dreads. (Dennett 2006: 114)

### **2.3 Zemach’s Third Objection**

Zemach’s third objection is that given that in fiction “Fa” is not that a is F, but according to the fiction a is F, then even universal truths like “man is noble in reason” proffered by literature should have different meanings in fiction as they do not have the same truth conditions. As Zemach (2003: 430) says: “in fiction, no sentence means what it normally does; it has entirely different truth conditions. How can I identify with a novel’s hero or heroine if they cannot possibly make any statement I make?” Given that the truth condition is not that “man is noble in reason” iff man is noble in reason but it is true iff according to the fiction, man is noble in reason. A universal truth like “man is noble in reason” would then be rendered “trivially true and entirely banal (*ibid.*, 430) since it is merely true because the fiction says so and not true because the world is a particular way. As Zemach says, while authors intend that we take these statements as true simpliciter with respect to the world, “Thomasson’s semantics makes that illegitimate.” (*ibid.*, 430)

Thomasson might reply that if a statement p is prefixed by the fictional operator “according to the fiction S”, and the statement “according to the fiction S, p” is true, then it does not imply that the statement p in and of itself necessarily has a different truth condition within the scope of the operator. The function (of the operator) is just that because it helps to say what is true according to the fiction, the truth of fictional statements like “Sherlock Holmes lives on 221B Baker Street” with respect to the fiction can be accounted for even though the statement is false simpliciter. It also does not show that a statement p within the scope of the

operator is necessarily false simpliciter with respect to the world (but merely that it might be false) or possesses a different truth condition: it merely shows that a statement that is false simpliciter can be true according to the fiction. Thus, to prefix any statement with the fictional operator is merely to signify that the fiction says such and such, not that the statement is necessarily false or necessarily has a different truth condition. The truth condition for such a statement is “According to the fiction S, p” is true iff the fiction S says (or depicts) that p. Despite being prefixed by the operator, “man is noble in reason” is true for independent reasons, namely that man is noble in reason and the fiction merely depicts the proposition in the fictional context

But this would lead to an inconsistency. She says:

Often some pretense is involved in such discussions [in fictional contexts], as we pretend that what the story says is true; for example, that it describes not mere fictional characters but instead actual people whom we could praise, criticize, or psychoanalyze. Such pretense renders the use of such phrases as "according to the story" unnecessary, as this is simply understood. I refer to these as "fictional contexts." (Thomasson 1999: 105)

There can be two ways of interpreting Thomasson here. The charitable interpretation would be that the regular recourse to pretense that characterizes a work of fiction makes the use of “according to the story” operators redundant. But this is very awkward. One can have a partly fictionalized biography with the recourse to pretense not being regular. The other interpretation would be to say that the operator only signifies an implicit pretense in only cases of statements involving pretense, not those that are true. But this becomes arbitrary, since the fictional operator was rendered redundant because of the presence of pretense. And the reason why it is redundant is because the fictionality signified by the “according to the fiction” operator entails pretense. So it becomes arbitrary as to why the operator signifies pretense in only statements involving pretense and why it should not in statements that are true simpliciter. Furthermore, given that the fictionality signified by the operator implies pretense, it will involve a contradiction with the proposed reply since, statements that are true outside fiction and are simultaneously true according to the fiction will both involve a pretense (given the fictionality) and not involve a pretense (as per stipulation that the operator signifies pretense only for statements involving pretense).

My account has a better solution. In my account, statements in fiction only presuppose propositions like (MET<sub>FIC</sub>) (in which the comparison to the relevant kind can be specified) and the intentional stance, and not the fictional operator. There can be assertions in fiction only if the presupposition is made. The quantifiers in such presuppositions are not restricted

and the “is fictional” predicate is added only in the context of this sort of unrestricted quantification. Thus the default position in fiction is that one is in fact making statements about the world. Since for “man is noble in reason” the truth condition is merely, “man is noble in reason” iff man is noble in reason, there is no problem that my account faces. This looks mysterious, but appealing to Matravers (2014), I would add that fictionality does not enter at the level of semantics but is linked to the lack of a fidelity constraint (where we assume that what is included in the narrative is constrained and governed by the narrator’s belief that the events have occurred), which explains why it figures as a predicate in (MET<sub>FIC</sub>)<sup>15</sup>. Statements appearing in fiction are to be evaluated in the light of the lack of the fidelity constraint. If this is so, then a reader of fiction would not be quick to judge false statements as false, but the truth conditions for statements in fiction can in fact be still be intact in terms of them pertaining to the world.

### 1. Friend’s Objections

Like Zemach, Stacie Friend (2000) has an objection to Thomasson’s distinction between fictional (internal) and real (external) contexts. While Thomasson would say that statements regarding fictional characters in external contexts need to be taken literally, Friend points out that statements like “I pity Anna Karenina” are not literally true, since I do not really pity an abstract object (this applies to fictional characters as ideas too). It can also not be taken as an internal statement, since it is not true that according to the fiction, that I pity Anna Karenina.

Thomasson might respond, using Reicher (2006) by revising her strategy so as to allow for the story operator to be a predicate modifier rather than a sentence operator. Thus, sentences like ‘Pegasus is a flying horse’ is true iff Pegasus is according to the story, a flying horse. This account has the advantage of explaining why an abstract object cannot literally have a property like being a horse (unlike Thomasson’s account, as we saw) because Pegasus is according to the fiction a flying horse. This account has internal sentences saying things from the point of view of the story as to what is happening in the fiction and external sentences about the world. However, external sentences come in two varieties. One are “quasi-internal sentences” that cannot be literally true and at times describe what is happening in the story, like “Pegasus is according to the fiction a horse”. The other are sentences like “Pegasus is a character from Greek Mythology” which are literally true.

---

<sup>15</sup> I elaborate on this in the section on Matravers in the chapter on Walton’s Pretense Theory (Chapter 5).

Thus, a possible response to Friend's objection by Thomasson might be as follows. "I pity Anna Karenina" does not need to be taken literally despite being true and as a full-fledged external sentence because it is paraphrasable as the "quasi-internal" sentence "I pity Anna Karenina who according to the fiction is a person". Thus, the distinction between external and internal contexts is preserved.

My objection to this account is as follows. If we take Anna Karenina outside the scope of the fictional operator, the sentence "I pity Anna Karenina who according to the fiction is a person" would entail "Anna Karenina" is the name of a fictional character" and "'Anna Karenina" refers to a fictional character". Given this entailment, one still lands up with the result that the direct content of one's pity is an abstract object, and not a flesh and blood person, which it should be. One reply to this might be that I am pretending of an abstract object that it is a person and the predicate modifier merely reveals the pretense. Moreover, it is *de re* pretense and not *de dicto*, so the abstract object is not the content of my thought (that is, it is not stating that an abstract object is the content of my thought, but a flesh and blood person. This shall be made clear in the next paragraph).

However, the objection that Friend raised is that if "I pity Anna Karenina" is not literally true, then we cannot make the distinction between internal and external contexts. But is the paraphrase literally true? In "I pity Anna Karenina who according to the fiction is a person", one cannot literally use the anaphora "who" for an abstract object, since "who" cannot be used literally for anything inanimate. Thus, taken literally, the paraphrase is false. Another paraphrase might be "I pity Anna Karenina and according to the fiction, she is a person". Here, even though "she" is in the scope of the operator, to refer successfully to the same kind of object, it should be "it". One reply might be "I pity Anna Karenina that according to the fiction is a person". However, this seems artificial and therefore merely an ad hoc move to accommodate one's ontology. Moreover, the "that" clause makes the pretense *de dicto*, since, "that" reveals that the content of my pity is an abstract object and I am aware of it.

Thus, Thomasson's distinction between internal and external contexts is given by her is untenable. My uniform account in which  $(MET)_{FIC}$  and the intentional stance are sufficient to explain why "I pity Anna Karenina" is true and that it is about a flesh and blood person.

Given the weakening of Thomasson's distinction of internal/external contexts, how would I account for contradictions like the round square? Thomasson handles these cases by means of predication prefixed by the "according to the fiction operator" in internal contexts. My account might be better able to account for this because there is no such contradictory object

but merely a bundle of mental representations that encode mind-dependent properties like *being a thought of being round* and *being a thought of being square*, in which the properties encoded in the mental representations are likened to that of round objects and square objects and the bundle is likened to an object.

Friend's other objection is as follows. Thomasson argues that if theories can allow for the existence of dependent abstracta like literary works but reject fictional characters on grounds of parsimony, then they are not really being parsimonious. Literary works and fictional characters can be of the same kind, given that if we postulated them, they would be abstract artefacts that are "cultural creations". They are both created by intentionality, dependent on copies of literary works to maintain existence. Hence if one accepts the former, it would be arbitrary to eliminate the latter, in the same way that we cannot reject baseball games in our ontology but allow board games. "False parsimony occurs not only if some objects in a particular category are excluded and others included, but also if a single category is arbitrarily singled out for exclusion but others with the same relevant categories are retained." (Thomasson, 1999: 144)

Friend's objection is that if this is so, then one also has to accept false posits like Phlogiston and mythical creatures like Zeus, given that it would be false parsimony to reject them as they depend on theories and myths and Thomasson puts theories and myths in the same category as literary works. But there are no such entities and that does not give reason enough to accept them as abstract artefacts. "Failure to pick out a concrete object is not success in creating an abstract one." (*ibid.*, 1000)

However, again invoking Reicher (2006: 160), Thomasson might reply that to deny that mythical objects or false posits exist concretely is not the same as denying that they are abstract objects. To conflate the two is a category mistake. What we do when we deny that Vulcan exists is to deny that it is a concrete object. Not that it is not an abstract object.

However, the problem is deeper. Friend also says that "failure to pick out a concrete object is not success in creating an abstract one." That is a further objection and I shall attempt to flesh it out in my own terms. Either all posits are abstract objects or they are not. If all posits are abstract objects, then if I posited a planet that exists in my hypothesis (like Mercury) then that refers to an abstract object. However, when I confirm my hypothesis and I discover Mercury, I do not discover that the abstract object that "Mercury" refers to is identical to the planet. Neither is there a reference shift. When I look back, I will think that my name "Mercury" referred to the planet Mercury which I later discovered. One might argue that I

was merely looking for a planet that satisfies the properties of my posit. Fair enough. But when I was looking for the referent of “Mercury”, under this model of the abstract object, I was not looking for the abstract object but for the referent of “Mercury”. This is why there was no reference shift.

Now suppose all posits are not abstract objects and only failed posits are. Again, it is arbitrary to say that in the case of Mercury there was no abstract object created but in the case of Vulcan, it was. I was in no way disappointed because there was no planet Vulcan identical to my abstract object. One might say that only when I failed to pick out Vulcan that I created an abstract object and “Vulcan” referred to the future creation all along. But this is very awkward. When “Mercury” was posited, it referred to a (an epistemically) possible planet which in fact turned out to be an actual planet and it is arbitrary to say that whereas “Mercury” referred to a possible planet, “Vulcan” did not refer to a possible planet but to an abstract object. On the other hand, the artefactualist might reply that given that “Vulcan” referred to a possible planet, it seems odd to say that it referred to nothing at all.

I think that the artefactualist’s worry in the previous line is a genuine one, but that artefactualism itself will have a tough time saving itself here. Nevertheless, I can still have the realist cake and eat it too if I say that “Vulcan” refers to an idea (the way Everett and Schroeder (2015) say that a perpetual motion machine is merely an idea) and so does Mercury” before it is discovered. When I discover Mercury, I find out that the properties of the planet Mercury matches the idea (more specifically it matches the mind-dependent properties of mental representations). Thus, I say not that the idea is identical to Mercury but that Mercury matches my idea. There is reference shift here that has a plausible explanation, because unlike saying that I was looking for an abstract object like in the previous case, I can merely say that I was looking for an object that would match my idea, and once I have found it, the reference shifts to it. Nevertheless, I can still refer to my idea of Mercury as well as Mercury now. In the case of “Vulcan”, I’ll always end up referring to an idea of Vulcan, since there’s nothing that matches my idea. Saying all of this is not counterintuitive. We keep saying things like “”What I saw lived up (or matched up) to my idea of it”, or that “What I saw did not live up (or matched up) to my idea of it.” Moreover, how do I account for the phenomenology of looking for the planet Mercury, when I have an idea of it and which seems to be the referent of “Mercury”? Again, I invoke something like (PRE) but in this case:

(HYP) Mercury is likened to a real planet and is and ought to be treated/counted as such, for all relevant intents and purposes.

The relevant intents and purposes are for the sake of the hypothesis and testing it.

In the case of “Vulcan”, there is nothing that matches the properties postulated. I still presuppose something like (HYP), but with Vulcan instead of Mercury. Thus, if fictional objects, like posited objects, are ideas, then I can change artefactualism to Everett and Schroeder’s theory and save it from Friend’s objection.

## **2. Iacona and Voltolini’s Objections**

In this section, I shall examine Andrea Iacona and Alberto Voltolini’s (2002) objections to Thomasson on her position on the ontological dependence of fictional characters which I find to be incomplete. I will then use that as a launchpad to take up my complete argument against Thomasson.

### **4.1 Iacona and Voltolini’s First Objection**

Iacona and Voltolini object to the rigid historical dependence of a fictional object’s origin on an intentional act(s) by an author(s). They argue that if it was another mental act by the same author, then on Thomasson’s account, it would have been a different fictional character. However, intuitively, it seems that another mental act which replaced the actual mental act with which the author created the fictional character could have just as well have created the same character. Thus, the dependency would have to be generic and not rigid. Iacona and Voltolini allow for rigid dependence on the author but not on the mental act.

But there is a further problem with even accepting that there is rigid dependence on the author. If it is another mental act of the same type, there is no reason why the same type of mental act cannot be tokened in another author, thereby not allowing for rigid dependence at all on the author. However, the reason why Iacona and Voltolini do not want to reject rigid dependence on the author is the Menard case (from Borges). Imagine that at the same time as Cervantes, Pierre Menard was in another house, and unbeknownst to Cervantes, Menard wrote “Don Quixote” and it happened to be identical word for word with Cervantes’ “Don Quixote”. One argument is, since they are different literary works because of possibly differing aesthetic properties, the characters too are different. The other argument is that intuitively, they are different. The intuition would be based on a case as was given above, by Thomasson, that if Fred Jones, centuries after Richardson’s Pamela was written, writes about a character remarkably similar to Pamela, it would not be the same character.

Here is a counterexample. Imagine that centuries later a man (say Friar Smith) finds a copy of Menard's *Quixote* as well as Cervantes' *Quixote*, without knowing the authorship. He identifies both as the same character and writes a sequel, with Quixote as the protagonist. The question then is, is it the same character as Cervantes' Quixote or Menard's Quixote? The problem with saying that it is the same as either one of them is that it would be arbitrary to do so. I will set the question aside as to whether this would make them the same work or not, but my case is that semantically they are the same work. Suppose Friar Smith does not make a sequel but copies half of Cervantes' copy and half of Menard's copy, thinking that it is the same work. Would the latter work be identical to Cervantes' copy or Menard's copy? It would be arbitrary to say it is identical to either of them. Perhaps this is not a warranted conclusion, because Cervantes' and Menard's Quixote might have had different aesthetic features or that Menard's personal history would add aspects to the interpretation of his Quixote, that would not be available for Cervantes' and vice versa. However, they might not be the same work, but they are semantically the same, since Friar Smith can write his sequel with Don Quixote as his protagonist, which is identical to both Menard's and Cervantes' Quixote. He can do this only because he understands the words by virtue of his knowledge of the language. Thus, both the works are semantically identical, even though they are not pragmatically identical or not identical with respect to aesthetic features.

However, Thomasson may object, what about Fred Jones' Pamela's case? Wouldn't that be a case of rigid historical dependence on the author which determines that Richardson's and Fielding's Pamela are identical? Perhaps Friar Smith is mistaken and it is merely a case of vague identity (vague identity being very controversial and a big bullet to bite), that there is no fact of the matter as to whether Friar Smith's character is identical to Menard's or Cervantes' Quixote, but only vaguely identical. Notice, this move is being made by the person who holds on to rigid historical dependence. There might be other reasons for vague identity, but in this case it is a consequence of holding on to rigid historical dependence. Or perhaps it is the rules which are not exact. As Friend (2007) says, "Thomasson ('Fictional Characters') says that we should not expect answers to these questions, because the existence and identity conditions for fictitious objects are determined by practices whose informal rules are vague and imprecise." (*ibid.*, 149) However, there might be another independent reason to undermine rigid historical dependence. If that holds good, then only vague identity or interest relative identity can be accepted, but not rigid historical dependence for the origin of a character.

Consider the computer game *No Man's Sky*. When a programmer programs the game, all the characters are not encoded in the program. No character is determined. As the game



proceeds, new characters are created based on previous situations encountered. The game encompasses the whole universe, so one can find bizarre ones on various planets which match the habitat of the planet. It is extremely unlikely that two players can even encounter the same character. The question then is, did the programmer create each character?

Answers to this can vary. If you point at a character and ask the programmer, did you create this? The response might be yes, but indirectly. Or it is an accidental creation. Or that it was generated by him. Or the programmer might say that he did not create it but the program did. Or the algorithm did. Intuitions vary. Moreover, if you ask, could another algorithm have created it, the response is yes. Any Turing equivalent algorithm could have created it. What it has to do is merely generate a character as a combination of attributes, that is, colour, size, number of legs etc. Another response available to the programmer is that it was created by an evolutionary process. Natural selection was simulated where characters develop as an answer to evolutionary pressures. A bigger planet with higher gravity would lead to bigger animals. There will be different selection pressures. Thus, simulated environmental pressures created the characters.

One observation to be made is that intuitions vary as to whether a particular algorithm rigidly created the character and that the programmer created it. Both intuitions are undermined by the point about selection pressure: it was not the algorithm or the programmer, but the selection pressure that created it. From this, it is difficult to even generalize that all fictional characters depend upon their authors, let alone rigidly. Furthermore, the point about Turing equivalence is analogous to my point above that if a token of the same type of mental act can be instantiated in another author, then she can also create the same character. If creation is contingent in this manner, where a character is dependent on an author but it is still not a sufficient condition for rigidity, then Thomasson cannot revise her thesis on rigid historical dependence to “if a character depends on an author, then it depends rigidly on the author.” The possibility of rigid historical dependence on an author for the creation of a character is thus seriously undermined by the case of *No Man's Sky*.

Perhaps the reply by Thomasson might be that the clear intuition in the Menard case can settle this. But there is an error theory about why the intuition is so strong. One says that Einstein's Theory of Relativity could have been given by someone else, and it is thus not rigidly historically dependent on Einstein. The analogy with *No Man's Sky* is pertinent. Like the context of science where theories can be easily translated without loss of meaning because it is impersonal, *No Man's Sky* too is an impersonal medium. Thus even if the

computer language is changed, there is translation of the level of Turing Equivalence. The medium is an impersonal one. However, in the case of literature, which is a personal medium, the reason why the intuition that there is rigid historical dependence is so strong is that the piece of literature cannot be translated without loss of its literary and aesthetic qualities. More pertinently, it is because the literary and aesthetic qualities include the aesthetic qualities produced by biographical details of the author and an appreciation of his or her originality. This is precisely why intuitions vary in the *No Man's Sky* case but not that much in the Menard case. Therefore, it is likely that the intuition that there is rigid historical dependence in the Menard case is not because there is actual rigid historical dependence but because of the heresy of paraphrase (which occurs due to the loss of aesthetic qualities). This has nothing to do with the nature of the characters or the semantics of fiction but due to an extraneous reason due to the more personalized nature of the medium of literature/fiction and the confusion of aesthetic qualities with semantical properties of fictional characters<sup>16</sup>. Therefore, since the intuition on rigid historical dependence stems from another source rather than actual rigid dependence, and given the *No Man's Sky* case as well as the Friar Smith case, there is no rigid historical dependence of fictional characters on their authors.

This brings us back to the Fred Jones' Pamela objection. My rejection of rigid historical dependence account predicts that Menard's and Cervantes' Quixote can be identical in the Friar Smith context, but Fred Jones' Pamela is still not identical to Fielding and Richardson's. Rigid historical dependence on the other hand, would have ensured that characters sufficiently similar, yet created by different authors would be not identical, to account for the intuitions in the Fred Jones' case. If that is so, then we have to accept that the characters are not identical in the Menard case. However, I do not see why this should be necessarily so. If identity in this context is merely interest relative (also see Lamarque 2010), based on certain rules of genre and other literary conventions, then to look at rigid historical dependence, especially given my arguments against it is merely superfluous.

It is also to be observed here that my arguments above still account for the truth of "If Arthur Conan Doyle had been busier with his medical practice, Sherlock Holmes might never have existed". Doyle's mental state did give in fact give rise to Holmes. But is consistent with the statement to say that that while Holmes might never have existed

---

<sup>16</sup> However, this extraneous reason might account for why it is legitimate to have interest relative identity and say Menard's Quixote and Cervantes' Quixote are counted as different characters in the initial Menard case. However, since the reason about the personalized nature of literature is extraneous, we can thus count both Menard's and Cervantes' Quixote the same in the Friar Smith context.

because it is less likely that someone else would create him, Holmes could still have existed if some other author came up with him, since Holmes is not rigidly dependent on Doyle.

#### **4.2 Iacona and Voltolini's Second Objection**

The second objection by Iacona and Voltolini is as follows. A fictional character as per Thomasson is generically constantly dependent on a literary work. But a literary work is an abstract object that is individuated semantically. That is, it is not just the symbols due to which a work is that particular work, since two works can have the same symbols but differ in meaning. It is symbols and their meanings. But if fictional names refer to fictional characters, then they partly constitute literary works. The question then is, how can a fictional character depend on something it itself constitutes?

One reply could be that since the work depends upon itself (dependence is reflexive) and a fictional character constitutes the work, the work depends upon the fictional character (dependence is transitive). But unless a temporal sequence of the creation of the literary work first being created and then the character is created is assumed (and I doubt if Thomasson assumes that) and that the dependence is asymmetric because of the causal relationship, there seems to be no reason why this relation cannot be symmetric (if both are created at once). If the work depends on the character so does the character on the work. This is what explains why, if a work is destroyed, the fictional character is simultaneously destroyed. Thus, Iacona and Voltolini have confused causation with ontological dependence.

There is a problem, however, with the relation's being symmetric as a result of both being created at once. The dependence shows that they are necessarily created at once. If a work is created, then so is a fictional character and vice versa. But there is no reason why I cannot create a character without intending to write a fictional work. I am bored, I feel like it. Or I am in a creative writing class where I am given a task to look at a picture, attribute characteristics that I imagine and name it. The creative writing teacher asks me, "Let's see what character have you created?" Another case might be that Arthur Conan Doyle thinks of various characters, writes them down as rough sketches in his diary (he does not even have to write them down), and then finally chooses Holmes. Thus when I say "Sherlock Holmes could have been a frog named Talbot Holmes", the statement is true iff Doyle created a frog character named Talbot Holmes amongst the many he chose from to be the protagonist in his novels and short stories. It might or might not have been a part of a literary work. Therefore both are not necessarily created at once.

This transfers into the constant dependence relation too. If the collected works of Sherlock Holmes are destroyed, but Doyle still remembers the Talbot Holmes or that the sketches are still there, then Talbot Holmes is not constantly dependent on the work. Or that the collected works can remain while Talbot Holmes is destroyed the moment Doyle forgets about it or the sketches are destroyed. Thus there is no generic constant dependence of a fictional character on a literary work.

### **3. Reicher's Objection**

Thomasson's competent readership condition for the continued existence of a fictional character is the clause that a fictional character is constantly dependent on at least one extant copy of a literary work and a competent readership or a memory of it. Maria Reicher (1999) objects to this as follows. Suppose the competent readership is all dead, that is, that there is no one left to interpret a particular work. In that case, in Thomasson's account, the condition is satisfied for the fictional character to cease to exist. Again, suppose that scientists reconstruct the ancient language and recover background knowledge. They thus become competent readers which ensures that the character has been resurrected. Thomasson however, holds that objects cannot "flit in and out of existence." (Thomasson 1999: 22) Thus, this case goes against a basic tenet of her account. This according to me is a definite counterexample.

### **4. Howell's Objections**

Robert Howell (2002) objects to Thomasson's idea of the baptism and quasi indexical reference to a fictional character via a text. The example given by Thomasson is that there is a quasi-indexical reference to the character dependent on the text when George Eliot utters "The character founded on these very words is to be called "Silas Marner" (Thomason 1999: 48) at the time of baptism. Not only is it unclear how the baptism works but also:

Speaking of 'these very words' sounds indexical. But is there some further indexicality involved in achieving reference to the character via its dependence relation to the words? And how does appeal to that relation achieve the reference? Do we fix the reference of 'Silas Marner' via the description "the character founded on these very words"? But Thomasson says only that it is "as if" we use that description (which in any case Eliot and others unfamiliar with Thomasson's theory cannot employ). So what do we really do? (ibid., 283)

I have already argued that there is neither a rigid historical dependence an author for fictional character, nor a generic constant dependence on a text. Given my worries about the two dependencies, I don't see why the baptism has to be done via the text as a

spatiotemporal foundation at all. The concern expressed by Howell becomes more pressing in that light.

### 6.1 The Problem with a Non-Spatiotemporal Fictional Character

The reason why Thomasson (1999) wants the spatiotemporal foundation is because she wants to undercut Kripke's point against the baptism of fictional characters that "because there is no spatiotemporal thing to be baptized there can be no baptism process, and hence no rigid reference (of the sort names require)." (*ibid.*, 47)

Another reason why she wants a spatiotemporal foundation is because a fictional character is non-spatiotemporal. The reason for this is that one cannot bump into a fictional character. The problem, however is with the assumption: if something is not a physical object, then it is not spatiotemporal. Therefore, the false dichotomy: either something is a physical object or it is non spatiotemporal. This leads to odd conclusions. Consider Descartes' dualism.

- a) The mind (or a mental event) is not a physical object.
- b) Anything that is not a physical object is nonspatiotemporal.
- c) Therefore, the mind is nonspatiotemporal.

One reason that might be given for this, like the Holmes case, is that one cannot bump into the mind. Therefore, it is non-spatiotemporal. But the problem is that the mind is presupposed when a person walks down the road and bumps into things. One cannot bump into one's mind only because it is a person with a mind who bumps into other objects as well as comes into contact with one's own body. At most, one can conclude from the fact that the mind is not a physical object is that the mind is not spatiotemporal in a particular way (Rosen, 2001) (this is compatible with any noneliminativist yet nondualist concept of mind). Similarly, if a fictional character is not a concrete object in the external world that one can bump into, then it is safe to give a likely explanation about why we do not bump into it; that this is so merely because it is not spatiotemporal in a certain way, since it is an idea in the mind (which is spatiotemporal). To say that it is non-spatiotemporal would be too strong a conclusion, especially since a more parsimonious and naturalistic conclusion is already available, like in the case of the mind. Moreover, contrary to Thomasson's account where fictional objects cannot "flit in and out of existence", Reicher's scientists' reconstruction case in fact shows that they can do so. Even if one accepts that concrete objects in the external world can flit in and out of existence, one cannot identify fictional characters with any of them. (if so, which one? A person doing things Holmes did? That

option was already eliminated by Kripke as we saw above). Since fictional characters are not concrete objects in the external world, and our thoughts do in fact flit in and out of existence, then ideas indeed are likely candidates for fictional characters.

### 5. Everett's Objection

Thomasson (2003b) looks to give an account of fictional discourse that tackles the following inconsistencies in fictional discourse (*ibid.*, 205):

- (1) Frankenstein's monster was created by Dr. Frankenstein.
- (2) Frankenstein's monster was created by Mary Shelly.
- (3) Sherlock Holmes is a detective.
- (4) Sherlock Holmes is a fictional character due to which one cannot call him to solve crimes.
- (5) Emma Woodhouse does not exist.
- (6) There are fictional characters such as Emma and her sister while there is no such fictional character as Emma's pesky kid brother.

If both statements in each pair ((1) and (2), (3) and (4), (5) and (6)) are taken literally, it would lead to straightforward contradictions. Thomasson, like Walton (1990), suggests that we need to take at least one statement in each pair non-literally to understand the speaker charitably.

Thomasson outlines four types of discourse about fiction (*ibid.*, 207):

- 1) Fictionalizing discourse: Discourse which is part of a fiction.
- 2) Internal Discourse: Discourse by readers about the fiction's content.
- 3) External Discourse: Literary critical discourse like (2) and (4)
- 4) Nonexistence Claims: Like "Sherlock Holmes does not exist."

The first two can be said to involve pretense with an implicit fictional operator. Thus, (1) and (2) as well as (3) and (4) can be consistent: (1) and (3) involve pretense, (3) and (4) do not.

Thomasson in (2003b) also adds another desideratum to her analysis of nonexistence claims. The analysis of a nonexistence claim not only must account for why (5) is true but also why "The fictional character Emma Woodhouse doesn't exist" is false. (*ibid.*, 214)

Thomasson's solution arises from the fact that nonexistence claims "presuppose and implicitly comment on a separate range of prior uses of the name by speakers with perhaps inappropriate, intentions to refer to a thing of a certain ontological kind." (*ibid.*, 217) Thus, when someone says "Santa is coming", and someone else corrects her by saying "Santa does not exist", they correct her mistaken intention to refer to a person.

Thus, Thomasson's (2003b) analysis of nonexistence claims is:

If N is a proper name that has been used in predicative statements with the intention to refer to some entity of an ontological kind K, then "N does not exist" is true iff the history of those uses does not meet the conditions for referring to an entity of kind K. (*ibid.*, 217)

In the case of "Moses does not exist", the statement is true given that the practice of the name's use traces back merely to a story and not the baptism of a real person.

This analysis accounts for why in a situation where *Hamlet* is being taught and the teacher said "Shakespeare modelled Hamlet on a 13<sup>th</sup> Century character of Sano Grammaticus" (*ibid.*, 218), denying Hamlet's existence would be absurd. The intention was by the teacher to refer to a fictional character and the reference is successful since the practice of the name refers back to the text. This analysis also accounts for (6) where Emma's pesky little brother, say "Fred", is denied existence since the intention was to refer to a fictional character. But the story does not have one. Thus, (5) and (6) are consistent in this account.

Anthony Everett (2007) objects to both of Thomasson's (1999 and 2003b) analyses of negative existentials. Thomasson's (1999) analysis, according to Everett is:

T1. A may use an utterance of " $\alpha$  does not exist" to convey the claim that that  $\alpha$  is not a real K, where (i) K is a conversationally salient kind and (ii)  $\alpha$  is fictionally characterized as being a K in a conversationally salient fiction. (*ibid.*, 65)

(where A is a speaker and K is a kind described by the fiction)

To recall, Thomasson (1999) had equated "Hamlet does not exist" with "there is no real person as Hamlet", where "person" is the value of K in this case. For a character such as Holmes, Thomasson's account would get the prediction right. However, Everett's counterexample is as follows. Consider the following (*ibid.*, 68):

(7) Yugo is a fictional character who has not been attributed any kind in the fiction. Moreover, it cannot be known by the reader whether Yugo is an animal or vegetable or of any kind at all. Yet, the statement "Yugo does not exist" is true.

Furthermore, Everett asks us to consider a statement in which the value of K is fixed to “thing”: “Fictional objects, though often discussed by literary critics, do not exist.” (*ibid.*, 68)

To account for this, Everett suggests that Thomasson might peg her analysis to:

T2 A may use an utterance of “ $\alpha$  does not exist” to convey the claim that  $\alpha$  is not a real thing. (*ibid.*, 69)

But in this case, Thomasson's prediction will be wrong, since that statement will turn out to be false. This is because Thomason's theory itself does say that every fictional character is a real thing, since it is an abstract artefact. Another revision by Thomasson that Everett proposes might be:

T3 A may use an utterance of “ $\alpha$  does not exist” to convey the claim that  $\alpha$  is a fictional object. (*ibid.*, 70)

But then the statement “Fictional objects do not exist” which is not trivially true will be trivially true, since then fictional objects are fictional objects.

### **7.1 Solodkoff's Modification of Thomasson**

Tatjana von Solodkoff's (2014) response is targeted at Everett's response to T2. Her argument is that since ‘real’ can be used in different ways in ordinary language, like something that is “existent; occurring in the physical world; not imaginary, fictitious, pretended or theoretical; actual” (*ibid.*, 340), there is no reason why ‘real’ in ‘real thing’ cannot be interpreted as concrete. “After all, we normally think of real things as being the kind of things you can meet, bump into and interact with in the physical world.” (*ibid.*, 340) Also, names tend to be associated with concrete objects. The denial that Holmes is not concrete is to remove such an association with the name “Holmes”. Alternatively, “Holmes does not exist” can be understood by the realist as “Holmes is a pretend object”, which is compatible with realism. Solodkoff argues that this does not diminish its status as an object, since “pretend” merely modifies “object” to “narrow down the class of objects being talked about” (*ibid.*, 341) and so a pretend object can still be an object (the modifier merely eliminates the possibility that it is a concrete object or any other kind of object). It is also not implausible to make the distinction between a pretend and a serious object, since Holmes comes into existence as a result of a pretend use of a name whereas Obama exists irrespective of any pretend use of a name. Thus, the value of K can be assigned to either



“serious object” or “concrete object”. (*ibid.*, 344) Consequently, for Solodkoff, the truth condition for a negative existential is (where  $K_c$  is a contextually specified kind in context  $c$ )

“ $n$  doesn’t exist” is true in a context  $c$  iff  $n$  isn’t a  $K_c$  (*ibid.* 338)

The advantage of a contextual specification is that the value of  $K_c$  is open, that is, it can account for Thomasson’s intuition about Holmes not being a real person where  $K_c$  is fixed to “person” in that context. Yugo cases can also be accounted for with “concrete object” as the value of  $K_c$ .

For general negative existentials like “Superheroes don’t exist” and “Hobbits don’t exist”, the truth condition she gives is:

‘ $F$ ’s don’t exist’ is true in  $c$  iff there is no  $x$ , such that  $x$  is a  $K_c$  and  $x$  is an  $F$  (*ibid.*, 346)

The truth condition for “Fictional characters don’t exist” for Solodkoff is:

“Fictional characters don’t exist” is true in  $c$  iff there is no  $x$  such that  $x$  is a  $K_c$  and  $x$  is a fictional character (*ibid.*, 347)

If the value of  $K_c$  is fixed as the kind “serious object”, then the right hand side will read as  $x$  *being not a serious object and a fictional character*. This analysis accounts for the truth of statements like “All fictional individuals are unreal”, “All fictional individuals are non-existent”, “No fictional individual exists” (*ibid.*, 346) even though the realist holds that fictional characters exist.

## 7.2 My Objection to Solodkoff

However, Solodkoff’s distinction between a pretend and serious “object” is untenable. One major reason is because our everyday use of the word “thing” is that of a concrete object and we use it in what Lakoff and Johnson (2003) call ontological metaphors by means of which we understand “events, activities, emotions, ideas etc.”.

Understanding our experiences in terms of objects and substances allows us to pick out parts of our experience and treat them as discrete entities or substances of a uniform kind. Once we can identify our experiences as entities or substances, we can refer to them, categorize them, group them, and quantify them—and, by this means, reason about them. (Lakoff and Johnson 2003: 25)

If that is so, then every way in which we talk about Yugo using the word “object”, whether serious or pretend, will likely be pervaded by the ontological metaphor and the burden will be on Solodkoff to explain why it is not one. In fact, the worry is, that the ontological

metaphor pervades even our talk of abstract objects. Perhaps it is not inescapably so, but the burden is on Solodkoff to show how she knows that what she is talking about isn't pervaded by ontological metaphor and is not parasitic on our concept of a concrete object, since it might as well be. This worry extends to when the value of K is "concrete object", since if we claim that Holmes isn't a concrete object, the artefactualist implies that Holmes is an abstract "object". We also saw this sort of mistake that had been made by Jeffrey Goodman (2004) in Chapter 3, when he took statements like "James Tiberius Kirk is a fictional individual" literally and from which he concluded that therefore "There are fictional individuals". What he missed out was the evident presupposition behind this statement, namely that "James Tiberius Kirk is a fictional person". The reason why the presupposition held was because literary critics, unlike philosophers who go on a holiday with language, and ordinary people ordinarily use the word "individual" in this context with the normal meaning "a single human being as distinct from a group" (Oxford Dictionary of English (2010)). So the application of the predicate "is a fictional individual" is not literal and is a category mistake, especially since it violates (CA).

(CA) A predicate that correctly ascribes a property that can belong only to a concrete individual can be literally predicated only of a concrete individual.

Similarly, with the predicate "is a pretend object", Solodkoff cannot really rule out that she is not violating (CA), especially when she's making the usage parasitic on the ordinary usage of "object" as a concrete object.

Allow me to elaborate on ontological metaphors. The use of ontological metaphors becomes clearer in the case of things which are not "discrete or bounded", like "mountains, street corners, hedges, etc." where we "impose artificial boundaries" to understand them as discrete, we thus understand them as "entities bounded by a surface". For instance, the ontological metaphor INFLATION IS AN OBJECT, is systematically presupposed by the following (*ibid.*, 26):

*Inflation is lowering our standard of living.*

We need to *combat inflation*.

*Inflation is backing us into a corner.*

*Inflation is taking its toll at the checkout counter and the gas pump.*

Buying land is the best way of *dealing with inflation*.

In the examples above, we use the metaphor so we can “refer to it [inflation], quantify it, identify a particular aspect of it, see it as a cause, act with respect to it, and perhaps even believe that we understand it.” (*ibid.*, 26)

We tend not to notice that these are metaphors, even in statements which are true, like “He cracked under pressure”, which presuppose the metaphor THE MIND IS AN ENTITY which is presupposed by THE MIND IS A BRITTLE OBJECT, which again has systematic instances. More examples of the THE MIND IS A BRITTLE OBJECT metaphor given by Lakoff and Johnson are as follows (*ibid.*, 28):

Her ego is very *fragile*.

You have to *handle him with care* since his wife’s death.

He *broke* under cross-examination.

She is *easily crushed*.

The experience *shattered* him.

His mind *snapped*.

Furthermore, we also use container metaphors in ontological metaphors, which assume that like physical things, what we are talking about has an inside and an outside and other stuff can be put inside them or taken out. For instance, when I say I get *into* the water, I use WATER IS A CONTAINER SUBSTANCE as a metaphor. Or we see a state like Kansas as a container, when we say “There is a lot of land in Kansas.” To put a clearer example, even when we understand mundane sentences like “There is a butterfly in the garden”:

We have to conceptualize the boundaries of the garden as a three-dimensional container with an interior that extends into the air. We also have to locate the butterfly as a figure (or trajector) relative to that conceptual container, which serves as a ground (or landmark). (Lakoff and Johnson 1999: 33)

The problem that Solodkoff faces is to show how when we use phrases like “pretend object”, we do not use “object” in a metaphorical sense. The reason for this is that even though we might say “Yugo is a pretend object”, we will inevitably conceptualize Yugo to have an inside, an outside and a boundary like the human body and in general, as a container. The burden is on Solodkoff to show how that is not inevitable and there is

additionally, a non-metaphorical sense in which we can apply the predicate “is a pretend object” to Yugo.

One reply available to Solodkoff is that “object” in this sense really means “entity”, which is a philosophical term of art and which can be used to refer to properties, events, states of affairs, kinds etc. In fact she does say that “pretend” modifies “object” in the way “fake” modifies “barn” such that the way a fake barn is not a barn, a pretend object isn’t a concrete object. However, as Lakoff and Johnson observe, our talk of even events, kinds etc. is infected with our everyday notions of boundedness and discreteness as well as container metaphors. Apart from the examples with water and the garden there are also examples with events like the metaphor RACE IS A CONTAINER OBJECT in sentences like “Are you *in* the race on Sunday?” and “He’s *out* of the race now.” (Lakoff and Johnson 2003: 31) Moreover, this is rather revisionary. We use object in the same sense at least insofar as we talk in terms of discreteness and boundedness and with an inside and outside as modified with serious as with pretend. It’s just that we use “pretend” to modify “object”. Why should the sense change, since there is no explicit ambiguity?

A more promising line of thought might be to say that even though we use such ontological metaphors to talk about God, that we can use such analogical language does not entail that God does not exist. Even with water, that we understand it in this way is merely a matter of convenience, and it does negate the fact that when we are asking about the nature of water, we really find out the nature of water. Moreover, we realize that the reason why we recognize the metaphor as a metaphor is because we know that water is not literally bounded like everyday objects. Nevertheless, since we use “object” in the same sense in serious and pretend object, the phrase “pretend object” is rather unfortunate, because we merely think of it as an object in the bounded sense, unaware of the metaphor implicit in it. A neater solution would be to presuppose (MET<sub>FIC</sub>)\*\* (which is a generalization of MET<sub>FIC</sub>) when we talk about Yugo as a pretended object and in fictional statements about Yugo, even though we do not know what kind Yugo is.

(MET<sub>FIC</sub>)\*\* There is an x such that x is pretended/fictional and x is like a real object and x is and ought to be treated/counted as a real object for all relevant intents and purposes.

The explicit comparison in the presupposition explains the implicit metaphor whenever “object” is used in this case as well as the implicit metaphor in “pretend object”. It also explains why Yugo is conceptualized as discrete, bounded with an inner/outer distinction.

I think Solodkoff's analysis is right at the level of explanation that it intends to have, where "object" is used as bounded and discrete and with an inner and outside, given the serious/pretense object distinction. As I said, the right hand side of her analysis will only be true because it presupposed (MET<sub>FIC</sub>)\*\*.

## 6. Van Inwagen's Objection

Van Inwagen (2003) questions whether it is even metaphysically possible for abstract objects to be created. One might say that it might be possible to create (even for God) an abstract object like a set, since if A and B (concrete objects) are created, then if sets exist, then the set {A, {A,B}} is also created (the set depends on A and B). If that is so, then analogically, if a manuscript is created, and if a fictional character is dependent for its existence on the former, then the author does create a fictional character. Van Inwagen however points out that not only are Thomasson's fictional characters unlike sets, but they are also so different from sets that one can hardly come across uncontroversial examples of such abstract objects that are created.

However, Thomasson's examples of such creation are marriages, laws, contracts and other social entities. Since they are created by acts of representation, there is no reason why a fictional character cannot also be created by acts of representation.

Van Inwagen on the other hand does not think it obvious that a new entity comes into existence when a marriage takes place. On the contrary, it can be construed as new properties being added to already existing objects (in the case of a promise, for instance, it is *having promised to teach Alice how to drive*) or relational properties (like *is married to*). (*ibid.*, 154) If this is so, then sentences like "some marriages are better than others" can be paraphrased with the quantifier only ranging over "people, properties relations and times." (*ibid.*, 154) Furthermore, the problem with even construing marriage as a new set involving a man and a woman which is an abstract object, is that there is no clear intuition that a marriage comes into existence. Thus, given the vague conditions of intuitions regarding this, van Inwagen concludes that "many cultural and institutional entities...can be brought into being represented as existing" (*ibid.* 155) is not even a philosophical datum.

Thomasson(2003a) argues for why a theory of fictional characters has to be consistent with a common conception of them and that the theory has to fit the data supplied by our ordinary beliefs and practices concerning fiction, unlike Meinongian and Platonist theories. In reply to the objection that since our commonsense notions can be wrong, especially in science (for example about the geocentric view), we should really look for the real truth rather than just

follow common sense, her first argument is as follows. Unlike empirical evidence for scientific views, there can be no such empirical evidence to settle a theory of fiction. For instance, there can be no empirical evidence to decide that they are not created but are necessarily exist in a Platonic realm or nonexistent. The only available data is that of common sense views “about the nature, existence or survival conditions of fictional characters.” (*ibid.*, 144)

Moreover, if we go by direct reference theories, the content of natural kind terms like “whale” is not determined by analyzing its concept and associated descriptions, but by discovering its nature by scientific investigation. This is because the term refers as a result of a “causal relation between those who ground the use of a term and a certain sample of entities.” (*ibid.*, 144) But in the case of fictional characters, there is no ostension and no such causal relation and thus there can be no scientific investigation to determine the nature of fictional entities.

As a result, in this matter, the only thing we can point at are words in copies of literary texts. But we need to disambiguate what it is that we are referring to (the qua problem). For instance, whether we are using the kind term “fictional character” to refer to the kind of ink or paper, or a genre kind, or even a kind of fictional character that the text represents. In order to do the disambiguation, the people who ground the use of the term need to know the rules that govern how the term “fictional character” is to be used, so that it is distinguished from other kind terms like “copy of text”). They also need to understand the ways in which fictional characters are connected to texts, to identify and reidentify them in and across texts and when one can refer to one. In order to do all this and “ground the reference for the kind term”, they must have a “substantive ontological concept of the sort of thing a fictional character is.” (*ibid.*, 145)

Thomasson draws an analogy with “inning”. In order to determine its reference, the “would-be grounder” needs to know what it is, identify the beginning and end of an inning, and possess necessary conditions for knowing when an inning takes place. All these amount to the grounder possessing a substantive concept of “inning”.

Thus, a possession of substantive concepts is to possess “existence, identity and survival conditions for members of a kind” (*ibid.*, 145). Having them is a precondition for reference and they also define the kind of referent involved. Thus, if concepts define the ontological kind, what the kind is does not need to be understood by recourse to discoveries. In order to

find out what fictional characters are, how they are to be individuated and whether they exist (their ontological status), we must then look to our ordinary beliefs and practices.

One argument that Thomasson mentions against her view is that we can be wrong about our beliefs that fictional characters are created just in the way that people who believed that alchemists created gold out of straw were wrong.

Thomasson's reply is that the existence conditions implicit in our concept of "fictional character" show how a fictional character is created. Thus, the condition for the creation of a fictional character is:

A fictional character is created iff the author is not referring to real people or a character in another story and describing the seeming actions of a named individual.

This amounts to the same as "pretending to refer to and assert things about a person, as part of an understood tradition of story-telling pretense." (*ibid.*, 148) Since that the sentence, "Jane Austen wrote a work of fiction pretending to refer to and describe a young woman named "Emma Woodhouse"" (*ibid.*, 148) is true, given the substantive concepts involved in "fictional character", it entails reference to the "fictional character "Emma Woodhouse"".

Thus, whereas no such conditions for creating gold are met in the case of alchemists, conditions that are sufficient are met in the case of fictional characters. One can only deny this if one believes that there is a conspiracy theory, saying that "the appearances are deceptive and no one in actuality writes a fictional story".

To deny that a fictional character exists, Thomasson argues, is also a "distortion of ordinary usage" (*ibid.*, 149). If one accepts that an author has written sentences in a story with a pretend reference to Emma, but there is no such fictional character violates common usage. This is akin to saying that one accepts that teams take turns to bat and bowl but there are no such thing such as an inning. The reason why the anti-realist seems to be denying it is that their stakes are too high to fulfill these conditions: either there must be a "practice-independent object existing in another mode of being and referred to" (*ibid.*, 149) or that there has to be a person who conforms to what the story says about the character and is causally related to the name. But this is to misunderstand the rules of how "fictional character" is used.

Thomasson's argument against the argument by van Inwagen is as follows. She puts up the following sentence for consideration (*ibid.*, 150-151):

(8) Jane Austen wrote a work of fiction pretending to refer to an Emma Woodhouse (and was not referring back to an extant person) and the fictional character Emma Woodhouse appears in a work of fiction.

She argues that anyone who denies that fictional characters exist will have to say that the second part of the conjunction does not have the same truth conditions or truth value as the first part of the conjunction. However, saying that the second part is false or has no truth value would be to reject the intuition that it is redundant to utter the second part of the conjunction, when used with the first part. It would also be to say that the inference from the first to the second conjunct is illegitimate, when it is not.

Thomasson's point, is thus that if one were to reject fictional characters, one would distort "the ordinary uses of terms such as "fictional character", severing the ordinarily permitted inferential connections between talk about fictional stories and what they say, and talk about the fictional characters about which things are said in the relevant stories." (*ibid.*, 150)

In reply to Thomasson, I would say that when she says that "Innings exist" or that "Fictional characters exist", ontological metaphors are presupposed. If this is so, then it is unclear whether the existence conditions arise from our psychological act of metaphorically conceiving of them as objects and as being created in the way objects are. In other words, if we conceive of them as objects, it might appear to us that their existence and identity conditions are like that of objects.

#### INNINGS IS A THING

The beginning of the innings. (like the beginning of a table)

The end of the innings (like the end of a table)

The innings drew to a close. (The door was closed)

I do not doubt that such innings or fictional characters exist. But not in the way they are commonly thought of in our ordinary practices. As my discussion in this chapter has shown, the nature of these things is more complex than it first appears (they might be ideas) and that Thomasson's account needs to be more fine grained to explain why existence and identity conditions arise. Perhaps the presence of ontological metaphor in our thought is not a strong objection to her view. However, that they are present and Thomasson might be mistakenly deriving existence and identity conditions from the metaphor can be construed to be a warning that ontological categories cannot simply be read off the way we speak.



Furthermore, perhaps a precondition for such an ontological metaphor is a bundle of mental representations which encode mind-dependent properties which keep changing (which constitute an idea, which is what a fictional character is) whereas the metaphorical attribution of objecthood is what is constant to the changing bundle, which shows how fictional characters change through time. I have not completely answered this question yet, but my account does seem to have its advantages. I shall explore this in the ensuing chapters. Moreover, Thomasson's question is, if fictional objects existed what would they be, given our literary practices. However, in the light of what I have said till now, especially the lack of dependence of fictional characters on literary works and the dependence on thought, the reformulated question I ask is: If fictional characters existed, what would they be, given our cognitive mechanisms?

## Chapter 4

### Salmon's Single Tier Theory and its Discontents

#### 1. Introduction

Salmon (1998) argues for the position that fictional names within fiction as well as metafictional sentences refer to abstract objects. This view is primarily in contrast with Kripke's (2013) and van Inwagen's (1977) which hold to a two tier theory of fiction, where fictional names in metafictional sentences refer to abstract objects. Like the previous chapters, this chapter too follows a pattern like Hegel's dialectic, of thesis, antithesis and synthesis, where antitheses, whether mine or by other philosophers will be laid out to Salmon's theses, from which will emerge a synthesis. Among other things that emerge in the synthesis, I begin to look more explicitly at mental representations constituting fictional characters, the divergence of facts about reference and facts on aboutness<sup>17</sup>, a possible solution to the problem of negative existentials and the nature of the proposition expressed by statements about fictional characters.

Kripke's (2013) view on names within fiction entails that a fictional name like Sherlock Holmes is a "rigid nondesignator". Salmon outlines a "homework problem" from Kaplan (1973), that if fictional names do not denote, then how are negative existential statements containing fictional names (like (1)) true?

1) Sherlock Holmes does not exist.

Both van Inwagen and Kripke have two tiered theories of fictional names, in which they allow for reference to fictional characters in metafictional statements (thus excluding fictional names in sentences prefixed with "according to the fiction") but fictional names in intrafictional statements do not refer. Thus, (1) for them turns out to be false, since fictional characters do exist. In this account, as per Salmon's gloss on the two tier theory, (1) comes out as true in contexts in which "exists" is not literal, but means real. Real in this context means "not merely a character in the story, but an entity of the sort depicted." (Salmon

---

<sup>17</sup> Crane (2013) makes the distinction between facts about aboutness and reference diverging and I am inspired by this distinction he makes. To the best of my knowledge, he makes the distinction first.

1998: 293) So (1) means that Sherlock Holmes is a fictional character but not a real person (as in Thomasson (1999)).

Thus, when we use names in fictional contexts, we are merely pretending to refer to a person and pretend to assert propositions. In van Inwagen's account, such fictional statements do not express propositions. In his account, "according to the fiction, Mrs. Gamp was fat" has as its logical form a three place ascription relation  $(\exists x)A(\text{fatness}, \text{Mrs. Gamp}, x)$ . For van Inwagen, as per Sarah Sawyer the truth condition for these sentences is:

It is true that a property is ascribed to a fictional character at a place in a work of fiction iff there is in/at that place a sentence which says of that character that it has that property. (Sawyer 2002: 187)

Thus, van Inwagen's account is what Sawyer (2002: 184) calls the "asymmetry thesis", where intrafiction and "according to the fiction, p" statements involve an ascription relation and literary critical statements involve a predication relation.

Kripke describes a two stage process which shows how statements outside fiction that are used to talk about fiction succeed in referring to abstract objects. The first stage is a metaphysical one, in which an abstract object is postulated by language when we use it while pretending, in the world of the fiction. At this stage, the name does not refer. The next stage, when we use the fictional name to talk about the fictional character (that is, in literary criticism) is a semantical one, "the language allows a grammatical transformation...of a fictional name for a person into a name of a fictional person." (Salmon 1998: 294) This in effect shows that the name Holmes is ambiguous, that is, that there are two distinct names instead of one which ought to have had different spellings, with 'Holmes<sub>1</sub>' as the name for pretend use and 'Holmes<sub>2</sub>' as the name for the abstract fictional character (what Sawyer (2002) calls the ambiguity thesis). So (1) becomes (Salmon 1998: 295):

(1\*) Holmes<sub>1</sub> does not exist.

(1\*\*) Holmes<sub>2</sub> does not exist.

Thus, from within the fiction, (1\*) is false whereas (2\*) is true. ((2) however, is ambiguous.) (*ibid.*, 295)

(2) Holmes plays the violin.

(2\*) Holmes<sub>1</sub> plays the violin.

(2\*\*) According to the fiction, Holmes plays the violin.

(2\*\*\*) Holmes<sub>2</sub> plays the violin

In the context of this discussion, Salmon differentiates between object-fictional statements like (2), made within the pretense, and metafictional sentences, like (2\*\*)

(2\*\*\*) as a statement about reality is false, since an abstract object cannot play the violin. However, when the statement is made about the fiction, it is true because: ““Holmes<sub>2</sub> plays the violin” is true iff according to the fiction, Holmes<sub>1</sub>(*sic.*) plays the violin.” (Salmon 1998: 295)

(1\*\*) is false both made about the world (“exist” taken literally”), since it is about an abstract object. It is also false with respect to the fiction, (“exist” in its “extended sense”) made about the fictional person, since it is true that according to the fiction, Holmes<sub>1</sub> exists.

However, (1\*) when taken about the world comes out to be true. The true nonreferring negative existential is thus (1\*), that we make when we say (1) and it is true.

But Salmon says that the homework problem still remains. How is a statement with a name that does not denote Holmes<sub>1</sub>, since there is no such thing true? In other words, Salmon says: “We have attempted to deal with the problem of negative existentials by concentrating on “Holmes<sub>2</sub> does not exist”. But it is Holmes<sub>1</sub>, not Holmes<sub>2</sub>, who literally does not exist. The homework problem requires more work. Kripke says that it is “perhaps the worst problem in the area.” (*ibid.*, 296)

Salmon outlines Kripke’s (2013) solution as follows. When people state “Holmes<sub>1</sub> does not play the violin” about reality, it should be seen as a special sort of speech act. It ought not to be seen on the same lines as “Holmes<sub>1</sub> plays the violin” uttered about the real world and which does not express anything. People might be unsure that there is such a person, or think that there is such a person, but does not play the violin. What the sentence then expresses is that there is no proposition that Holmes<sub>1</sub> plays the violin or that the proposition exists but is not true. Thus, Salmon claims “in short, the sentence is interpreted as meaning

*there is no true proposition that Holmes<sub>1</sub> plays the violin*. A similar cautious interpretation is available whenever negation is employed.” (*ibid.*, 296)

Similarly, when people say, “Holmes<sub>1</sub> does not exist”, “what one really means is better expressed by” (*ibid.*, 296) “there is no true proposition that Holmes<sub>1</sub> exists”. Kripke does not say that “Holmes<sub>1</sub> does not exist” is true because “Holmes<sub>1</sub> exists” is false in English but that it expresses that *there is no true proposition that Holmes<sub>1</sub> exists* (*ibid.*, 296) for the reason that there is no such proposition. Salmon calls this Kripke’s intensional ascent, as opposed to a semantic ascent.

Salmon’s first objection to Kripke’s account is that if a term  $\alpha$  does not refer, then a propositional term like  $\ulcorner$  the proposition that  $\alpha$  is bald  $\urcorner$  (*ibid.*, 297) will be empty. Thus, with respect to the proposition *there is no true proposition that Holmes<sub>1</sub> exists*, since Holmes<sub>1</sub> does not refer and propositional terms like  $\ulcorner$  the proposition that Holmes<sub>1</sub> exists  $\urcorner$  and  $\ulcorner$  the proposition that Holmes<sub>1</sub> plays the violin  $\urcorner$  (*ibid.*, 297) are empty, the propositional constituent  $\langle$ Holmes<sub>1</sub> $\rangle$  ought to be missing. In Kripke’s account, the analysis of a negative existential “Holmes<sub>1</sub> does not exist”, by another negative existential *there is no true proposition that Holmes<sub>1</sub> exists* leads to an infinite regress, because the question as to what the proposition the expression  $\ulcorner$  there is no true proposition that Holmes<sub>1</sub> exists  $\urcorner$  expresses is still left open, given that Holmes<sub>1</sub> is empty.

The second objection by Salmon (to two tier theories including Kripke as well as van Inwagen) is that an object fictional statement like “Holmes<sub>1</sub> plays the violin” (*ibid.*, 297) does not express any semantic content. This makes it difficult to explain what semantic content metafictional statements like “According to the fiction, Sherlock Holmes plays the violin” (*ibid.*, 297) express. Analogously, Le Verrier believed that Vulcan influences Mercury’s orbit, but given that Vulcan does not exist, and the name “Vulcan” expressed no content, it would lead one to say that Le Verrier believed nothing, which is odd. Furthermore, “If object-fictional sentences...express nothing and only pretend to express things, how can they be true with respect...to the fiction, and how can meta-fictional sentences involving object-fictional subordinate clauses express anything at all, let alone something true?” (Salmon 1998: 298)

His third objection is that Kripke’s account invokes an “intensional apparatus”. In statements like “Sherlock Holmes was cleverer than Bertrand Russell” (*ibid.*, 298), ‘Holmes’ would not only be ambiguous, but it would also refer to the concept *the brilliant*

*detective who performed such and such exploits.* (*ibid.*, 298) This is how the name in metafictional sentences like “According to the fiction, Holmes<sub>1</sub> plays the violin” (*ibid.*, 298) has an indirect (*ungerade*) use referring back to the name in the object-fictional sentences and thus the concept. Salmon says “Kripke acknowledges this, calling it a ‘special sort of quasi-intensional use’.” (*ibid.*, 298) If this is so, then we get propositions like ‘Holmes<sub>1</sub> plays the violin’ that are true without the need for a fictional operator. Moreover, if this is so, then any person who fits the description would make (2\*) true. Additionally, this undermines Kripke’s direct reference theory itself.

As a prelude to Salmon’s account, he says the following about Kripke’s account, and two tier accounts in general, as a methodological consideration:

Once fictional characters have been countenanced as real entities, why hold onto an alleged use of their names that fails to refer to them? It is like buying a luxurious Italian sports car only to keep it garaged. I do not advocate driving recklessly, but I do advise that having paid for the car one should permit oneself to drive it, at least on special occasions. (*ibid.*, 298)

This is a serious inadequacy in two-tier theories. The more serious inadequacy is that given that there is only the pretend use of a name in fictional cases, the question of a nonreferring or a referring use for object fictional sentences does not arise. As Salmon puts it:

A name semantically refers to this or that individual only relative to a particular kind of use, a particular purpose for which the name was introduced. One might go so far as to say that a pretend use by itself does not even give rise to a real name at all, any more than it gives birth to a real detective....The problem with saying that ‘Sherlock Holmes’ is nonreferring on Conan Doyle’s use is that in merely pretending that the name had a particular use, no real use was yet attached to the name on which it may be said to refer or not to refer. (*ibid.*, 299)

Of course, Salmon acknowledges that in these cases, Doyle would weakly use a name while writing it, but at most it is like writers reciting while practicing. From this Salmon gives his account of how names in fiction refer to abstract entities.

### **1.1 Salmon’s Account**

Doyle, while creating a story, created a fictional character as the protagonist and a name “Sherlock Holmes” for the protagonist. This character is an abstract artefact which is man-made and which is only a man named “Sherlock Holmes”, according to the story. As he tells the story, Doyle pretends to use the name (does not really use it) “Sherlock Holmes” to refer to the person Sherlock Holmes: “he pretends to be Dr. Watson using “Sherlock Holmes”, much like an actor portraying Dr. Watson on stage.” (*ibid.*, 300) Later, “the use of the name

is imported from the fiction into reality, to name the very same thing that it is the name of according to the story.” (*ibid.*, 300)

“Sherlock Holmes”, when it refers to the abstract artefact corresponds to “Holmes<sub>2</sub>” in Kripke’s account. But “Holmes<sub>1</sub>” in Salmon’s account is not really a use of a name, but is a pretend use. Salmon also escapes one objection to van Inwagen that if we take literary critical discourse at face value where it is said “Based on the evidence, Holmes was not completely asexual” (*ibid.*, 301), then we will have to take Holmes having a sexual orientation literally, whereas an abstract object cannot have that sort of property. In order to escape this, he says that scholars are not really using the name, but are pretending to use a man’s name. Thus, he says:

The only genuine, nonpretend use that we ever give the name—of which I feel confident—is as a name for the character. And that use, as a name for that very thing, is the very use it has in the story—though according to the story, that very thing is a human being and not an abstract entity.” (*ibid.*, 300)

If the account is correct, it so happens that every name in the story refers to the abstract fictional character and the sentences express singular propositions. However, Doyle merely pretends to be Watson who asserts them. Thus, taken at face value, the sentences are false, because they are actually about an abstract artefact.

This again would fall prey to the objection that a work of fiction is not really a bunch of lies about abstract objects. Or one might object that the essence of fiction is not that we are pretending that abstract objects are real things. But this, Salmon says, is a misunderstanding of fiction. Analogous to actors who play roles and are parts of a play, characters too play roles and are parts of a fiction. In a historical fiction, a real person is said to fictionally play a role in the fiction.

In Salmon’s account, it is “the essence of a fictional character to be depicted in the fiction as a person who takes part in such-and-such-events, performs such-and-such actions, think such-and such-thoughts.” (*ibid.*, 302) This is borne out when we think that what suited the role of the patriarch in *The Godfather* was the character Puzo created, just like Clark Gable was well suited to play the role of Rhett Butler in *Gone with the Wind*. On the other hand, this role of James Bond from Fleming’s novels could not have been played by the character of the butler from Ishiguro’s *The Remains of the Day*. Being depicted this way is not just a necessary and sufficient condition to be Fleming’s character in any possible world but also reveals the essence of the character.

With this account in mind, Salmon takes “Sherlock Holmes” to be a rigid designator which refers to the fictional entity both in the fiction and in the real world. Object fictional statements are false with respect to the real world as abstract objects cannot play the violin but true with respect to the fictional world. Yet, they are true in the fiction because they are

entailed by the propositions, themselves about fictional characters, that comprise the fiction, taken together with supplementary propositions concerning such things as the ordinary physical-causal structure of the world, usual societal customs, etc., that are assumed as the background against which the fiction unfolds. (*ibid.*, 303)

Moreover, metafictional statements now get semantic content, because they are true with respect to the real world iff they are true in the fiction, given that in fiction the sentences are shorthand for metafictional sentences. (*ibid.*, 303) Like Thomasson’s (1999) account, Salmon takes it that we pretend that statements that are true in fiction are true simpliciter. For him, the subordinate clause in the metafictional sentence has a Fregean indirect reference<sup>18</sup> to the false proposition about Holmes. This analysis thus explains how metafictional “according to the fiction, p” sentences become truth-evaluable and deals with the problem as given in Salmon’s second objection to two tier theories.

As a result, Salmon can now give an analysis of (1) for at least those statements involving fictional names. (1) or “Sherlock Holmes does not exist; he is only a fictional character” (*ibid.*, 303) is analysed as meaning that the fictional character is not *the very sort of entity depicted* with what he calls a “Pickwickian use of a name” (*ibid.*, 304), that is, a nonliteral use. It is pertinent to note that Salmon says: “Since this interpretation requires a reinterpretation of the name, it might be more correct to say that the speaker expresses this proposition than to say that (0) ((1) above) itself does.” (*ibid.*, 304)

Moreover, unlike Kripke’s *ungerade* use of Holmes<sub>2</sub> as parasitic on the description used in fiction, Salmon says that just in the way that we say that this person is a Sherlock and the like, in which those names are used as descriptions, “Sherlock Holmes” is a substitute for a description like “Holmes more or less as he is depicted in the stories” (*ibid.*, 303). Thus, (1)

---

<sup>18</sup> In Frege’s view, when it comes to clauses and sentences in indirect speech contexts, expressions do not refer directly to what their referents would be in direct speech, but designate/refer to (*bedeuten*) the senses expressed by the words of another person’s utterance. Thus, the words in the indirect speech context refer to the customary sense expressed by the words in ordinary contexts, and this is their *indirect referent* (Frege 1948: 211-212). Furthermore, the sentences and clauses in indirect speech contexts also designate/refer to the sense or thought (their indirect referent) expressed by the sentence uttered by another person. Frege applies this also to clauses and sentences in belief contexts and contexts involving other propositional attitudes like thoughts, desires etc.



would mean that Holmes as he is depicted in the stories does not exist in reality. This is the proposition expressed by (1).

## 2. Sawyer's Objections

Sarah Sawyer's (2002) summary of Salmon's explanation of why metafictional "according to the fiction, p" sentences are truth evaluable is: "If object-fictional sentences express propositions which concern the very same fictional characters as the corresponding meta-fictional sentences, the latter could be true or false in virtue of the supposed truth or falsity of the former." (Sawyer 2002: 191)

Salmon's explanation of the truth evaluability of metafictional "according to the fiction, p" sentences requires pretense, since on the face of it, object-fictional sentences like "Sherlock Holmes plays the violin" are (i) about abstract objects and therefore (ii) false. According to Sawyer, there are two levels at which pretense works to make Salmon's explanation succeed, the first level is to pretend that the proposition expressed by the sentence does not involve an abstract object, but a real man. The first level handles the issue with it being an abstract object. The second level is to pretend that the proposition is true rather than false. This deals with the issue of the sentence being false.

Thus, on Sawyer's understanding of Salmon:

A metafictional sentence of the form "In the fiction F, P" is true iff in order to understand F we are to pretend that the sentences that comprise F express true propositions about real people and one such sentence is 'p' (Sawyer 2002: 192)

Sawyer first objects to the second level of pretense. It is plausible to pretend in plays and daydreams that "I'm on the beach in California" (*ibid.*, 192) is true. But, suppose the number two is an abstract object, it looks like we cannot conceive how to pretend that "The number two is a man who likes to play croquet" (*ibid.*, 192) is true. Salmon's second level of pretense is more on the lines of this. Sawyer says that one way of resolving this is by saying that there is no difference between the two levels of pretense, and that to pretend that "Sherlock Holmes" refers to a man and not to an abstract object is to pretend that the proposition "Sherlock Holmes is a man" (*ibid.*, 193) is true. However, this only makes the two levels of pretense one, and we end up with the same conceivability problem faced by the sentence about the number two above.

The second objection by Sawyer is as follows. If we pretend that “Sherlock Holmes” refers to a man, then to appeal to any abstract object as a referent is explanatorily redundant. Her reasoning is as follows. I can pretend that I have a magic wand and I can pretend that my pencil is a magic wand (that an object is a magic wand). Salmon’s account will deny this distinction, since if I pretend that a magic wand exists, then it becomes the same as pretending that an object has other properties (in Salmon’s account). Thus, Salmon requires an additional argument for why one should have this additional object of which we pretend that it is something else. Since he does not offer such an argument, there is no reason why we cannot just pretend that “Sherlock Holmes” refers to a person, without the need for an abstract object.

Furthermore, Salmon says that it is a *de re* pretense rather than *de dicto*, since we don’t pretend that an abstract object is a man but *of* an abstract object that it is a man. However, Sawyer says that *de dicto* thoughts involve the thinker with a conceptual relation with them but *de re* thoughts necessarily have a non-conceptual relation between a thinker and the object of the thought (even though there may be a conceptual relation). “On this view, elements of the content of a given *de re* thought will not by themselves determine which object the thought concerns; some non-conceptual relation between the thinker and the *res* must be involved.” (*ibid.*, 194) However, neither is this non-conceptual relation a perceptual relation, nor can relations with other fictional objects be invoked without being question begging. Thus, “we are owed some account of what makes a particular abstract object the one relevant to the truth-conditions of any given object-fictional sentence”. (*ibid.*, 195) On the other hand if the *de re/de dicto* distinction is not about thoughts but kinds of “attribution of single kind of thought” (*ibid.*, 195), the account of relevance to truth conditions is required.

Thus, Sawyer’s final step in the argument presents a dilemma for Salmon. Either Salmon’s theory of pretense explains why metafictional statements are truth evaluable or it does not. If it accounts for it, then the asymmetry theorist can appeal to mere pretense in object-fictional statements without the need for abstract objects to explain the truth conditions. If Salmon does not provide such an account, then Salmon’s theory is inadequate based on his own criteria. Thus, on the first count, Salmon’s account is redundant and on the second, it is insufficient to deal with two tier theories.

To Sawyer’s argument about explanatory redundancy in the case of pretending of an abstract object that it is a man, Contessa (2008) objects that it is not a serious objection,

since the abstract object is required to provide a uniform account of reference in internal and external sentences, even if in internal sentences, one pretends that an abstract object is a man, in the case of Sherlock Holmes. Additionally, Contessa sees no reason why one cannot pretend that the number two is a man who likes to play croquet. Nevertheless, one possible reason he adduces on Sawyer's behalf for why pretending that the abstract object Sherlock Holmes is a person might be a problem is that since Holmes the person cannot exist in any possible world, it is difficult to pretend that something metaphysically impossible is the case. However, even though one might not be able to pretend that a logical impossibility is true, like a hula hoop is both round and not round, children do often pretend that metaphysical impossibilities are true. For instance, it is metaphysically impossible for a dog to be a horse, but children do pretend of their dogs that they are horses.

While Contessa has a point regarding the overall explanatory advantage of Salmon's theory, the reply shifts the goalpost from the specific objections Sawyer levels at Salmon to the general objective of the theory. According to Sawyer's arguments, Salmon does not have an account of the non-conceptual relation of a thinker with the objects of thought, his account cannot show why one cannot just pretend that "Sherlock Holmes" refers to a man rather than it referring to an abstract object and that his account involves circularity. These are very specific arguments, and in the absence of a specific response to each of these arguments, Salmon cannot merely legislate that names in object-fictional sentences refer to abstract objects because it is generally advantageous to have a single tier theory. The two tier theory would be advantageous especially if the single tier theory fails to provide a specific account for "what makes a particular abstract object the one relevant to the truth-conditions of any given object-fictional sentence" (*ibid.*, 195). When I speak of a specific account, I mean an account of how these specific parts work, like how there can be non-conceptual relations of the abstract object with thoughts and how one can still make a difference between pretending a magic wand exists and pretending of a pencil that it is a wand. In the absence of such an account, the two tier theory will fit the data from fiction more than the single tier theory, since Salmon does not have the machinery to specifically account for the data.

Now, Sawyer might concede to Contessa's point and say that it is possible to pretend that metaphysical impossibilities are true. Nevertheless, she might bring up her point about the absence of an account of how thought can be related *de re* to abstract objects that fictions are supposed to be about. If such a relation is absent, then it is tough to conceive and provide an account of how we can be thinking of such an abstract object *de re* and then pretend of it that it is a man. In that case it is explanatorily simple that we just pretend that

“Sherlock Holmes” is a man and this account is already available. This is further reinforced by Sawyer’s point about Salmon’s account removing the distinction between pretending that a magic wand exists and a pencil being a magic wand. If this distinction persists, it is tough to see why pretending of an abstract object that it is a man is on the same lines of pretending of a pencil that it is a magic wand and not on the lines of pretending that a magic wand exists.

To this, Salmon might reply that pretending that the magic wand exists and pretending of a pencil that it is a wand are the same type of pretense, namely *de re* pretense of something that it is something else, in the former case of an abstract object and in the latter case of a pencil. This way, he can claim that a distinction is maintained between pretending that a magic wand exists and pretending of a pencil that it is a magic wand. Within this, he makes a further distinction, between pretending *de re* of concrete objects and pretending *de re* of abstract objects.

This seems unsatisfactory. What are we pretending of in the magic wand case that it is a magic wand? In the “Holmes” case, we are pretending of the abstract referent of “Holmes” that it is a man. The magic wand case is not that clear. If we pretend of an abstract or concrete object that it is a magic wand, then it is assumed by the pretense that we pretend of that it is a token of the type [Magic Wand]. Let’s look at a general definition of what it is to be a token of a type:

X is a token of a type y iff x possesses all or most defining properties associated with y.<sup>19</sup>

Thus, Socrates is a token of type [MORTAL] since Socrates possesses the property of *being able to die*.

In the case of the magic wand, a magic wand would be the token of the type [MAGIC WAND] iff it possesses properties like *being able to cause action at a distance*, *being a rod*, *being able to cast spells* and the like. As mentioned before, when I pretend of a concrete or abstract object that it is a magic wand, I pretend that it is a token of the type [MAGIC

---

<sup>19</sup> I prefer Jackendoff’s (1988) *preference rule system* to judge tokens of types, which does not require an account of necessary and sufficient conditions for counting a token as a member of a type. However, his account is to allow for exceptions. Since we are not dealing with exceptions now of necessary and sufficient conditions, like borderline cases of red and orange, necessary and sufficient conditions are sufficient for this task, since he does acknowledge that such conditions account for categorization in standard cases.

WAND]. In order to do that, by definition, I pretend that it possesses all or most defining properties a token of type [MAGIC WAND]. However, in order to do that, I already have to imagine a token of the type [MAGIC WAND]. This is where Sawyer's objection comes in again. If I can already imagine a token of type [MAGIC WAND] then the abstract object is explanatorily redundant. Even in the case of pretending of a concrete object that it is a token of type [MAGIC WAND], all there is the object and the pretense but no abstract object.

Salmon's response to this would be to reiterate his second and third objections against Kripke, namely that the problem of semantic content of object fictional sentences still remains and that this will lead to an "intensional ascent". Sawyer still might reply to this that it is the fact of the mental act of pretense that lends semantic content to object-fictional statements, so that the truthmaker for metafictional statements of the form "According to the F, p" are the facts of someone (or a story) representing that p, without p having to exist, the properties being represented being representation-dependent (see Crane 2013: 135). Representation-dependent properties are "properties which depend upon the fact that the object is being represented in some way: in thought, language, pictures, and so on." (*ibid.*, 2013: 68) Perhaps they are pleonastic properties, that is, properties like being prime that follow from the truth of statements like "3 is a prime" (from this it follows that "3 has the property of being prime"), "without any further metaphysical assumptions" (*ibid.*, 68) Or perhaps she can take up my claim, which is that the best possible explanation for such representation, at least in fiction, would be because the properties are mind-dependent, that is, they cannot exist without being thought of. Taking a cue from Thagard (2019: 30), I take mind-dependent features like sensory features like *how cats look and sound like*, motor features like *how it feels to pick up a cat*, emotional features like *how much you like cats*, and verbal features like cats being a kind of mammal. To this I'd also add properties derived from innate ideas which are not derived from sense modalities. I express these features as *being a thought of being a cat*, *being a thought of being round* etc. Thus, Sawyer can account for the semantic content of object fictional sentences, which is given by these features. (I have already given a glimpse of what these features are in the previous chapters and I shall clarify more on their nature in Chapter 6)

However, the problem with this sort of account is that if the term "Sherlock Holmes" does not refer to anything at all, then part of the truth conditions for "According to the fiction, Sherlock Holmes lives on 221B Baker Street" merely becomes "According to the fiction, Sherlock Holmes lives on 221B Baker Street" is true only if a work of fiction X says

“Sherlock Holmes lives on 221B Baker Street”.<sup>20</sup> But in that case “Sherlock Holmes” becomes mere mention (rather than used) even though it is not.

Sawyer’s account has further problems. If we ask her what it is to pretend that a magic wand exists, she will say that it is to pretend that a token of the type [MAGIC WAND] exists. If I ask her, what it is to be a token of the type [MAGIC WAND], she would have to give the account of what it is for a magic wand to be a token of the type as given above, the understanding of which is presupposed by the pretense. She thus has to accept that even if one cannot infer from “I pretend that a token of the type [MAGIC WAND] exists” or “a magic wand exists” that “(a token of) a magic wand exists”, one can at least infer that the representation exists (the representation is very also useful in the explanation of what it is for a magic wand to be the token of a type [MAGIC WAND]) Let us see how so.

When I pretend that a magic wand exists, I minimally represent the object of my pretense. That can be formulated as x represents y where x is me and y is the magic wand. However, the relation x represents y would entail the existence of y, given that relations entail their relata. Thus, this is not what the pretense involves. However, the case of representation itself seems a bit different. X can represent y, where y is a magic wand, but y does not exist due to an ambiguity in “x represents y”. It can be:

1) Where y is a magic wand (or a unicorn or Holmes), where y seems to be the object of the representation. However, in this case, the object of the mental act is not the magic wand itself, but a representation. So from x represents y, one cannot infer that the magic wand exists, but minimally, that the representation does.

2) Where y is a representation, y exists. So in “x represents y”, y is the representation. This can be strange too. One can say that given that the representation of a magic wand exists, given that it is *of* a magic wand, therefore the magic wand exists. To avoid this confusion so as not to invoke the magic wand, it is therefore clearer to say that the representation with mind-dependent properties like *being a thought of being a rod* etc. is likened to a concrete object which allows us to talk about a representation of a magic wand. Thus, the representation of the magic wand exists.

The problem Sawyer has is that whether or not she sees abstract objects as necessary to the

---

<sup>20</sup> I say only if, since there will be other conditions like there is such a such representation. But the one necessary condition suffices for my argument.

pretense, she is bound to see mental representations as indispensable to a pretense. Of course, she might need a further analysis of negative existentials (which I give below) and something like my master argument from (MET<sub>FIC</sub>)\* to more fully accept that the magic wand exists (reiterated below and in other chapters). Nevertheless, in the light of this discussion, one can say that we do in fact refer to our mental representations when it comes to fictional characters to salvage a semblance of a realist view, otherwise one (especially Sawyer since she is bound to accept mental representations as indispensable to the pretense) is open to my use mention objection above. This does perhaps bolster the case for Everett and Schroeder's (2015) view that fictional characters are ideas constituted by mental representations, where we do refer to mental representations when we think of fictional characters. Additionally, if one needs to explain what it is for a magic wand to be a token of the type [MAGIC WAND] my attempt at the explanation would be that the representation possesses mind-dependent properties as their content (when it comes to fictional entities) associated with the type, which are likened to a concrete object.

### **3. Interlude: My Objection to Salmon**

Observe that I have preferred mind-dependent properties that are encoded by a mental representation being likened to a concrete object rather than it being pretended of it that it is a concrete object, like in Salmon's account. As of now, I see two issues with Salmon's account. One is that if one notices above, that in the case of our representation of Holmes,  $x$  represents  $y$  does not entail that  $y$  exists, but there is a representation of  $y$ .

Keeping this in mind, the problem with Salmon's account is as follows. Suppose one pretend-uses the name "Holmes" in an object fictional sentence. In that case, " $x$  represents  $y$ " where  $y$  is Holmes is false, it would not entail that Holmes the abstract object exists, since there is only pretend use of the name. If this is so, then there are some possible worlds in which Doyle's pretend use does not create an abstract object. If in those worlds the abstract object is redundant and pretense is sufficient for fictional names, there is no reason why in the other worlds it is not. This seems to be in the spirit of Sawyer's objection.

Moreover, Salmon would be hard-pressed to deny that Doyle while writing object fictional sentences and pretend-using the name at least represents Holmes to himself. On Salmon's account, this by itself does not entail the existence of Holmes (the abstract object), since the pretend use of a sentence like "Holmes plays the violin" taken on its own does not entail the existence of Holmes. This is something Salmon would agree with, since his motivation to

add to the two-tier account does not come from such entailment. Moreover, the abstract object hasn't been created while Doyle is pretend using the name.

But let us consider the following dilemma. Either the representation is identical with the abstract Holmes or it is not. If the representation is identical with the abstract object, Salmon's account would reject this, since the name is pretend-used and the abstract object has not been created yet. If the representation is not identical with the abstract object Holmes, then since it is not itself a flesh and blood human being, it must be pretended of it that it is one. There are two problems with this. First, it seems that at every step, even after the abstract object is created, if one represents Holmes as well, one has to pretend of the abstract object as well as the representation that it is a man. The abstract object seems redundant, given that one can make do with the representation as the referent. Second, if one has to pretend of the representation that it is a man, the pretense itself is redundant. This is because what Salmon calls the "pretense" would be preceded by x representing y, which does not entail that a concrete human being exists. If we however, look at the phenomenology of our representations, the representation itself appears to be that of a flesh and blood man (say playing the violin), rather than us pretending of our representation that it is a flesh and blood man. If the representation precedes the pretense, then something other than pretense is required to explain its phenomenology.

This is why I have proposed that the representation itself be composed of representation-dependent properties (that are mind-dependent) which are likened to a concrete object, due to an ontological metaphor (as discussed in my earlier chapters) to account for this. This is however, not a knockdown argument, and will be bolstered in the ensuing discussion. Unlike Salmon, there is no need for pretense in the semantics or the aboutness of fictional entities in my account, and it thus escapes the spirit of Sawyer's objection to him.

#### **4. Contessa's Objection**

In this section, I outline the problem that Contessa (2009) has with Salmon's account. As we have seen, it follows from Salmon's account that if any referring expression does not refer to a concrete object in an object fictional sentence, then it refers to an abstract object. "Sherlock Holmes' brother" refers to an abstract object of which we pretend that it is Holmes' brother and "Sherlock Holmes' violin" (*ibid.*, 260) refers to an abstract object of which we pretend that it is a violin. From this, Contessa points out that "his head" in "Sherlock Holmes shook his head" (*ibid.*, 260) will refer to another abstract object of which



we pretend that it is Holmes' head, and this would show that Doyle created two abstract objects, Holmes and his head.

However, Contessa imagines Salmon as replying that when we pretend that the abstract object is a person, we simultaneously pretend that it has a head (Contessa 2009: 260). "His head" refers to the head that the pretend person has. But the problem Contessa finds with this is as follows. If the "pretended" head does not exist, then this violates the Axiom of Existence, namely that "Whatever is referred to must exist". If the pretended head is a real object then there are two possibilities. If it is an abstract object, then it runs into the problem that Doyle created two distinct abstract objects<sup>21</sup>. If it is another kind of real object, then Salmon owes an account of what kind of object it is. Moreover, this might not look like much of a problem at first, but it gets aggravated below.

Suppose in response, Salmon merely bites the bullet and says that yes, Doyle did create at least two objects, Holmes and his head and Holmes is a mereological fusion of the abstract object that is referred to by "his head" and "his torso". There are also other objects like "Holmes' nose" etc. However, the problem with this reply that I see would be that unless otherwise indicated, we import our assumptions about the concrete world into the fictionalized world. We will assume that Holmes' body is composed of cells, which will ultimately be composed of quarks. If Salmon is right in his reply, then if abstract object that "Holmes" refers to is a mereological fusion of the ultimate constituents of the world, then there are innumerable abstract objects of which we pretend that they are the ultimate constituents of the world that make up Holmes' body. This seems very unpalatable and non-parsimonious. Now, even though it is implied that the content of the fiction includes all the relevant scientific truths of the world, the problem really is not that these truths are imported. The real problem is that the way in which Doyle creates two abstract objects of which Holmes is the mereological fusion, Doyle also ends up creating the innumerable abstract objects of which it is pretended that they are the ultimate constituents of the world making up Holmes' body. This is in addition to the innumerable constituents of the world. In effect, Salmon will have to show how the human mind in the act of pretense can create so many abstract objects. Even though he might have such an account, he will run into the following problem Frederick Kroon (2013) points out.

---

<sup>21</sup> While it does not seem to be a deep problem for Salmon, I have included this for the sake of the dialectic that follows.

Anthony Everett (2005) in his challenge to the artefactualist account, lays out two criteria for the identity of fictional objects which the artefactualist's account violates.

P1) If the world of a story concerns creature a, and if a is not a real thing, then a is a fictional character.

P2) If a story concerns a and b, and if a and b are not real things, then a and b are identical in the world of the story iff the fictional character of a is identical to the fictional character of b. (ibid., 627)

Kroon takes these principles and argues that given the way groups are described in fiction, in the artefactualist's picture, they would violate P2). Suppose there are multitudes of dwarves taking part in the War of Dwarves and Orcs in *The Lord of the Rings*. Since there is no name or description that would enable a reader to distinguish one from another, they are in principle indistinguishable. Thus, there is no criteria for a to be identical with itself or to be distinguished from any other. The same applies to the innumerable abstract objects that Doyle creates of which it is pretended that Holmes' body is constituted by them. Just like the case of the Dwarves and Orcs, there is no name or description available to the reader to distinguish each of them from another, and there is no criteria of identity. If there is no entity without identity, such abstract objects cannot exist or come into existence.

Nevertheless, Contessa's criticism of Salmon is not convincing, since terms like "Sherlock Holmes' violin" might be descriptions that are not referring terms. That being said, it would be in Contessa's favour if one says that the author merely stipulates the descriptions as proper names in the fiction and merely stipulates that there would be as many abstract objects as in the concrete world, like in the case of cells. Surely such things can be stipulated, which is why Kripke (2013) says that names from within fiction cannot adjudicate between choosing a descriptivist theory or a direct reference theory, since descriptions can be pretended to be names and vice versa in the fiction.

Thus, Salmon might need to look for another account of what kind of entity a fictional character is. Perhaps Everett and Schroeder's account that fictional characters are ideas that are constituted by mental representations might be something that can resolve this issue. A representation can merely encode a mind-dependent property like *being a thought of being made up of cells*, etc. One does not need to have innumerable abstract objects of which it is pretended that they are cells that compose his body since the content is given by properties encoded by our mental representations. Like the properties of our imagination that show objects as incomplete, as in there is no determinate fact of the matter on how many hairs a unicorn, which is an object of our imagination has, there is no fact of the matter about the number of cells that needs to be given by properties of our mental representations.

## 5. Hayaki's Objection

Hayaki's (2009) objection to Salmon stems from taking (3) to (5) as disanalogous, which reveals a tension between them. She puts forward the following statements (*ibid.*, 141)

(3) The Canon is about Sherlock Holmes.

(4) The Canon is about a brilliant private detective who solves many crimes that baffle Scotland Yard.

(5) The Canon is also about Queen Victoria (among other things).

Subsequently, (5) entails (5') which is that Queen Victoria exists and that the Canon is about her. So as per Hayaki, (5') is (*ibid.*, 145):

(5')  $(\exists x)(x = \text{Queen Victoria} \ \& \ \text{the Canon is about } x)$

Moreover, if we accept that (6) entails (6') (*ibid.*, 144), especially given Salmon's picture where "Sherlock Holmes" refers to an abstract object in metafictional statements prefixed by "according to the fiction" then, (3) entails (3') (*ibid.*, 145):

(6) It is true according to the Canon that Sherlock Holmes is a detective.

(6')  $(\exists x)(x = \text{Sherlock Holmes} \ \& \ \text{it is true according to the Canon that } x \text{ is a detective})$

(3')  $(\exists x)(x = \text{Sherlock Holmes} \ \& \ \text{the Canon is about } x)$

If (3') is true, then aboutness can be "a genuine relation between fictions and objects" (*ibid.*, 145)

But (4') is false, because the Canon is not about a real detective (*ibid.*, 145):

(4')  $(\exists x)(x \text{ is a brilliant private detective who solves many crimes that baffle Scotland Yard} \ \& \ \text{the Canon is about } x)$

Given the falsity of (4'), it shows that aboutness is not a genuine relation, since there is no relation between the canon and a detective despite being about a detective.

From this, Hayaki concludes that "The Canon is about..." is an intensional context" (*ibid.*, 145), in which "Holmes" is not referential. Furthermore, the truth conditions for "the Canon is about x" as per Hayaki are "The Canon is about x" is true iff it is true according to the Canon that x exists and "The Canon is about an F" is true iff it is true according to the Canon that there is an F.

Hayaki appeals to a general principle, assumed by the truth conditions, let's call it (CREAL) (*ibid.*, 145):

(CREAL) If the Canon is about x, and x is real, then there is a real object, x, such that the Canon is about x.

But, if x is an abstract object, and x is real, then the Canon should be about the abstract object. Thus, in the case of Sherlock Holmes, which is abstract and real for Salmon, the Canon should be about the abstract object and not a flesh and blood person. But that is a wrong result.

One objection that Hayaki thinks might be levelled against her is that if aboutness is not a relation, then the existential quantification over Queen Victoria in (5') will be impossible, which is odd. Hayaki's reply is that even if it is not a relation, there "does seem to be some relation in the vicinity." In a sense, Queen Victoria appears in the novel the way Napoleon appears in War and Peace. They both have the common property of appearing in a novel and this is extensional, since co-referring terms will preserve truth value. Thus appearing in it would also treat Queen Victoria and Holmes alike as both appearing in novels in the Canon. But again, if we take Holmes to be an abstract object, the tension between (3) and (5) appears again. An abstract object will appear in the Canon which will make it about the abstract object, but (3) and (5) both are about flesh and blood people.

One reply by the Salmon might be to say that (5') can be accepted but not (3'). But this would be to reject (CREAL). But a good reason would have to be given to reject it, for which Salmon will have to appeal to a two tier theory of names. But that would be unacceptable to him.

Thus Salmon would have to see (3) and (5) as disanalogous, whereas they are not.

In my view, Salmon might be able to answer this objection if he modified his response to another objection of mine to his account. Suppose an author decides to create a fictional character that is fused in the sense of it becoming like conjoined twins of Queen Victoria and Holmes called Schlomes (not in the mereological sense). This should not be that unfamiliar. Think of this in analogy with a two-headed monster which in a sense is one character but in another sense two characters. One might suppose the following metafictional principle for the sake of this work of fiction:

(F) For any actual or fictional characters  $x$  and  $y$  appearing in a work of fiction  $Z$ , if  $x$  and  $y$  are fused into conjoined twins, a new object is thereby created in the world of the fiction.

If the sort of “joining” that I talk about above can happen, one wonders how it is even metaphysically possible to fuse an abstract object together with a concrete person. This is way worse than the question of how mind and body interact if they were two different kinds of entities. Or it is pretended that it is a fusion. There was an abstract object of which it is pretended that it is a person, and there was Queen Victoria and it was pretended that they were fused together as Schlomes. However, by means of this fusion, a new fictional object ought to have been created as per Salmon’s theory, given (F). If a new fictional object is created, then it is a literal fusion to form a unitary whole new object, and it ought to be wholly abstract. As we saw above however, it is impossible for there to be a literal fusion and obviously, the new object is part concrete and part abstract. Thus, a new fictional object is not created. Since a new fictional object ought to have been created and it is not, Salmon’s theory is false.

However, Salmon might have a response to this. He would reply that “Queen Victoria” and “Holmes” would in fact refer to separate entities and it is pretended of Queen Victoria that she is a fictional character. The pretense of Queen Victoria as a fictional character fits with the intuition that the author is giving a fictionalized account of Queen Victoria. From there, it is just one step to say that the author is pretending of Queen Victoria and the fictional character Holmes that they are fused like twins and that they are one character. Salmon would thus claim that no new fictional character has in fact been created. (F) is only held in pretense, and is an object fictional assumption.

There are two things to be observed from this. In this picture, facts about reference, that “Holmes” and “Queen Victoria” refer to separate entities, and facts on aboutness (that the author can pretend that they are fused and are one character) seem to diverge. This is a crucial step. The second is that the aboutness facts seem to have a mode of presentation, the way our mind presents something to us.

To this revision, the modified Salmon would have to still face two objections. The first one would be to the proposal that it is pretended of Victoria and Holmes that their pretended fusion is one character. Salmon would use a principle of character generation like this:

(PCG) If an author pretend uses a name and that name is imported into a proper use in reality, then a new character is created.

Thus, if the author pretend uses “Schlomes” and the “use of the name is imported from the fiction into reality, to name *the very same thing* that it is the name of according to the story” (*ibid.*, 300), he thereby creates a new character, as per Salmon’s own criterion. Unlike the reply Salmon might have given that one pretends of the fusion that it is one character, there is in fact a new character now<sup>22</sup>.

In my account, I propose that fictional characters are constituted by bundles of mental representations which encode mind-dependent properties. This avoids the above problem. What makes the two the same character is that the properties that are mentally abstracted from the concrete Queen Victoria as well as the mind-dependent properties that are attributed to Holmes that are further likened to a single concrete person. It is the combination of mind-dependent properties abstracted from Queen Victoria and Holmes that combine to form a unified mental representation which would account for the quantification over a single entity as well as the objectual metaphor that presupposes mental association with a concrete person (metaphor in cognitive linguistics is understood as a “pattern of conceptual association, rather than to an individual metaphorical usage of linguistic convention” (Grady 2010: 188)) that accounts for Schlomes being a new fictional object. This account is also advantageous because Salmon seems to falter with my objection because the fusion is difficult for two different kinds of objects, one abstract and one concrete. In this case, the bringing together of the same kind of properties by means of

---

<sup>22</sup> Brock (2010) goes against a principle like (PCG) using many versions of it and with many counterexamples.

bundling mental representations, which accounts for how the combination takes place and it thus forms one fictional object.

The second objection would be for Salmon still having to deal with Hayaki's objection that "the Canon is about..." is an intensional context" (*ibid.*, 145) since (4') is false as the canon is about a detective but there is no relation between the Canon and any detective.

(4')  $(\exists x)(x \text{ is a brilliant private detective who solves many crimes that baffle Scotland Yard \& the Canon is about } x)$

However, this should not be a difficulty with my account, since all statements like (4') will presuppose "Sherlock Holmes is a fictional person" which is elaborated as (MET<sub>FIC</sub>) (see Chapter 1 and 2):

(MET<sub>FIC</sub>) Sherlock Holmes is fictional but is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

Which entails:

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

Thus, the application of the predicate "is a brilliant detective" will be applied to the ontological metaphor in which a mental representation is likened to a person. To apply it otherwise would be a category mistake, since a mental representation cannot literally be a brilliant detective. If that is so, the Canon is about x and (4') is true, given the presupposition.

It should be observed that Salmon does not have my reply directly available to him to escape Hayaki's objection. (4') would be true for him, since the first part of the conjunct is not within the scope of pretense. However, he can modify his account by saying that like (MET<sub>FIC</sub>), (4) can presuppose (MET)<sub>P</sub>

(MET)<sub>P</sub> Sherlock Holmes is fictional of which it is pretended that it is a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

There are two advantages of this presupposition account here. First, the predication in the first part of the conjunction of (4') that *x* is a brilliant private detective who solves many crimes that baffle Scotland Yard is made true because there is in fact such an entity presupposed by the statement. This kind of relation to the abstract entity (in my picture, a relation to a mental representation, not an abstract entity) accounts for the reference facts. Given that I indicated above that there is a divergence between reference facts and aboutness facts, the aboutness fact is given by the relation "is likened to" in "x is likened to a real person" in (MET<sub>FIC</sub>) and "of which is pretended that" in the second part of the conjunction "the Canon is about x". This explains why the Canon is about a flesh and blood person.

However, compared to my account, Salmon would be at a disadvantage. Imagine that Conan Doyle, like in one of my objections to Thomasson in Chapter 3, in a state of drug induced delusion, writes down the contents of his delusions. When he comes out of his delusion, he realizes that he has written a wonderful story and passes it off as fiction. If you ask Conan Doyle, if the character of fiction and the character of his delusions is the same, he will say "Of course it is!" One might ask if the delusion could bring an abstract object into existence and doubt Doyle's version. On the face of it there is no reason why it should not (but perhaps it is more plausible to say that an idea is created, since an idea is not causally inert and part of the causal nexus of the world but an abstract object is questionable in that respect). In Salmon's account where "the use of the name is imported from the fiction into reality, to name *the very same thing* that it is the name of according to the story", there is no reason why the use of the name cannot be imported from the delusion to reality, especially when a doctor needs to talk about the object of Doyle's delusion in a sentence like "Holmes is a delusionary character." The sentence would involve quantification over Holmes too, as in a metafictional sentence. Let us call this a metadelusionary sentence (see discussion below from Caplan (2004)). In a manner like this, even by Salmon's lights, one can say plausibly that the delusion brings an abstract object (by mine and Everett and Schroeder's lights, an idea) into existence. As for believing Doyle, it's the plausible intuition that he merely took the same object of his thought from his delusion and imported it into fiction. Saying otherwise would raise questions such as would it not be more parsimonious for it to be the same object given that it's part of the same narrative by the same author and with the same characteristics? Would it not be also more parsimonious then, to think of the same object now as an object of delusion and now as an object of fiction? Moreover, Doyle had merely written down the content of his thoughts. He takes the same words and passes them off as fiction. He thinks to himself this would be good fiction. People understand the same



thing when they read the words, only that Doyle is aware that it was the contents of his delusion. Even if someone later discovered that after all the sequels that have been written on Holmes that the original text was the contents of Doyle's delusion, it would at best be an interesting curiosity and some people might want to study this phenomenon or even think of the text as the contents of Doyle's delusion, but it would be absurd to say that Holmes as a fictional character would thereby cease to exist since he is now thought to be a delusionary object or that he is not well cherished. At worst, people would think of him as originating in the contents of Doyle's delusion but that entails nothing as to its ontological status as a separate object (like one would think of an adult as originating in a womb).

Observe that for Salmon, it is essential that it is pretended of an abstract object that it is a flesh and blood person to account for its aboutness. However, in the Conan Doyle case, the aboutness of the Holmes of delusion and of fiction is the same, and there is no pretense involved in the delusion. Yet, the characters are identical. My account is neutral and can account for the same process that accounts for the same aboutness in the delusion and the fictional case, since it indicates a deeper cognitive process of objectual metaphor with the "is likened to" relation (see Lakoff and Johnson (2003) who argue that that these metaphorical processes underlie all our thought).

In the light of Ben Caplan (2004)'s pointing out a gap in Salmon's theory, my arguments gain more significance. Suppose there is the following inference (Caplan 2004: 335, n 12).

(P1) Authors create creatures of fiction.

(P2) If authors create creatures of fiction, then myth-makers create creatures of myth.

(P4) If myth-makers create creatures of myth, then imaginers create creatures of imagination.

(C2) So imaginers create creatures of imagination.

Suppose, somebody rejects (P2). Given that the difference between myth and fiction is in the propositional attitudes, belief and make-belief respectively, the one who rejects it must say that make-believing is "ontologically special that enables the creation of abstract objects whereas believing is not." (*ibid.*, 334) Caplan says that it is difficult to see why make-believing has a special ontological role but not believing. The fiction-only creationist thus

minimally owes an account of why this is so. According to Caplan, Salmon would accept (P2) but either reject or be neutral towards (P4). If he rejects it, then he would have to say that make-believing and believing are ontologically special but imagining (entertaining in one's mind) is not. Moreover, he will have to give an account of why both make believing as well as believing are ontologically special but not imagining.

There is no such issue with my account above, since the "is likened to" relation can be common to imagining, believing, make-believing and is present even in a delusionary state. Given this, it is tough to see why neutrality towards (P4) should be allowed.

## 6. Predelli's Objection to Salmon and Modification

Stefano Predelli (2002) lays out the key tenets of a position like Salmon's, which he calls the Character Reference thesis (MCR Thesis). The following are its tenets with respect to Holmes' example, where H is the literary character Holmes and <H, smoked a pipe> is the proposition expressed by "Sherlock Holmes smoked a pipe":

(MCR1) All genuine uses of 'Holmes' are character-referring uses, i.e., they refer to H.

(MCR2) According to the Holmes' stories, H is a pipe smoking detective who befriended a doctor named 'Watson' and who helped the police solve a variety of baffling crimes.

(MCR3) 'Holmes' is a proper name which, consistently with the Millian theory of names, has its referent H as its semantic value. (*ibid.*, 267)

Predelli points out that Salmon's analysis of (1) is different from that of the MCR thesis. In Salmon's analysis, (1), that is "Sherlock Holmes does not exist" is true iff "there is no one who is both H and sufficiently like that (as depicted in fiction)." (*ibid.*, 269)

However, with respect to the actual circumstance, since there is no referent of "Holmes", the sentence fails to have a Millian compositional analysis, the desired result. If, however, the Millian proposition is given pragmatically, "Holmes" refers in that case, which will make (1) false in actual circumstances. The other problem with Salmon's account that Predelli points out is the following. Suppose someone who watches the movie *Amadeus* says, "I wish I had been alive back then, so that I could meet such a talented young man." (*ibid.*, 270) Since the listener thinks that the film is fictional, he says, "You're confused. Mozart did not exist." (*ibid.*, 270) If we apply Salmon's account, it would come out as true, since the actual

Mozart did not have many key properties attributed to Mozart in the film and the actual Mozart is not thus identical to the Mozart as depicted in the film.

Suppose Salmon revises his analysis of (1) and says that it merely expresses the proposition <Holmes, nonexistence> as per the MCR thesis. Predelli shows that it too will run into problems as we shall see below.

According to Predelli, one advantage that this account has is that it accounts for the truth values of sentences like (1) (“Sherlock Holmes does not exist”). Consider the following to see the advantage of the revised thesis (*ibid.*, 267):

(7) Holmes helped the police solve many baffling crimes.

Let *u* be an utterance of sentence *S* in context *c*, typically represented by a sentence-context pair <*S*, *c*>.

*u* is true iff the content expressed by *S* in *c* is true with respect to circumstance determined by *c* (*ibid.*, 262)

An utterance of (7) in a factual circumstance will come out as false. But in circumstances determined by a fictional context, it will be true, since according to the fiction, Holmes helped the police solve many baffling crimes. Predelli considers (8) (*ibid.*, 268):

(8) Watson must have personally indulged in the unhealthy habits he would later attribute to his imaginary friend. How else can one explain the obsessive insistence with which he continues to believe in the existence of the mystical detective, roommate and companion of adventure? No, my dear Watson: Holmes does not exist!

But (1) is false in (8) because in the fictional circumstance, Holmes is flesh and blood and not imaginary. Furthermore, Predelli brings up the following (*ibid.*, 268):

(9) Contrary to widespread belief, Doyle did not actually write much, with the exception of a rather dull autobiographical account of the early years of his life. Holmes, the literary character who is alleged to have inspired legions of successive writers and film-makers, does not exist. All we have is an elaborated fraud, built up by benevolent biographers from bits and pieces collected from inferior novelettes popular at the time.

(1) is false in (9) too, since in factual circumstances, Doyle did create Holmes.

The problems that crop up for the MCR Thesis that Predelli discusses are as follows.

Consider (10) (*ibid.*, 269):

(10) There is no point in searching all the apartments in Baker Street, looking for stacks of cocaine. Holmes does not exist.

(10) ought to come out as true when evaluated with respect to actual circumstances.

However, since  $\langle H, \text{nonexistence} \rangle$  is expressed by (1) as per the MCR thesis, it would be false.

Predelli now asks us to suppose that someone wrote the following about Moriarty, Holmes' antagonist (*ibid.*, 270):

(11) It is doubtful that Moriarty is at all interested in the dark tales of popular gothic literature. Perhaps, he would be attracted by the refined profile of Fitzgerald's Gatsby, were that literary character available to him. Paradoxically, his favorite literary character would be the character of Holmes, that contradictory synthesis of cold rationality and decadent qualities. But, of course, Holmes does not exist.

Assuming that the world of the fiction includes works by Shakespeare, Hamlet would be a literary character in it but not a flesh and blood person. Given that Holmes would be a flesh and blood person in the world of Doyle's fiction and not a literary character, as per (MCR2), (1) would be true for (11). However,  $\langle H, \text{nonexistence} \rangle$  would be false in this scenario, which is not correct. Predelli considers (*ibid.* 272):

(12) According to the CIA, law enforcement in England has been infiltrated by communist spies. Detailed information is being collected on the private life of everyone who has ever worked for Scotland Yard, including private collaborators. Why, even Holmes is of interest to the director of the Agency!

In (12), "even Holmes is of interest to the director" of the Agency is false, since he did not actually work for Scotland Yard. However, as per the MCR thesis, the content  $\langle \text{Holmes, being of interest to the director of the CIA} \rangle$  would be true just in case the CIA director is interested in the literary character. This too is an undesirable outcome.

## 6.1 Predelli's Solution

Predelli tries to solve these issues by restricting the domain of the quantifier. Consider the following statement made in the context of a philosophy department picnic (*ibid.* 272):

(12) No professor talked to any student.

This is true even if any other professor outside of this context, spoke to any other student in Germany or any other part of the world. Predelli mentions that this restriction of the quantifier is compatible with three approaches to quantifier domain restriction—the ellipsis approach, the indexical approach and the unarticulated constituent approach. Under the ellipsis approach, D being a class of relevant individuals, (12) is elliptical for “no professor in D spoke to any student in D.” (*ibid.*, 273) On the indexical approach, the quantifier expression is an indexical whose “semantic value is sensitive to an appropriate contextual parameter.” (*ibid.*, 273) Thus, where D is a “subclass of relevant individuals”, the semantic value of “no professor” in the context “is a complex item pertaining only to individuals in D who are professors.” (*ibid.*, 273) In the unarticulated constituent approach, there is no implicit or explicit expression that shows that D is the semantic value but the content has D nevertheless.

However, the restriction of the domain of the quantifier is not sufficient to account for the scenario in which one of the professors, Smith, who is also a student of art history, speaks to another professor. This scenario does not make (12) false. Smith is an individual who is a member of the class of individuals who are professors in the setting. It is this aspect which is salient in the context and not that of him being a student. Likewise, qua students too, only a specific kind of student is picked out in this context. Thus, the ellipsis approach will analyse (12) as short for “no professor talked to any student of the philosophy department.” (*ibid.*, 274) In the indexical approach, “is a student” will be a predicate that is indexical and contributes the relevant class of students salient in the context. The unarticulated constituent approach includes in the content expressed by (12) the relevant class of students.

Predelli makes the case that “exists” works like the predicate “is a student”, wherein when (1) is used in (11), “it must be interpreted as a comment on the individuals endowed of the kind of existence relevant to the topic at hand.” (*ibid.*, 276) His rationale for this is that even if the semantic contribution of existence is:

the property an item  $x$  has if and only if, for some item  $x=y$ ”, it “does not entail that existence, or the class of things endowed of existence, always provides the semantic contribution offered by occurrences of ‘exist’ or cognate expressions. Consequently, it does not entail that utterances of a sentence of the form ‘ $n$  exists’ must be true whenever the item contributed by  $n$  has existence, i.e., whenever it is something” (*ibid.*, 275-276)

As a result, the semantic contribution of “Holmes” and “exist” may not coincide. Thus, an utterance of “Holmes does not exist” is true only if the semantic contribution of “Holmes” is not a member of the class associated with the predicate “exist”. (*ibid.*, 276) Predelli says that this is consistent with a sense of exists in the Oxford English Dictionary meaning “to have being...under specified conditions.” (*ibid.*, 276, n. 12)

Hence, the ellipsis approach would say that in (11), (1) is elliptical for “Holmes does not exist as a flesh and blood detective”. (*ibid.*, 276) The indexical approach would say that “exist” would have the subclass of items in the domain which in the context are human beings. In the unarticulated constituent approach, the content has an item “pertaining to the property of being a flesh and blood individual”. (*ibid.*, 276) Even in (12), <Holmes, nonexistence> comes out as true because of the salience of the property *being a literary character* in the context, rather than *being a flesh and blood human being* because the circumstance determined by context has it as such.

## 6.2 My Objections to Predelli

The first objection to Predelli would be to his appeal to circumstances determined by fictional contexts. As opposed to factual circumstances, let’s call them “fictional circumstances”. It is over these circumstances that the domain of the quantifier is being restricted and where relevant properties become salient. Given that he is not a modal realist, the view that concrete possible worlds exist, he would have to think of a fictional circumstance as an abstract entity, since it is not a concrete event. This however is a category mistake, since circumstances transpire or happen. Predicates like “transpire” and “happen” can literally only be applied to concrete events in time and space (which would include mental states), since they presuppose temporal concepts (this objection is like the one Yagisawa made against van Inwagen). On the same grounds, only spatiotemporal entities and not fiction determine literal circumstances. The burden is on Predelli to show how it can be applied to an abstract entity, unless it’s a conjuring trick in which philosophers stipulate a philosophical term of art for their purposes without explaining why or how.

Of course, he might reply that pretheoretically, one would think that “Tyrion Lannister murdered his father”, is a fictional circumstance which happened. But given that it did not literally happen, and it would be a category mistake to say so, it seems more accurate to say that it is depicted. If it is depicted and did not really happen, then it is not really a circumstance but represented as one.

It is important to clarify this before what follows, because Everett (2007) makes a general objection to this type of realist view. In particular, he maintains that when we make negative existential claims like (1), we are making a claim about the world which is not tethered to any particular context: “we typically judge utterances of fictional negative existentials to be true independently of what we know about the contexts in which they are uttered” (Everett 2007: 62).<sup>23</sup> (This objection also makes more sense now, given my objection about fictional circumstances not being literal.) Thus, if there are no such abstract circumstances, then there is no option but to utter (1) as a statement about the world, contrary to Predelli.

But to this, Predelli would say that there is such a statement made in the fiction. The best way to explain its truth is to say that it is true in the pretend circumstances. As I have argued against Solodkoff in Chapter 2 that since our ordinary concept of a “thing” or “object” is that of a concrete object and it is a category mistake to literally say that there is such a thing as a pretend circumstance. At best, it is an ontological metaphor. As I have said before about any fictional entity, including a fictional circumstance, that any statement about it is about an abstract entity that is likened to a concrete circumstance.

Thus, since there are no such literal fictional circumstances making fictional statements true, it makes sense that whether the statement (1) is made in a fictional work or made about the actual world are both about the world. This would imply that there is no restriction of the domain of the quantifier either with respect to the fictional circumstance, since there is no such circumstance.

Furthermore, Reimer (2001: 503) objects to Salmon by pointing out a counterexample to sentences like (13) which Salmon uses to account for the existence of fictional entities.

---

<sup>23</sup> Muller(2012) uses this to object to Predelli’s account.

(13) There is a fictional detective who (according to the Conan Doyle stories) is addicted to cocaine.

The counterexample is that we also say things like (*ibid.*, 503):

(14) There is a fictional detective, named ‘Sherlock Holmes’, but he does not exist—he is merely fictional after all.

How then can Salmon or Predelli rescue their theory given a sentence like (14), especially given that the quantifier only ranges over the actual world?

Predelli might paraphrase (14) as “there is a fictional detective, but he is not among the flesh-and-blood objects.” However, this again violates a correct application principle like (CA), since literally speaking, “is a detective” can only be applied to a concrete object. (CA) A predicate that correctly ascribes a property that can belong only to a concrete individual can be literally predicated only of a concrete individual.

To look for an alternative solution, let us look at an example. An article in the Economist (Greene, 2014) argued that what is generally thought of as “mansplaining” defined as “(of a man) to comment on or explain something to a woman in a condescending, overconfident, and often inaccurate or oversimplified manner” (dictionary.com) is not really a sexist phenomenon in which men do it to women due to patriarchal dominance but is merely an assertion of male status through dominance of conversation which men do not only to women but to other men as well. The linguist Deborah Tannen explains this phenomenon as a particular way in which men get status on one hand, and how women gain status on the other by means connections and networks.

One may agree or disagree with this sort of analysis, but if I agree, I might just claim to someone, “Mansplaining does not exist!” In reply, the interlocutor who does not know the entire explanation yet, would tell me, referring to the phenomenon in which men interrupt women, “Of course not, there’s this particular phenomenon and it exists.” The reply would be “The particular phenomenon that you think of as mansplaining is not really mansplaining but something else. Thus there is no such thing as mansplaining” and I proceed with my explanation. The word ‘exist’ functions here to deny of something that it has a particular nature (a person might say “Communism never existed”, by which he would mean “true



Communism never existed”) rather than to categorically deny the existence of that phenomenon itself.

Now let’s take this analogy and apply it to (14). To say in the same breath that “Holmes” does not exist and that he is a fictional character is merely to say, “what you thought of as an object in the external world is not really an object in the external world”. This is what explains Thomasson’s (1999) intuition that when we say that “Holmes does not exist” what we are really saying is that “Holmes is not a real person”. But her account invokes domain restriction through restricted quantification over the domain of real persons and in Predelli’s case to factual circumstances and fictional circumstances.

However, Walton (2003) argues against accounts like Thomasson (which would be applicable to Predelli too) that while “there is/are no” locutions have their domains restricted in contexts like a zoo where someone says “There are no Cheetah’s” even though they exist elsewhere, it is very odd to say “Cheetahs don’t exist” in the same context. This is because we understand the domain as unrestricted. As per Walton, if “Charlie” is a Cheetah’s name and one says “Charlie doesn’t exist” (*ibid.*, 241) in the zoo, one is not merely saying that Charlie does not exist in the context of the zoo, but he is not there anywhere (similarly with “There exist no Cheetahs”).

I have already argued that the statement being fictional is not relevant to the semantics. Even in this case, given there is a problem with fictional circumstances as such, even fictional circumstances should not literally figure in the semantics. In my solution, all statements in and outside fiction about fictional entities involve unrestricted quantification in the presupposition.

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

So what are we doing when we make a negative existential statement like “Holmes does not exist”? Normally whenever we think about Holmes, we are thinking about an ontological metaphor, and as Lakoff and Johnson observe, we are mostly unaware of them explicitly as metaphors. Even though we are not aware, we tend to be aware that that it has a different nature from a concrete object which explains Thomasson’s example “there is no (real) person such as Lear”. But unlike her, what I claim is that this sentence as well as a negative existential about Holmes presupposes (MET<sub>FIC</sub>)\* with a substituted variable, and merely

says minimally, like the mansplaining case, one is denying that it has a particular nature (namely that it is a concrete object in the external world). Similarly, when I say that “fictional circumstances do not exist”, I am saying that it lacks the nature of being a concrete circumstance in the external world and yet, one can quantify over such entities. Since they are mentally constructed out of bundles of mental representations, they are not literally circumstances. Furthermore, by means of the “likened to” relation to concrete circumstances that involves a metaphor, we obtain what Predelli thinks of as fictional circumstances.

### 6.3 Why Are Properties Salient

In what follows, I will try to figure out what makes the properties salient in contexts, especially since I reject Predelli’s notion of circumstance, in which properties pertaining to fictional characters become salient. In the absence of a better explanation, I will argue that it is the design stance that is primary for when we think of the fictional character as a fictional character and when that property becomes of salient and intentional stance is primary for when we think of a fictional characters more directly as a flesh and blood person and that property becomes salient (In Chapter 6, I shall elaborate further that it is the mind-dependent property being a *thought of being flesh and blood* that is encoded by a mental representation that becomes salient and it appears that an object has it because of ontological metaphor). In the contexts in which Predelli appeals to properties becoming salient, I would ask, what it is that makes the properties salient? One explanation might be that I associate a description with the name. Because I do that, the context makes the description salient to me and conditions fulfilled by the object help me pick out that property and makes it salient.

However, Predelli is giving a Millian account, so having an associated description is out of the question for him. Predelli might reply that it is my knowledge of the context that makes the property salient. In the professor case, it is known through contextual clues and inference to the best explanation that *being a professor* is salient in the context and it is mutually agreed that such and such feature is salient. For instance, a contextual clue might be that x being a professor is more salient for him because one might infer say, that in a context of professors interacting with students, that the general hierarchical structure of the interaction would make the feature [+PROFESSOR] salient versus [+STUDENT]. This is because when one says “No student spoke to a professor” in the same context, it is highly likely that the [+STUDENT] of the professor who is also a student will not be highlighted.

Now, given this knowledge, all professors satisfy the description “professor” and (12) is true. So far so good.

Let’s see how this works for a proper name like “Holmes”, also because there is no such literal circumstance like the one in the professor case (which Predelli takes for granted that there is such a circumstance). I figure out through contextual clues like preceding sentences, the book I am reading etc. that it is a statement in a literary critical context or even in a fictional context, where like in (11) the previous sentences talk of literary characters. However, prior to the property of Holmes being a literary character to be salient, reference must be secured to the mental representation. From what I already know of the context, I make an inference that it is that property that must be salient when we are talking about Holmes, because the context activates the design stance. So, the process is as follows. I see the word “Holmes”. The word secures reference to the mental representation. My knowledge of the context makes the design stance primary. Based on the heuristic, I cognize Holmes primarily in terms of him being a literary character, rather than a flesh and blood person. This is what enables me to think of Holmes primarily as a literary character and makes this feature salient.

I have already argued that it is puzzling for there to be anything more than the representation as the reference when it comes to fictional characters. If that is so, my picture of identifying a property and its salience faces a problem. The representation itself projects Holmes being a literary character as salient. It seems odd to say that I identify this property first and infer that it is salient. Or that it becomes salient to me because of my inference, since even before making the inference, one has to refer to the representation which already has the salience as its feature. There has to be some other explanation for the salience in the representation.

My explanation for the salience (which I have given in the chapter 3) is the design stance as mentioned above and well as inference from surrounding context such as surrounding words, my knowledge of the world etc. (I’ll expand on this in my chapter on Walton). If one holds the design stance already<sup>24</sup>, the representation at any time will enable one to cognize

---

<sup>24</sup> The design stance can be triggered in the case of written or spoken fiction by the presence of verbal cues like “fictional character” or “literary character”, where the semantic feature “-ANIMATE”, that is, inanimate, is present in the adjective “fictional” and “literary”. (“Animate” is “a term in the grammatical classification of words (especially nouns) to refer to a subclass whose reference is to persons and animals, as opposed to **inanimate** entities and concepts.” (Crystal 2008: 26)). The case of “character” in the phrase is more interesting. I have argued in my chapter on van Inwagen that in the case of characters like Holmes, x is a

Holmes as a literary character and that feature becomes salient. This also explains why in (11) Holmes being a literary character is salient despite the sentence being in a fictional context, which ideally takes any character as primarily flesh and blood.

#### 6.4 The Problem with Negative Existentials

One thing that is worth asking of Predelli's account is why it is that  $\langle H, \text{nonexistence} \rangle$  comes out as false in (10). In the light of this, it is also pertinent to ask why it is that statements like (1) always need to be understood in another light (like the truth condition Salmon gives or Thomasson equivalence to "there is no (real) person" in the realist's account?) Why can't (1) be taken at face value? In what follows, I will try to show why there is a problem by looking at the problem of empty names and what sort of solution has to be offered at the level of the nature of the proposition.

Direct reference theories face a major problem with empty names. This is because if the semantic value of a proper name is the individual that it refers to, then empty names like "Vulcan" would be meaningless because they lack semantic value. But they are meaningful. Therefore, the semantic value of a proper name is not its referent. A further problem arises when, if one holds that a sentence with a name expresses a singular proposition and the proposition has no semantic value in the subject slot, then a sentence with an empty name cannot express a proposition, unlike the sentence "Bush is human" that expresses  $\langle \text{Bush}, \text{being-human} \rangle$  (Braun 1993: 450).

An analogous problem that is faced with empty names is faced with Jeff King's (2007) theory of propositions, when he reduces propositions to facts. But the problem is that facts are not the sorts of things that can be false, because facts exist in only "one way" (Speaks 2011). An argument from analogy can be made here. The reason why there is a problem with the negation of facts is that they exist in only one way, whereas propositions can be negated. The reason why there is a problem with so-called empty names is because the objects in the subject slot can exist in only one way, because of which, the meaning of a name can only be given in one way. To clarify, the meaning can be given in only one way,

---

character only if  $x$  is a fictional person. In contrast with some languages where the +ANIMATE feature is distinguished morphologically (*ibid.*, 26), in English it is a semantic feature of the word "person" that it is +ANIMATE. In the context of fiction, "character" would then involve both + and -ANIMATE features derived from "person" and "fictional" respectively. While the +ANIMATE feature would trigger the intentional stance, the combination of both features would indicate that it is the design of fictional characters to involve both + and - ANIMATE features. So in these contexts, even if the intentional stance is activated, and used to predict and understand beliefs and desires, the primacy is of the design stance, since there is the -ANIMATE feature present.

that is, only when an object that it refers to is present. This is why there is a problem with the meaning in the absence of the purported object. That is, while “Holmes” purports to refer to a flesh and blood person, the flesh and blood person is absent. Thus, the reason why even in the case of a realist view, even though an abstract object is present in the object slot, (1) seems to come out true because of the absence of the purported object but false because of the presence of the abstract object. The meaning of the name thus, can be given in two ways. This is the reason why the conflict between its truth and its falsity arises.

Therefore, instead of propositions as facts for King and meaning slots as objects, one needs to look for some kind of entity that has at least two modes of existence. For Speaks, the best candidates are properties. The proposition is a property like *being such that grass is green*. If the property is instantiated, then a statement is true, if it is not instantiated, then it is false. Analogously, but even more radical, is my proposal that the content in the name slot in the empty names case can be seen as mind-dependent properties. What would be required in order to avoid the above problems is a direct reference view in which mental representations are the direct referents of fictional names and mind-dependent properties that encode and provide the content (Everett and Schroeder (2015) too account for fictional reference in terms of reference to mental representations). This is the reason why “the meaning can be given in two ways”. Once as referring to mental representation in which the content is given in terms of mind-dependent properties (as in not taken as a concrete object in the external world as seen in sentences like “You didn’t see a unicorn. It’s in your head”) and another time, since facts about reference and aboutness diverge, as a flesh and blood person, since as I have already said, the mental representation is taken in terms of an ontological metaphor.

For fictional names and so-called empty names, the names refer to mental representations. For names which refer to concrete objects, I leave it open whether Speaks’ theory is correct. The point of my argument was in the analogy with Speaks’ argument which can be taken as a plausible one, but the truth of his theory not being relevant to my argument (even though the plausibility of his argument is). As for the case of referring to mental representations, I shall elaborate on this issue in Chapter 6.

## Chapter 5

### Walton's Pretense Theory and its Discontents

#### 1. A Brief Overview

Kendall Walton (1990) argues that fictional statements are true only in a game of pretense and reference to characters is pretend reference. His picture is that fictional works are props in a game of make believe just in the way that a stick is a prop when a child counts it as a bear in a game of make believe. In Walton's view, actual objects play various roles with respect to the imagination. Walton says that "they *prompt* imaginings; they are *objects* of imaginings; and they *generate fictional truths*." (*ibid.*, 21) Props aid in collective imagining and enable people to coordinate their imaginings so that each participant is imagining the same things as others are. They also enable people to imagine more than what they would have otherwise if they were imagining alone (without the props). They are also objects of the imagination in the way one imagines of the stump that it is a bear; although this is not always so, since a faucet can randomly prompt one to imagine a trip to Italy but is not the object of imagining. Similarly, actors in a theater are props and objects of imagining.

For Walton, representational works of art as well as novels are props. In the case of painting, the pattern on the canvas (that is, "the painting itself") "makes it fictional in *La Grande Jatte* that couple is strolling in the park." (*ibid.*, 38) For novels, the words make it fictional in, say, *Gulliver's Travels* "that there is a society of six-inch tall people who go to war over how eggs are to be broken." (*ibid.*, 38)

The rules of the game of make believe (in the case of the bear, it can be rules like 'if there is a stump then it is a bear', also called principles of generation) authorize some types of make believe and exclude others. One can seriously convey the content of *Sherlock Holmes*, like saying that "Sherlock Holmes lives in 221B Baker Street" because it is appropriate to do so as one authorized by the Arthur Conan Doyle's works in the "official" game of make believe. This explains the apparent truth of such statements. While the stump case has an explicit principle of generation which is a stipulation like "Let's say the stumps are bears" (*ibid.*, 38), Walton says that most are not explicit and he does not:

assume the principles of generation are in general, or even normally "conventional" or "arbitrary", nor that they must be learned. Nevertheless, what principles of generation there are depends on which ones people accept in various contexts. The principles that are in force are those that are understood, at least implicitly, to be in force." (*ibid.*, 38)

For Walton then, fictional truths occur when there is a “prescription or mandate in some context to imagine something.” (*ibid.* 39) However, Walton rejects the claim that fictional truth ought to be a species of truth. For instance, “In the world of the Unicorn Tapestries, a unicorn was captured” (*ibid.*, 40-41), there is no real realm or place in the world where a unicorn was captured. Thus, for Walton, the statement “A unicorn was captured” is merely not true. However, one can also have fictional statements which are true, which are real world truths and are fictionally true in novels. Thus, for Walton, it is better to say that fictional truths are what is appropriate.

The advantage of this picture where props generate fictional truths, Walton claims, is that it gives an air of objectivity to the fictional world, where even if we are not aware of the fictional truths, or experience them or even be wrong about them, they still hold. This explains why participants are surprised when it is fictional that there is a bear in the thicket.

As Walton says:

It is not thinking that makes it so; the prop does. Fictional worlds, like reality, are "out there," to be investigated and explored if we choose and to the extent that we are able. To dismiss them as "figments of people's imaginations" would be to insult and underestimate them. (*ibid.*, 42)

Walton also distinguishes between work worlds and game worlds. Suppose Richard is an appreciator looking at *La Grande Jatte*. That there is a couple strolling in the park is fictional in his game, but it is also fictional that he sees them. Richard can even decide to see the patterns in the painting and imagine a hippopotamus wallowing in the mud as part of his unauthorized game. The distinction is invoked so that one does not have to decide arbitrarily which of the various appreciator’s games “is to be identified with the world of *La Grande Jatte*.” (*ibid.*, 59) Thus, for Walton, that there is a couple strolling in the park is fictional in every authorized game for the painting and is thus fictional in the work world. The work world comprises every proposition that is fictional in authorized games for a particular work. The game world, on the other hand comprises what is fictional in the games of the participants. For example, that Richard sees a couple strolling in the park is fictional in his game world, but not in the work world. (*ibid.*, 59-60)

One of the motivations for realist accounts that Walton mentions is to give a uniform account of *de re* statements in fiction like “Don Quixote mistook windmills for giants”. These statements are on the lines of non-fictional statements like “Julius Caesar was warned of the Ides of March by a seer.” (*ibid.*, 131) Now if the fictional statement on Don Quixote is paraphrased to remove reference to a fictional object, the paraphrase might look like “It is Don Quixote-fictional that there is someone named “Don Quixote”, who mistook windmills for giants.” (*ibid.*, 131)

However, Walton says that suppose *the* Don Quixote, who is the central character, does not in fact mistake windmills for giants but someone in the novel talks about a remote ancestor of his by the same name who does so. This not only makes the paraphrase which is about the central character false, but it also prevents us from differentiating the two without having to pick out a character (which would invoke realism). Moreover, Walton (*ibid.*, 132) says that “Don Quixote” is merely mentioned in the paraphrase but not used like in the original.

Walton then combines the virtue of the anti-realist paraphrase view to *prima facie* reduce the need for fictional entities by saying that *de re* fictional statements are fictional only in the game world (not in a work world) of a person, where one imagines seeing or knowing a fish or Don Quixote “from the inside.” (*ibid.*, 136) The lesson he draws from the paraphrase is that what is fictional in the story are just the *de dicto* statements “that there is something or are things-ghosts or fish-of a certain sort.” (*ibid.*, 135)

Moreover, suppose Stephen was pointing at the painting *Shore at Sheveningen* and said “this is a ship”, one may correct him, saying “this is not really a ship”. (*ibid.*, 219) Since he is not pointing to an actual ship, he is merely pretending to refer and pretending to make a claim. (*ibid.*, 219) It is also fictional that he is expressing a truth and “that is a ship” does not express a proposition but it is merely pretended that it does (it is also a pretend assertion). Analogously, since one cannot point at a ship by pointing at the text of *Moby Dick* the same way one cannot point to a ship in a painting, “there is no other reasonable candidate for the referent of “that””. (*ibid.*, 219)

However, we also make assertions like “Tom Sawyer attended his own funeral” (*ibid.*, 396), for example, by a critic while discussing Tom Sawyer. It is statements like these which can be easily prefixed by “in the fiction” clauses. This can be paraphrased as “The Adventures of Tom Sawyer is such that one who engages in pretense of kind *k* in a game authorized for it makes it fictional of himself in the game that he speaks truly.” (*ibid.*, 400)

Thus, if one engages in a pretense of kind *k*, that is, the pretense one exemplifies in uttering “Tom Sawyer attended his own funeral”, one makes it fictional of themselves that they speak truly. There is no description that can individuate it, and its reference is fixed by pointing to an instance of it in a Kripkean fashion, the way the reference of natural kinds is fixed by pointing to instances of water. Moreover, when one makes these assertions, one also simultaneously says that it is appropriate to engage in such a pretense. As for statements that appear to entail ontological commitment to fictional characters because they cannot be prefixed by the “in the fiction” operator, like “Robinson Crusoe was more resourceful than Gulliver” (*ibid.*, 407), Walton accounts for them by saying that in cases like



these, one engages in unofficial games of make believe, where both novels are props. Thus, one “is asserting something true if and only if it is fictional in” the unofficial game that one “speaks truly”. (*ibid.*, 408) Furthermore, a statement like “Gregor Samsa is a (purely fictional) character” is explained by Walton as a “betrayal” of the pretense, in which one is specifying that “Gregor Samsa” does not really pick out anything and is thus merely fictional. (*ibid.*, 422)

Walton’s starting point is the worry that theories that posit fictional entities are bound to reinterpret our pretheoretical commitment to people and things in the fiction. He expresses his dissatisfaction with paraphrase theories since they fail to explain why the particular sentence that is paraphrased is uttered the way it is by us, which for him are more basic than a paraphrase. For instance, given that we are not really talking about a person referred to in “King Lear has three daughters” and attributing him the property, Walton says that we need to understand “why we say it in the potentially misleading way we do.” (*ibid.*, 388) Moreover, when we say that there is no such thing as Grendel, why do we seem to be disguising our statements about what kind of things they are (fictional or nonexistent) by a simple denial of their existence? Another of Walton’s motivations is the thought that realist fictional theories add pretense as an afterthought, whereas for Walton, it looks as if we are so entrenched in make believe that we mistake pretending to refer as genuine reference.

The finer details of Walton’s account will be made clear in the course of this chapter. But more importantly, among other points, there are two key points that will be brought out in this chapter. In the course of the discussion of intersubjective identification, I shall bring out the point that it is interest relative identity for fictional characters (which I introduced in Chapter 1 and will further defend in Chapter 6), is determined by contextual rules and conventions, that allows for intersubjective identification (that is, how two or more people talk about the same fictional character). The other important point that will be brought out will be using Derek Matravers’(2014) arguments against Walton, that fictionality and pretense are irrelevant to semantics and therefore to ontology, since what is basic is not the fiction-non-fiction distinction but a more fundamental distinction.

## **2. Friend’s Objections to Walton**

The first objection Stacie Friend (2007) outlines in her survey article on fictional characters is something that Walton himself points out. More specifically, she claims that, on Walton’s account, even though the truth of a statement like “I pity Anna Karenina” can be accounted for by seeing it as an extension of a pretense, “Holmes is a fictional character” still comes

out as true. Walton's solution to this issue, as we saw above, is to say that sentences like these are parts of unofficial games where we pretend that the world has "real" as well as "fictional entities". The other solution is that Walton sees this as a betrayal of the pretense in which such a person is considered actual. The realist objection to this is that this is an ad hoc move.

The most important concern that Friend (2007) outlines is that Walton faces a difficulty about individuating the kinds of pretense that he appeals to. For a statement like "Isabel married Gilbert" (*ibid.*, 145) which expresses the gappy (incomplete, with no referent for an empty name) proposition "x married y" (*ibid.*, 145), it is a constant issue with antirealist theories that they cannot distinguish between "Romeo married Juliet" or "Isabel married Isabel". (*ibid.*, 145) Therefore, there is no way one can distinguish between thoughts about one fictional character versus another, unless Walton invokes fictional objects themselves to do so. Moreover, individuation cannot be done by appealing to types of names, since names of the same fictional character can be different in different languages.

Moreover, if he says that the content of descriptions connected with the name distinguishes characters, there is a problem. This is because two different characters can have the same descriptions like with Cervantes' and Menard's Quixote or there are two separate characters when in fact they are one and the same, like Odysseus in *The Odyssey* and *The Aeneid*.

To this, Walton does say that pointing to instances of the pretense, connected to a particular work (Friend 2007: 146, Walton *ibid.*, 402) is what individuates kinds of pretenses (Isabel-directed pretense, Gilbert-directed pretense etc). This too is insufficient, as Friend points out that the novel is a prop that prescribes imagining many different characters and merely appealing to this relation to the use of a fictional name to the novel will not individuate a particular character. If we appeal to "syntactic name types", then that is inadequate too, since the syntactic name types are uninterpreted and "it is only under interpretation that certain parts of the text count as about the character." (*ibid.*, 146) The moment semantic interpretation is conceded, then the question still remains as to how we can distinguish different fictional characters in different kinds of pretense without appealing to fictional characters.

## **2.1 The Problem of Intersubjective Identification**

In this section, I will take up Friend's criticism of the individuation of fictional characters in Walton, specifically on how two people can talk about the same character and her own

modification of the pretense account in the light of that. The aim of this section is to bring out aspects of the dialectic on this issue, so as to offer my own solution to the problem. In the course of this, I shall also take up Garcia-Carpintero's (2018) criticism of Friend's account, his own account within the framework of pretense theory as well as offer my own solution at the end.

### **2.1.1 Friend's Objection to Walton**

This problem is further elaborated by Friend (2011) as follows. She points out that when Nabokov says (1) and Smith the critic says (2), they are disagreeing and making contradictory claims about Kafka's *Metamorphosis* (*ibid.*, 187):

- 1) Gregor Samsa has been changed into a beetle.
- 2) No, Gregor Samsa has been changed into a cockroach.

Any theory of fiction needs to explain what accounts for them speaking about the same thing, that is, account for intersubjective identification. With respect to Walton's point about pointing to instances of the pretense for individuation both Nabokov and Smith claim that *The Metamorphosis* requires two different kinds of pretense. The question still remains as to what makes the two cases the same "Gregor directed pretense".

### **2.1.2 Friend's Account of Intersubjective Identification for Fictional Characters**

Friend then gives an account of it not in terms of what one is to imagine (in terms of propositional content) but how one is to imagine, providing "a framework for specifying appropriate ways of imagining." (*ibid.*, 184)

- 3) London is the chief city of a fascist state.

The above sentence (*ibid.*, 191) is true so long as 1984 "prescribes imagining that London is the chief city of a fascist state." (*ibid.* 191) In this, there is a prescription to imagine the singular proposition (*ibid.*, 192):

(L) <London, being-the chief-city-of-a-fascist-state>

While (L) shows the content we are to imagine, the specific way we ought to respond to 1984 needs to be explained. Friend lays out two criteria for the reader to satisfy the prescription to imagine. Firstly, it requires not just imagining (L) but imagining (L) "because so imagining is prescribed by 1984" (*ibid.* 193) (this eliminates accidental

imaginings, say one thinks of the horrors of Thatcher’s London when they hear the title). Secondly, to satisfy the prescription to imagine, the reader has to either be in a particular mental state or have a propositional attitude. As she says: “I assume that a person counts as imagining that London is the chief city of a fascist state so long as she stands in the relation constitutive of imagining, to a mental representation that means that London is the chief city of a fascist state.” (*ibid.*, 193)

Friend calls the structured mental representation<sup>25</sup> in imagining *fancy* (due to lack of an English noun), the mental representation of an individual—*notion* (connected with mental file folders with ideas) and the mental representation of a property or relation-- *idea*.

The reader of 1984’s fancy thus has content (*ibid.*, 194):

<<London-notion, idea-of-being-the-chief-city-of-a-fascist-state>>

Metamorphosis’ prescribing is to be seen with the content (*ibid.*, 197):

(1’) <\_\_\_\_\_ (*Gregor-notion*), having-been-changed-into-a-beetle>

(2’) <\_\_\_\_\_ (*Gregor-notion*), having-been-changed-into-a-cockroach>

The subscript specifies a way of imagining without there being an actual object. The aboutness of the Gregor-notion being a notion of Gregor is explained by appealing to Gareth Evans’ idea that the reference is secured by being causally linked to the dominant source of information. Friend appeals to notion networks, where for example, when Pierre uses “Londres” and Frank uses “London”, both are thinking about London because their notions are embedded in notion networks which support the “naming conventions” that are rooted in London. The advantage she has over name centric approaches (where chains of name using are used to account for how reference is secured) is that it accounts for co-identification of characters with different names like Clark Kent/Superman, for that of nameless characters (Garcia-Carpintero 2018: 11), and for co-identification of the name of the same character in

---

<sup>25</sup> On mental representations, Friend says:

The assumption that propositional attitudes can be understood as involving structured mental representations is not uncontroversial, but in my view it is the best explanation of many features of propositional attitudes; see, e.g., Fodor (1987,1990). Further, research on narrative comprehension in cognitive psychology and linguistics supports the hypothesis that we create and update mental representations of characters while reading (see, e.g., the papers in Goldman et al. 1999). I say nothing here about the relations that constitute attitude-types, such as imagining, believing, desiring, etc., though I assume they are distinguished functionally. (Friend 2011: 193, n. 19)

different languages. For fictional characters, the object-directedness of the imagining is possible by the following steps:

- a) Nabokov opens a mental file when he reads *Metamorphosis*.
- b) This is what guides him when he uses the name.
- c) Nabokov is a participant of a notion network which connects his use of the name to Kafka's "writing the story". There is no further relation of story to the root.

Thus, (1) is true iff the *Metamorphosis* prescribes imagining (1'). But how do we know our imagining is of Gregor, that is, "how do we know notions are embedded in the appropriate network" (*ibid.*, 200)? We don't because the content by itself has no indicator. Friend says:

If this answer worries you, ask yourself how you know which real individuals you are thinking about. There is nothing in the explicitly entertained content of a thought like "London is in England" that guarantees that your thought is directed in the right way at London itself. You are presumably aware of the descriptions and perceptions you associate with London—the ideas collected in your mental file—but if descriptivism is wrong, your thought is not about London in virtue of satisfying these. What matters is that the London-notion associated with these ideas is embedded in a notion network rooted in the city. Most of the facts that determine singular reference for thought and discourse are simply not "in the head." Nonetheless, some of these facts can be used to specify prescriptions to imagine, even about fictional characters. (*ibid.*, 200)

The problem for Friend, that she herself identifies, however, is that the root at the end of the network cannot be found in cases like "Gregor Samsa", unlike London. Her solution is that the root is Kafka's writing it, and in writing it, his introducing a new naming convention in which he associates ideas with Gregor. Thus, when Kafka describes Gregor, the descriptions are what are the dominant source of information in the notion network. The problem about the lack of a referent still remains, but that is a general problem for non-descriptivist theories. Her suggestion is to appeal to psychological processes involved in 'forming notions of individuals'.

Moreover, for her:

notion networks can be identified deictically, by reference to embedded representations. Even if you do not know whether or not the name 'Troy' refers, you do know that you are using the term 'Troy' to identify the same place as Homer's poem. So the Troy-network just is the practice that connects you to The Iliad. (*ibid.*, 201)

### 2.1.3 Garcia-Carpintero's Objections to Friend

Garcia-Carpintero (2018) formulates Friend's view thus:

(CIF) Two attitudes have a common focus if and only if the dominant source\* of information in the files they deploy is a single producer's notion in the network. (ibid., 11)

One problem he identifies with it is that characters can be created by multiple authors. His counterexamples undermine  $CI_F$  strongly. One case he gives from Frege is of two mountaineers who separately see the same mountain, and the locals tell the first one who sees it from the north horizon that it is named "Aphla" and writes this in his diary that it is 5000 metres and specifies its position in his map. The other one does the same, except that he sees it from the southern horizon and finds out that it is called "Ateb". Thus there is co-identification despite two different chains, making the condition about a single producer's notion in the network not a necessary condition for co-identification. This is so even in cases of empty names like 'Vulcan', Garcia-Carpintero invokes Edelberg (1992) in which two teams of scientists independently come up with the same posit 'Vulcan'. Another counterexample from the fictional case he gives is from Everett (2013), where two people who watch Hamlet don't realize that Ophelia is Polonius' daughter and yet co-identify Ophelia's mother and Polonius' wife as the focus of their thoughts (there is no mention of her in the play). Thus, "convergence in a single individual of the chains leading to the two representations" (*ibid.*, 12) is not a necessary condition for co-identification.

The counterexample Garcia-Carpintero presents to the right hand side of the biconditional being a sufficient condition is this. Suppose an author of a novel N, Alice, uses a file of herself to create two characters, Alex and Adrian. Now one reader A thinks Alex is gloomy and another B thinks Adrian is fun. The problem is that since the information in the "Alex" file and "Adrian" file have an identical single source,  $CI_F$  would legitimize the truth of (*ibid.*, 13):

(4) A thinks Alice/someone is gloomy, but B thinks she/the same person is fun.

But an account like (5) is more accurate (*ibid.*, 13):

(5) A thinks Alice – as represented by Alex/modeled as Alex in N – is gloomy, but B believes she – as represented by Adrian/modeled as Adrian in N – is fun.

Garcia-Carpintero says that Friend herself deals with this issue by the following example.

Thomas Kyd based his Hamlet on a Norse legend about the (possibly real) Prince Amleth of Denmark. Suppose that Amleth was real, and that Kyd intended to use 'Hamlet' to identify the same individual, if any, represented in the Norse legend; perhaps the English name is just a variation on the Norse one. And suppose (just

for the sake of argument) that Shakespeare, mistakenly assuming Hamlet to be invented or perhaps not caring one way or the other, intended to use the name 'Hamlet' to identify the same character as Kyd. Then it turns out that all this time we who talk about Hamlet are referring to a real individual. This looks like the wrong result. (Friend 2014: 316 as quoted by Garcia-Carpintero *ibid.*, 13-14)

As Garcia-Carpintero points out, it looks like there is a Hamlet-notion which includes fictionality not present in the Amleth-notion. Shakespeare then would have “is fictional” in the Hamlet file but not in the Amleth file. Similarly, to deal with the Alice case, there would be “is fictional” in the Alex and Adrian files but not in the Alice one. Thus the co-identification needs to be based on different files. The problem with this that Garcia-Carpintero shows is that there is no reason why Alice cannot contain “is fictional” in her file, especially if she thinks that she thinks she is the fictional characters Alex and Adrian and this makes Alice herself co-identify with them so as to say “Alex—and Adrian--, c’est moi!” (*ibid.*, 15)

Next, Garcia-Carpintero brings up Everett’s (2013) account:

(CIE) If two representations . . . are associated with the same fiction then we will take them to be ‘about’ the same thing if and only if, within the scope of the pretense associated with that fiction, they count as being about the same thing. (*ibid.*, 15):

This deals with the mother of Ophelia/Polonius’ daughter counterexample, since they would be within the scope of the pretense, they are identical. Moreover Nabokov’s and Smith’s use can be traced via pretend use networks to Kafka’s pretend use, which accounts for their aboutness by being traced via networks to Kafka’s pretend use.

Garcia-Carpintero’s objection to this is the same as his objection to Walton. Walton and Everett are Millians, so since the pretense does not include propositional contents, there is no explanation for why the content within the pretense arises. Moreover, like two tiered realist accounts, Everett seems to give an explanation of co-identification in a metafictional claim like (6), but not claims within the fiction.

(6) A critic thinks Samsa/a character of Kafka’s was transformed into a cockroach, but Nabokov believes he/the same character was transformed into a beetle. <sup>26</sup>

---

<sup>26</sup> (6) is a paraphrase by Garcia-Carpintero (2018: 2) of Friend (2011)’s explanada for intersubjective identification in (1) and (2)

Thus, as per Garcia-Carpintero, he “puts the cart before the horse”, since aboutness still needs to be explained. This also fails to account for the above Vulcan case, since they are theoretical posits that failed rather than fictions. One can still count them as within the scope of the pretense, but the cart before the horse objection would still apply—“what does it take for two representations to count, within the pretense, as being ‘about the same thing?’” (*ibid.*, 17)

#### 2.1.4 Garcia-Carpintero’s Account

Garcia-Carpintero’s account is as follows:

(CIGC) If two representations are associated with fiction F, then they co-identify if and only if the rf-presuppositions of F “pick out” the same “object”. (*ibid.*, 22)

“Rf-presuppositions” are reference fixing presuppositions. For instance, the way, unlike the sentence “John stole the camera”, the sentence “it was John who stole the camera” presupposes “someone stole the camera” (*ibid.*, 19). To take another example, “x is hungry” when assigned a value, presupposes the singular proposition expressed by “x is the male picked out by the demonstration associated with ‘he’” (*ibid.*, 20) (where ‘he’ is a Kaplanian character— a function from context to content). This is a semantically triggered presupposition, and pragmatically triggered presuppositions are triggered additionally, and they do the work of “specifying additional features of intended demonstrated referent, perceptually accessible or accessible from previous discourse. The descriptive identification embodied in such presuppositions is “reference-fixing” and not “meaning giving” in Kripke’s sense.” (*ibid.*, 20)

He denies that this is a “reductive descriptivist view” since it involves a referent as the content of presuppositions. There are also semantically triggered presuppositions like “x is called N” for “N is hungry”, where a “naming practice” at the “pre-semantic” level individuates a name which helps pick out the right homonym in a particular context. As I understand this, for instance, if one says “John is hungry”, it triggers the presupposition “x is called John” which connects to a naming practice, which helps select which John one is talking about in that particular context. There are also pragmatically presupposed descriptions like *causing perturbations in Mercury’s orbit* for Vulcan and *the author of the Tractatus* for Wittgenstein. For fiction, since it involves pretend assertions, there are pretend presuppositions.



Thus, the reason why Nabokov and Smith's "Gregor Samsa" co-identify is because their "Samsa" files have rf-presuppositions which are causally linked to via Burge's "quasi anaphoric links" (*ibid.*, 23) to Kafka's text of *Metamorphosis* where in turn there are further rf-presuppositions linked to representations of "Samsa" which pick out the "object".

Intersubjectively, (quasi anaphoric) links are constituted by intentions to use referential devices in accordance with the meaning that expressions have in the usage of the interlocutors on whom one relies. Intrasubjectively, the relations between referential vehicles that mental files are intended to capture constitute such links. (*ibid.*, 3-4)

Thus, despite different descriptions both associate with the name "Samsa", their rf-presupposition picks out the same entity Kafka's pretend dubbing picks out, which in turn is what the text invites one to imagine. Garcia-Carpintero invokes possible world semantics for this (as a useful tool, where he doesn't think that possible worlds are primitive). So what he says is:

To resort again to the possibilist ideology, if any of their representations were about something, it would be the entity picked out by the (pretend) dubbing on which the name 'Gregor Samsa' in Kafka's text relies; then both would pick out this entity." (*ibid.*, 23)

His account is neutral between realism and anti-realism, but he prefers that these "entities" are pretend ones. He also seems to prefer the existential quantifier outside the scope of "thinks" and "believes" in extra fictional statements like (6) and intentional identity statements containing empty terms like (*ibid.*, 3):

(7) Hob believes a witch has blighted Bob's mare, and Nob believes that she (the same witch) killed Cob's sow.

His analysis of this is (*ibid.*, 19)

(8)  $\exists \mu \exists v \exists \alpha (R(\mu, \alpha, \text{Hob}) \ \& \ R(v, \alpha, \text{Nob}) \ \& \ W(\alpha) \ \& \ \text{Hob believes } B(\mu, b) \ \& \ \text{Nob believes } K(v, c)).$

"R" is a relation between belief states and their semantic values, which grounds *de re* ascriptions. "μ" and "v" range over Bob's mare and Cob's sow respectively and "α" ranges

over “objects” corresponding to *de jure* referential “expressions” in language or in thought, i.e., objects *prima facie* determined by them”. (*ibid.* 19) The objects can be concrete or exotic objects, where in the case of exotic “objects”, Garcia-Carpintero says, “are merely pretend, our references to “them” and quantifications over “them” just being a hypostatizing figure of speech.” (*ibid.*, 19)

Garcia-Carpintero’s account is “a non-Millian, descriptivism-friendly view of names and other referential expressions” (*ibid.*, 6) and a version of Yablo’s (2001) fictionalism:

on which the semantic referential apparatus (de jure directly referential expressions such as names and indexicals, quantifiers generalizing over the positions they occupy, expressions for identity) is used metaphorically...deploying the figure of speech called hypostatization (Garcia-Carpintero 2010b). It is a dead, conventionalized kind of metaphor. In contrast with pretense-theoretic fictionalist proposals, on this view utterances in metatextual discourse are thus straightforward assertions with truth-conditions. (*ibid.*, 6)

Thus, his account deals with the mother of Ophelia/Polonius’ wife case as well as with the Clark/Kent case because the rf-presuppositions pick out the worlds of the play which contain the individual. In the Vulcan case, given that both scientific communities which made the posit have the same pragmatic rf-presuppositions, they will pick out the same object in close worlds.

To the sufficiency objection to Friend, Garcia-Carpintero explains that the reason why (5) turns out to be false is because rf-presuppositions that are connected to Alice’s representation of herself in her own file from which she creates Alex and Adrian, in which “Alice” picks out Alice in a possible world. It is thus clear that is not what accounts for the truth of statements like “Alex is gloomy” and “Adrian is fun” but rf-presuppositions triggered by (5), in which “Alex” and “Adrian” pick up separate objects in possible worlds. This is also applicable to the Hamlet/Amleth case.

## **2.2 My Objections to Friend and Garcia-Carpintero on Intersubjective Identification**

If we look at the structure of Friend’s argument above<sup>27</sup> for why we do not know that our imagining is of Gregor via content, we will discover that the conclusion does not follow

---

<sup>27</sup> To repeat for the sake of convenience:

from the premises. The reason why she puts this argument is to show by analogy that if meanings aren't in the head in the case of London and we cannot know that our thoughts are about London by means of the content, then the same applies to fictional names. The following is my reconstruction of Friend's argument:

Premise 1. If the content of our thought satisfies all the wrong descriptions we associate with London, our thought is not about London by virtue of this.<sup>28</sup>

Premise 2. The content of our thought satisfies all the wrong descriptions we associate with London,

Premise 3. Our thought is not about London by virtue of this.

Premise 4: If premise 3 is true, then the explicit "content of a thought like "London is in England"" does not guarantee that our thought is "directed in the right way at London itself".

Conclusion 1. The explicit "content of a thought like "London is in England"" does not guarantee that our thought is "directed in the right way at London itself".

Premise 5: If conclusion 1 is true, then we do not know our imagining is of Gregor, given the content.

Conclusion 2. We do not know our imagining is of Gregor, given the content.

The problem with the argument is that Conclusion 2 does not follow from Conclusion 1 and Premise 5. It does not follow because even if the explicit "content of a thought like "London is in England"" does not guarantee that our thought is "directed in the right way at London itself" (*ibid.*, 200), knowing accurate properties of London can still make it likely that our thought is about London by virtue of a Bayesian inference (based on knowledge of some or socially salient properties that belong to London and knowledge of contextual rules) that it is so<sup>29</sup>, and given the content, we can know that our thought is about London by virtue of

---

If this answer worries you, ask yourself how you know which real individuals you are thinking about. There is nothing in the explicitly entertained content of a thought like "London is in England" that guarantees that your thought is directed in the right way at London itself. You are presumably aware of the descriptions and perceptions you associate with London—the ideas collected in your mental file—but if descriptivism is wrong, your thought is not about London in virtue of satisfying these. What matters is that the London-notion associated with these ideas is embedded in a notion network rooted in the city. Most of the facts that determine singular reference for thought and discourse are simply not "in the head." Nonetheless, some of these facts can be used to specify prescriptions to imagine, even about fictional characters. (*ibid.*, 200)

<sup>28</sup> Friend is not a descriptivist about nonfictional names.

this guess. The point here that undercuts Friend's point that we cannot know that our thought is Gregor is that we can in fact know that it is about Gregor, given that we might be making a Bayesian inference. It also follows that if one knows that it is about Gregor, one's thought is in fact about Gregor by virtue of getting it right as a result of the inference.

In fact, this is on the lines of Chomsky's view on I-Language, that is, the idea that language is a property of the mind rather than the commonsense and mainstream philosophical view as an E-Language which is a shared social object<sup>30</sup>. The objections against Chomsky's views is about how communication can take place without there being a "shared language" or "public language" with "shared reference" and "shared meanings", especially Fodor and Lepore's (1992) problem about how there can be "communication between time slices of an idiolect". This is Chomsky's reply:

It may be that when he listens to Mary speak, Peter proceeds by assuming that she is identical to him, modulo M, some array of modifications that he must work out. Sometimes the task is easy, sometimes hard, sometimes hopeless. To work out M, Peter will use any artifice available to him, though much of the process is doubtless automatic and unreflective.' Having settled on M, Peter will, similarly, use any artifice to construct a "passing theory" - even if M is null. Insofar as Peter succeeds in these tasks, he understands what Mary says as being what he means by his comparable expression. The only (virtually) "shared structure" among humans generally is the initial state of the language faculty. Beyond that we expect to find no more than approximations, as in the case of other natural objects that grow and develop. (Chomsky 2000: 30)

In another case, the author himself, right after having created Gregor can think of Gregor purely by virtue of the content and he obviously knows his imagining is of Gregor. No notion network is required. Moreover, 5 hours later, he forgets his character for a bit, and thinks to himself, "What was the character I thought of that transformed into an insect? Ah! Gregor!" He knows it is Gregor by virtue of inferring from the content. But from this it also follows that the criteria he uses to infer that it is Gregor and determine that his thought is identical to the other person's can also be used specify the criteria to determine intersubjective identity.

One might object here that a person can be talking about something without knowing any property belonging to an object or even getting all of them wrong. But the above case of the Bayesian inference on the basis of said knowledge is just a sufficient condition but not a

---

<sup>30</sup> In Chapter 6, I canvass arguments as to why language and meanings are primarily understood as idiolectal and why the I-language conception is best. If we take Chomsky's arguments seriously, it follows that any conception of E-Language as primitive or it being anything other than an idealization is pseudoscientific and is akin to belief in ghosts. It also follows that much of mainstream analytic philosophy that works on this assumption thrives on pseudoscience.

necessary condition. The point of it was not only to show that a person can know what he or she is thinking of or imagining but also to show that intersubjective identification is made on the basis of such criteria which form the basis for one's Bayesian inference, the correct application of which enable us to judge that someone gets their inference right or wrong. But even then, it is arguable that one is attempting to make a Bayesian inference on other grounds. In the case of us possessing all the wrong descriptions, the probability of us thinking about London by virtue of descriptions is 0, even though we can still be thinking about London, since we might have other contextual clues, like one based on probabilistic calculations via anaphora (like "he" referring to Socrates the previous person is talking about) of what the previous person was talking about. For instance, if one overhears someone talking about "Socrates" without knowing about any description and asks, "So which year was he born, this Socrates?". In this I merely infer, that "he" refers to what the other person is talking about name is "Socrates". Moreover, in a case like this or in any case where one talks about a person by means of his or her proper name without possessing a description, like in the Gödel-Schmidt case below, it can still be argued that one doesn't use descriptive features but indexical features (Jackendoff's (2002) term for a minimum identifiable characteristic presupposed in a sentence like "I don't know what that was, but here it comes again!" (*ibid.*, 204)). Of course, it does not need to be meaning constituting, so one does not fall into descriptivism, but even in the Socrates case above, one has to eliminate polysemy to zero in on the particular Socrates the other speaker is talking about rather like zeroing in on the correct meaning of a polysemous word, like "shot", where it can be a gun shot, shot of vodka etc. through contextual cues. (Miller 1999) So piggybacking on the other person's use of Socrates, one opens up an object file to which further descriptions can be added, on the basis of an indexical feature the hearer has identified and this allows for minimal co-identification of the speaker and the hearer's thought.

It has been seen that both properties that belong to an object as well as contextual rules which form the basis of a correct inference. I will now try to show below how contextual rules figure in the determination of intersubjective identity. The main idea is based on knowledge of contextual cues as well as conventions, people determine the identity of what they are thinking about, especially since the same properties one is thinking of can count as two different objects depending on the context<sup>31</sup>. If what I am saying is right, then it will lead to the correct prediction in the following case, where the intersubjective identification of London is based on context. Suppose Linda thinks she's thinking about "London". She thinks about the Big Ben but all the other properties she thinks about are properties of Paris.

---

<sup>31</sup> I shall expand on this in Chapter 6.

She goes to Joe (who has knowledge about some properties that belong to London) since one can get a few properties wrong and still be thinking about London) and talks to him about a place called "London". She then goes on to tell him about all the properties of Paris as well as about the Big Ben. Now in a context where Joe had brought up the topic of London (saying, such a lovely place, London) and she continued the conversation, Joe would correct her, saying, "You're not thinking about London, you're thinking about Paris". Even if she does not say anything about the Big Ben, Joe is entitled to say the same thing to her, and one explanation is that she merely picked up an indexical feature in her previous conversation with her interlocutor. But suppose Linda goes to a travel agent and tells him, "I want to go to London. I'm really excited about seeing the Big Ben." But then she goes on to list all the properties of Paris. It is within the rights of the travel agent to correct her saying, "Ma'am you're thinking about London, you want to go to London, but you've got the properties wrong. Those are the properties of Paris." In this case, Linda seems to have imported an indexical feature of London with the wrong properties in her object file except one, which is why the travel agent can in fact tell her that she is thinking about London. Furthermore, the travel agent can in fact be wrong. If Linda picked up the connection between London and the Big Ben in an epistemically accidental fashion, say, in a magazine with London and a Big Ben randomly introduced among descriptions about Paris, then the agent has one clue missing on his knowledge of her context to make the right inference that she was thinking of Paris.

This also leads to the right prediction in Kripke's Gödel-Schmidt case. Kripke undercuts descriptivism about names by arguing that if reference is determined by descriptions, then we will get the wrong prediction if Gödel means "the one who proved the incompleteness theorem". If Gödel stole the theorem from a man named Schmidt, then we will end up with the wrong result, where all the while we would have been referring to Schmidt with the name "Gödel". Now suppose someone is talking about the man who proved the incompleteness theorem in a context where it is evident that people don't know it was Schmidt. Based on my knowledge that it was Schmidt who proved the theorem and not Gödel and my knowledge of the context that people associate the proof with Gödel, I figure (perhaps not consciously, via the inference) that he is talking about Gödel and tell him that he should be actually talking about Schmidt. Similarly, if the man is waxing eloquent about Gödel as a genius while using "Gödel" to talk about the man who proved the incompleteness theorem, based on my knowledge of the context that people use "Gödel" to talk about the man who proved the incompleteness theorem and my knowledge of the facts,

I figure he is talking about Gödel and correct him to tell him that he should be talking about Schmidt as a genius.

My account looks circular, since the Bayesian inference will already refer to what the inference is made about. But it needs to be noted that this does not say that referentialism is false and descriptivism is true, at least at the level of E-Language, thought of as an idealization from I-Language. The meaning of “London” is still London. We are asking a different question. What in this case makes our thought about London? I have argued in the previous chapter that reference and aboutness come apart. My point here too might reinforce that. Imagine someone saying all sorts of wrong things about Holmes. Taking on from Pautz (2008) (on which I elaborate more below), suppose one hears the name Holmes from someone, knows one correct description attached (“wears a hat”) and forgets all other descriptions. Then they think of all kinds of other properties and write another novel with the name “Holmes” who is now a superhero. The person then meets a Holmes nerd who happens to be talking about Holmes and talks enthusiastically about how they heard about Holmes and decided to write a novel on him and goes on to elaborate. It is in place for the nerd to ask the person, “Are we talking about the same character?” in some contexts because the person miscalculated or had the wrong descriptions. So in those contexts, it’s either indeterminate or the person would not be talking about Holmes. In other instances, in case someone hasn’t written a new novel, and they just go on to tell a nerd about Holmes, the nerd can correct, saying, “No, that’s not how Holmes is, Holmes is like this”. It seems like the nerd is correcting them based on knowledge of the referent, which determines aboutness. But in this case as well as the Gödel case above, it looks as though reference and aboutness diverge, and conditions on aboutness seem to do more with context and a speaker’s knowledge of context. There is at least one person with knowledge of the properties of the referent<sup>32</sup> as well as knowledge of the context who determines what another person is talking about

If we also look at why Garcia-Carpintero invokes descriptions in reference fixing presuppositions to modify Friend’s account, I have a ready explanation. The reason why he does that is that one needs these descriptions so that one may bet, faced with accurate content, that one’s thought is about London. They are also required to zero in on the correct homonym of words. Thus, contrary to Friend’s saying that “most of the facts that determine

---

<sup>32</sup> This referent, in Chapter 6, is construed as normative and which I argue is the normative basis of meaning. There is also a descriptive referent (descriptive in the sense of descriptive vs normative grammar), which involves reference to a mental representation in an idiolect.

singular reference for thought and discourse are simply not “in the head”, it is plausible to think they are in the head.

Here’s a contrary picture of how it works. Word usage is transmitted via texts and conventions by causal relations. All of them are dependent on people’s minds for being interpreted and for their existence. Without people’s minds, words are just marks on paper, or vibrations in the air. Causal transmission happens only via this. The causal relations merely happen with physical elements. Even if there was no such causal relation, or there was a sudden break, but contextual cues existed, we would be able to determine the identity of the object of our thought. Therefore, what actually determines what I am thinking of, is not this impersonal causal transmission, but that I bet using contextual cues as to the identity of what I am thinking. I think about the same thing as Nabokov because there are contextual cues that give me criteria to determine whether I am thinking about the same thing or not. It is pertinent to mention here that Miller (1999: 14-15) outlines three kinds of contexts from which contextual cues can be gleaned. One is a situational context, is the sort of information needed for deixis as well as understanding the “purposes and goals of the communicative interaction.” (*ibid.*, 14) This information is drawn from one’s knowledge of the people and the world they inhabit. The second is the topical context, which is the specific domain of discourse. For instance, a polysemous word like shot would mean one thing while talking about marksmanship, another to a bartender and another to a doctor. The third is local context, which is the information that is given by words that surround a particular word immediately. Word order and syntactic categories matter here and local context is dependent on them. Thus, one’s knowledge of information in these contexts aid in making inferences about the identity of what one is talking about, and therefore to determine co-identification.

Another circularity objection may arise for my account, since in order for my betting to be right, there has to another fact of the matter about whether it is the same object of thought or not, say something that is determined by the author. But in the presence of criteria given, for instance, such and such content in such and such context, why is there a further fact even required to say one thinks of something as the same? For instance, in the case of money, there is no further fact of the matter about how to count a piece of paper as a particular amount except social criteria like “x is counted a y in context c” in an account like Searle’s (2006). The reason why Garcia-Carpintero’s objection against Friend works and that her account cannot distinguish between Alice and Alex and Adrian is precisely because there



are no such criteria in her account<sup>33</sup>. I shall elaborate on this picture in the section on Pautz below.

There is a further question to be asked of Friend, which can be asked of Walton too. We know which fictional statement is appropriate and which one is not. Suppose I agree with Walton and Friend that yes, Nabokov is right in (1) because it is prescribed, in Friend's case, that it is (1') that is prescribed. Even if we say that most people can know that (1) is appropriate without knowing that it is prescribed, it is plausible to say that at least one person knows that it is true because it is prescribed (Walton and Friend for instance, who appear to know the "true" theory behind why it is fictional). The question is: how does one know that it is prescribed? There is a circularity problem. The prescription to imagine explains the content of one's thought. But one cannot identify what one is prescribed to imagine unless one already knows the content.

Friend's reply would be that since the prescription to imagine makes it fictional because I am in the right relation to a notion network, one does not need to know how whether it is prescribed. It is sufficient that one is in the right notion network. But I have just argued against the thesis that one cannot know that one's imagining is of Gregor, which for her is the same as saying that we do not know "that our notions are embedded in the appropriate network." (*ibid.*, 200) However, confusingly, she also says that her account does presuppose that notion networks can be identified. She says, "notion networks can be identified deictically, by reference to embedded representations. Even if you do not know whether or not the name 'Troy' refers, you do know that you are using the term 'Troy' to identify the same place as Homer's poem. So the Troy-network just is the practice that connects you to *The Iliad*." (*ibid.*, 201) On the face of it, this looks contradictory to what she is said before that we cannot know. But the charitable interpretation is that we do know the right network not via the content but deictically. So Friend's reply to my objection would be that one knows that one is prescribed to imagine that P by deictically pointing to the right notion network.

This too is fraught. In the case of concrete objects one can point to a referent and say "there". While faced with fictional names and "embedded representations" one has to eliminate homonyms first to zero in on the correct name. If this is so, then to even refer deictically using the correct homonym, one has to narrow down possibilities using at least some minimal notion of what it is about, or having a minimal knowledge of its meaning

---

<sup>33</sup> I do not think this falls prey to Kripke's arguments on error and ignorance which apply against Searle's cluster theory, which says that a name is defined by a cluster of some descriptions associated with it.

(what makes this token “John” a token that refers to the John one talks about and not any other), even if one does now possess full knowledge (for instance, whether Troy be a real place). In the case of Troy above, one clue about narrowing down on homonyms might be one’s knowledge that one uses “Troy” to “identify the same place as Homer’s poem” (*ibid.*, 201). This shows that Friend’s account is circular, since one cannot deictically refer to a notion network without already presupposing knowledge of what it is about, at least in a minimal sense.

It is pertinent to observe here that this is one reason why Garcia-Carpintero appeals to presuppositions to narrow down on homonyms. He writes:

when it comes to an utterance of ‘N is hungry’, the proposal has it that the semantically triggered presupposition is x is called N – and hence that there is someone called N. The name is individuated by a specific naming practice picked out at a “pre-semantic” level in order to identify which specific homonym is meant in the relevant context. (Garcia-Carpintero 2018: 20)

Although Garcia-Carpintero might be right on a lot of counts, his invoking of rf-presuppositions in Kafka’s text of *Metamorphosis* through which a “Samsa” representation picks out an object in a possible world looks like an ad hoc move. For one, why should the object in a possible world that satisfies rf-presuppositions in Kafka’s text be the same as the character Kafka is talking about, since there can be more than one object in the same possible world that satisfies the descriptions. The move is ad hoc then, since it just posits an object in a possible world to save his hypothesis even though there’s further explanation required. Why should the descriptions magically pick out one object as against another? One reply that Garcia-Carpintero might give is that the object is a hypostasized dead metaphor that is quantified over. That might be right, but since it is a hypostasized metaphor and not an abstract object which can have independent identity conditions, why should the descriptions pick out one that satisfies them? Moreover, Garcia-Carpintero says that his account is compatible with abstractionism, but abstractionism would make the need for possible worlds superfluous, since abstract objects would have their own identity conditions.

Garcia-Carpintero can of course, appeal to interest relative identity in possible worlds like Lewis does (Lewis 1986). But then, the whole process of the quasi-anaphoric links is rendered redundant because any reader can directly access content given by textual descriptions and based on contextual clues, identify it. This is not very different from what I have been arguing for.

Another problem Garcia-Carpintero faces is Everett's indeterminacy problem, which fictional realists face too. How can picking out an object in a possible world allow for co-identification when what is being picked out itself can be indeterminate? Everett's example is the following:

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. (Everett 2005: 629)

Assuming that no one is sure because of metaphysical indeterminacy, Everett argues that metaphysical indeterminacy leads to the contradiction that a is definitely distinct from b and a is not definitely distinct from b. If this is correct, then there is no determinate "object" to be picked in the possible world Garcia-Carpintero talks about.

According to me, the problem seems to stem from the entire literature, including fictional realism, assuming that there must be some feature that accounts for identity of fictional characters completely outside of our inferential processes and conventions that can be traced back to Kafka introducing a naming convention (intrinsic to the naming convention) or of some "object" (intrinsic to the "object) even if it is a hypostasized metaphor. The problem seems to dissolve the moment one denies such an intrinsic defining source of identity and allows for constitutive rules (like "x is counted as y in context c") and contextual clues that help to zero in on figuring out the identity of what one is thinking of. For instance, the situational context can trigger knowledge of constitutive rules to determine the inference for identity. It helps to see why constitutive rules eliminate the need for an intrinsic identity feature of a fictional character. Searle writes:

regulative rules regulate preexisting forms of behaviour, constitutive rules make possible new forms of behavior. They constitute the phenomena they regulate...the activity of driving exists independently of this rule; the rule regulates an antecedently existing activity. The rules of chess, on the other hand, do not just regulate, but they constitute the activity they regulate. (Searle 2018: 51).

It is because these rules constitute, in this case the identity of fictional characters that they become what determines the identity in particular contexts and allow for co-identification.

Whereas these criteria are public and shareable, acts of dubbing a fictional character and mental content are not publicly shareable the way the criteria are.

### 3. Zemach's Objections to Walton

To recall, for Walton, when I utter the sentence "Tom Sawyer attended his own funeral", I engage in a pretense of kind K. Walton says that there can be no "individuating description" of this kind of pretense. Moreover, since the proposition "Tom Sawyer attended his own funeral" doesn't exist, one cannot even pretend to assert a statement that actually expresses such a proposition. He says, "the reference of K can be fixed by pointing to examples, such as the pretense Sally herself displays...Sally specifies K by indicating, displaying, a single instance of it, her own act of pretense." (Walton 1990: 402 as quoted by Zemach 1998: 170) Walton invokes Kripke's account of natural kinds to account for the particular nature of the pretense. The way the term "water" names the essential nature of water, and not just a sample of water, the term "K" is the name of the essential nature of Sally's behaviour.

For Walton, the reason why a sentence in another language can be a translation of the English sentence at issue, especially since there is no proposition that they both express, is that the speaker of the other language's behavior shares the same essence while uttering that sentence—the K pretense. Since there is no fictional entity that the same name refers to in two different languages ("Tom Sawyer" in English and "Su Quitopo" in Indonesian), Walton would say that they refer in the "same manner" but not to the same entity where we "betray a kind of pretending to refer." (Walton 1990: 424 as quoted by Zemach 1998: 170) Zemach points out that they simply do not refer in the same manner. This is so because they behave in two different ways simply by virtue of uttering "Tom Sawyer" in one language and "Su Quitopo" in another. The two different utterances are sufficient to make the manner different. Moreover, this is not a necessary condition because Walton in even talking (which is a different manner) about Sally's pretense of referring to Tom Sawyer refers to Tom Sawyer in doing so. Now in saying that Sally \*refers to Tom Sawyer\*, Walton utters a meaningless sentence, since it expresses no proposition. Now if Sally utters a meaningless sentence with the name "Tom Sawyer" and Walton says that the Indonesian girl Diyan utters "Su Quitopo" has the same behaviour without referring to anything, that is incorrect, since there is no similarity in their behaviour.

Zemach's objection goes further. Walton merely gives an instance of Diyan's utterance, without giving any evidence like a videotape of her behavior. He did not observe the behaviour of every Indonesian while uttering the same sentence to determine that Diyan

shared the same behaviour to determine that the nature of the K-pretense is the same. Not only is there no way of determining this in order to say they share the same behaviour and therefore have the same K-pretense, all one needs to do is take “Tom Sawyer attended his own funeral” from English to Indonesian without any other behavioural concomitant that is similar to any other person. If that is so, then all one has to do while translating the sentence is to understand it, since translation preserves meaning. Zemach further contends that if that is so, to understand the sentence, one has to understand what the term denotes and know what its referent is. He therefore understands it, not merely quasi understands it or pretend understands it.

In my view, one response to Zemach that Walton can give is that the reason he does not need to observe the behavior of every Indonesian to determine that Diyan’s K-pretense has the same essence as everyone else’s is that the essence is dispositional. Thus, the mental state of the K-pretense is individuated not by the actual behavior of the people but by the tendency to do certain actions. Like what Ryle says of the mental state of “knowing”, similarly the pretense of kind K will be individuated by dispositions, “to say that a person knows something, or aspires to be something, is not to say that he is at a particular moment in process of doing or undergoing anything, but that he is able to do certain things, when the need arises, or that he is prone to do and feel certain things in situations of certain sorts.” (Ryle 1949: 177)

So when a person is in this state, they will say certain things, they will say that they are true, they will translate sentences from one language to another in a certain way etc. Moreover, given that they are pretending, they would not react to their “objects” of pretense the way they would for real people.

The problem with this view is that it commits Walton to ontological behaviourism, the view that mental states just are behaviours and dispositions to behave, at least for propositional attitudes. If he says that this view of mental states extends only to pretend states, it seems arbitrary that he would have this view for one type of mental state for one propositional attitude and think of it as *sui generis* and not for other propositional attitudes. Thus, there is no reason why this view of mental states should not extend for him to other propositional attitudes. He is also in sympathy with Kripke/Putnam causal views of reference, and his view arises out of that framework, or else he could have accepted Fregean senses or some theory with psychologically dependent content, but decided to account for the lack of a referent in terms of pretense. Now, Hauser (2002) in his entry on “Behaviorism” in the “Internet Encyclopedia of Philosophy” writes that:

semantic externalism is the view that "meanings ain't in the head" (Putnam 1975: 227) but depend, rather, on environmental factors; especially on sensory and behavioral intercourse with the referents of the referring thoughts or expressions. If emphasis on the outward or behavioral aspects of thought or intelligence -- and attendant de-emphasis of inward experiential or inner procedural aspects -- is the hallmark of behaviorism, semantic externalism is, on its face, behavioristic (though this is seldom remarked). Hauser (2002)

If this is so, it seems plausible that holding on to the causal theory of reference is consistent with being a behaviourist (since Hauser says "on its face") if not a fully behaviouristic theory, at least in terms of semantics. So it makes sense as to why Walton's view leads to what it does and makes Hauser's statement even more plausible. Given that his behaviorist view would extend to other mental states like belief etc., as well as with respect to meaning, this ought to turn him into a wholesale behaviorist.

Now not only is behaviorism a very controversial view, and is almost discredited by now, Walton also needs to explain why exactly is it that people behave in the way they do for specific content in the case of pretense. This also lines up with Garcia-Carpintero's objection above that Walton needs to explain why the content in the pretense arises and the fact that Friend is bound to talk about Gregor-notions that are mental. The reason why they're forced to do this is precisely because Walton ends up becoming a behaviorist with his view on pretense. Moreover, if he accepts liberal dispositionalism about pretense and belief, the view that beliefs are dispositions to have private mental episodes on pretense as well as behavioral ones, he faces the same problem of individuation and translation that Zemach remarks on. What, at all should be common between mental states of Diya and Sally so as to individuate K pretense? This is the task that remains for pretense theory.

#### **4. Richards's Objection to Walton**

Mark Richard (2000) questions Walton's claim that fictional statements do not say anything. Let's consider (9) (Richard 2013: 172):

(9) Ishmael survived the wreck of the Pequod.

For Walton, as we have seen, the pretend utterance of (9) informs listeners that this kind of pretense is appropriate for readers of the novel *Moby Dick*. The speaker thus makes it fictional of himself that he speaks truly in the authorized game. This is like indicating to people from a foreign culture that it is appropriate to eat snake livers with parrot's nest sauce by doing so. (Walton 1990: 399)

As we have seen, the one advantage of this account is that pretending to refer does not require a referent like Ishmael. However, the other significant advantage of the account

Richard points out is “its (logico-) syntactic conservativeness.” (Richard 2000: 205) This is so since a pretend utterance of (9) does not involve a reference to the novel or the “property of being fictionally true.” There is thus no need for it being elliptical for “it is fictional that” or “according to the fiction”. We take the surface syntax of (9) at face value since we are “saying by doing” and the syntax does not deceive us as to its true logical form.

Thus, given that one pretends to refer in uttering, one conveys that it is appropriate to engage in a kind of pretense, so when they say (9), it is (10) that is being genuinely asserted (the paraphrases that follow are from *ibid.*, 208-209):

(10) If you engage in a pretense of kind K in a game authorized for *Moby Dick*, then you make it true of yourself, in the game that you speak truly.

Analogously, in a sentence where there is no apparent reference to a fictional entity, a description is substituted for the proper name:

(11) Only one man lived to tell of how a white whale sank a ship with a one legged captain. What is genuinely asserted for (11), which expresses a proposition, unlike (9) is:

(12) If you pretend to say that only one man lived to tell of how a white whale sank a ship with a one legged captain, in a game authorized for *Moby Dick*, then you make it true of yourself, that in the game you speak truly.

A closer paraphrase is:

(11.1) It is fictional(ly true) in *Moby Dick* that only one man lived to tell of how a white whale sank a ship with a one legged captain.

It is to be noted that (11) is not an ellipsis of (11.1) because Walton would think that the elision of (9) would be

(9.1) It is fictional in *Moby Dick* that only Ishmael survived the wreck of the *Pequod*.

The statement that is elided for sentence S is of the form “it is fictional in f that S”. The problem then with the ellipsis account is that even though the paraphrase in (11.1), which is an ellipsis of (11) succeeds in saying something, if the sentence S in “it is fictional in f that S” contains a proper name, “the longer statement retains the apparent reference to [Ishmael and the *Pequod*]. (Walton 1990: 397 as quoted by Richard 2000: 210), which it shouldn’t. Since the fictional names are empty, such a sentence cannot “literally say anything”. Thus, Walton’s main motivation for rejecting paraphrase is that utterances like (9) with fictional names do not literally say anything since they have fictional names.

Richard's argument against Walton is as follows. Our pretheoretical intuitions say that one does not distinguish in kind between the following sentences (*ibid.*, 211):

"Snow is white" says that snow is white.

Mary said that the rain in Spain falls mainly on the ground.

In *Moby Dick*, it says that only Ishmael escaped the wreck of Pequod.

Richard says "there will be nothing about your usage or linguistic behavior which would justify thinking that you (or your "language module") perceive the last sentence to differ in some semantically significant way from the first two." (*ibid.*, 211) Even if the rest of Walton's account is true, Richard says that there should be no additional reason that one participates in a pretense to talk about the pretense itself (to say that such and such statement is appropriate) and that it is more parsimonious that one says that (9) and (11) are "true in *Moby Dick*". For Richard, Walton unnecessarily imports a theoretical fact that (9) does not literally say anything. On the face of it, especially given that our pre-theoretical intuitions actually disbelieve that (9) does not literally say anything, why think that this so called "theoretical fact" is even relevant to considering that the fictional operator is not the right analysis for (9) and (11)? "Certainly we can't appeal to it, to explain why the utterer of (1) ((9) here) engages in pretense, for the utterer, as just observed, rejects the theoretical fact." (*ibid.*, 211) Moreover, if (9) is not "semantically defective", that is, "if its literal unembedded use makes a truth evaluable claim," (*ibid.*, 212) then that (9.1) is true is merely explained by the fact that there is a mandate to imagine by the prop that Ishmael survived the wreck. If that is so, then when one pretend asserts (9) and tells the audience that that it is appropriate to the game of pretense, this is merely a more complex way of saying, "According to the fiction, only Ishmael survived." (*ibid.*, 212) Thus saying that relative to *Moby Dick*, p is the case is no different from saying that the rules of generation of the novel authorize p.

However, if Walton were to respond to Richard, he would still object on the lines that the fictional operator still retains the reference to Ishmael. The truth condition for (9) that Walton provides is that a genuine assertion of a statement like (9) is true iff it is fictional in the game of pretense that one speaks truly. To individuate the pretense without using a proper name or individuating description so the question of the ontology of the proper name does not crop up again, Walton appeals to one pointing to a pretense of kind K as seen while discussing Zemach. The motivation for this account is to avoid any commitment to fictional entities whatsoever.



But if Richard had to give a full account, he would say the truth condition is that *p* is the case iff *p* is true in the work *W* and there is a mandate to imagine by work *W*. The problem however, since Richard does not deny that statements like (9) do not literally say anything is that it is still underdetermined whether (9) expresses a proposition or not and thus whether the fictional proper name refers or not. This is so since, despite the pretense Richard appeals to, it is still possible that the fictional operator merely signifies that there is a pretense, and the pretense is that of an abstract object that it is a man (like in Salmon's (1998) account or a mental representation in my account, which is likened to a man). Richard still has to live with the possibility that the reference to Ishmael has not been eliminated, since he does not deny the "theoretical fact" (*ibid.*, 212) that the sentence does not express a proposition and therefore does not literally say anything. Thus, unlike his claim that the "theoretical fact" is not relevant to explaining why one engages in pretense, it in fact turns out to be relevant, since it explains the complete elimination of the possibility of Ishmael referring to anything. Moreover, Walton's overall account provides for the elimination of fictional entities in extra-fictional discourse that abstractionists like van Inwagen appeal to where Walton appeals to unofficial games of make believe, as written above. If Richard has to eliminate such entities in extra-fictional discourse, especially where the fictional operator does not apply, he will have a tough time with the current account he offers, even if he invokes unofficial games of make belief, since it will still be underdetermined as to whether the name refers to a fictional object or not.

The only way Richard can succeed with his argument then is to acknowledge that Ishmael still retains reference and accept an ontological commitment to an abstract referent or a mental representation. This is something I am willing to countenance as of now. Or he turns into a descriptivist saying that this is an intentional context and that enables him to say that "Holmes" has no reference but sense. For now, I will accept that Kripkean arguments do defeat descriptivism.

### **5. Pautz's Objections to Walton**

As per Anna Pautz (2008), every theory of fiction needs to satisfy the fictional coreference constraint, that is, provide an analysis of statements of this form (like the Geachian intentional identity problem above) (*ibid.*, 149):

(13) Bridget believes that Sherlock Holmes is smart and Caroline believes that he (the same person) is smart.

Realists (abstractionists) account for the truth of this by saying that since “Sherlock Holmes” refers to an abstract object, both Bridget and Caroline have the same object as an object of their belief.

Pautz says that for the pretense theorist, pretense has wider scope than belief, so the analysis would be that both (i) and (ii) taken together make (13) true (*ibid.*, 149-150):

(i) Bridget pretends that Sherlock Holmes exists, that ‘Sherlock Holmes’ refers to Sherlock Holmes, that ‘Sherlock Holmes is smart’ expresses a proposition, and that she believes the proposition.

(ii) Caroline pretends that Sherlock Holmes exists, that ‘Sherlock Holmes’ refers to Sherlock Holmes, that ‘Sherlock Holmes is smart’ expresses a proposition, and that she believes the proposition.

The counterexample to this is what Pautz calls the “problem of coincidental similarity”, which is of two people reading two different novels by two authors in which it’s by pure coincidence that the character has the same name.

Pautz says that one response to this is that they are part of an extended game of pretense (*ibid.*, 150):

(14) The pretenses of Bridget and Caroline are included in the same extended pretense game.

Thus, Bridget and Caroline are part of the same extended game of pretense iff (*ibid.*, 150):

(iii<sub>a</sub>) Bridget and Caroline engage in pretense in relation to the same novels.

Thus, (13) is made true by (i), (ii) and (iii<sub>a</sub>).

One problem this faces, as Pautz points out, is that if the pretense theory tries to provide an analysis of what makes a copy of a novel on Holmes in French the same as one in English, the analysis cannot say that the copies in different languages are the same because they are about the same character. The other issue is that this does not generalize to the case of

Conan Doyle not writing the novel but there being an oral tradition of Holmes stories. That is, that an oral tradition is not the same as the “same novels”.

The revision to this would be that “two acts of pretending belong to the same extended pretense game if and only if both acts of pretending have a common causal source: as it might be, the telling of a story by a certain person *P* at certain time *t*.” (*ibid.*, 151) The problem Pautz points out is that the Holmes stories might have been an oral tradition with various sources with no source in particular that can be traced, and stories on Holmes might be a mixture of stories told by different untraceable people.

Pautz’s final formulation of a plausible pretense position is that (iii<sub>b</sub>) along with (i) and (ii) make (13) true (*ibid.*, 152):

(iii<sub>b</sub>) Bridget’s use and Caroline’s use of ‘Sherlock Holmes’ while engaging in pretense are causally related to the initial pretend use of ‘Sherlock Holmes’ by Conan Doyle.

Pautz introduces what she calls an “intended pretend belief”, that is, a pretend belief that is mandated by the prop in the make believe and the principles of generation. Her counterexample to (iii<sub>b</sub>) is as follows. Suppose Bridget has read all the Holmes’ novels and her pretend use of “Holmes” is causally related to Conan Doyle’s. She also has corresponding intended pretend beliefs about Holmes acquired from the novel. Now suppose Caroline has merely heard the name and then ends up thinking and believing that it is about a story where Holmes is from California and is a rich surgeon with a Ferrari who helps the police solve cases. The only intended pretend belief she has is that Holmes is smart. Surely what she has in mind is a different character. But if the above causal thesis is correct, it would come out as the same character. Since the prediction is wrong, we need to reject (iii<sub>b</sub>) too.

Hicks (2010) partly rejects this intuition. He says that they are in fact about the same character because one can merely say that Caroline’s beliefs are false. That is, if they are false, the reason why they’re false is because Caroline’s beliefs are about the same character. Hicks thinks that when Pautz talks about unintended beliefs, she conflates anti-fictional beliefs (those that are false in the fiction, for instance, statements made by an unreliable narrator) with unauthorized beliefs (beliefs part of an unauthorized game). Once the distinction is made, both the same and different character intuitions are accounted for. Hick’s intuition that it is the same character is because he sees Caroline’s belief as an anti-

fictional one on the basis of the authorized game. It is false because it is anti-fictional. Pautz sees it as a different character because she sees it as part of a different unauthorized game.

In my view, both intuitions can be right in different contexts. Suppose on her notions of “Holmes”, Caroline goes on to write different novels. It seems clear that it is a different character. Then it just happens to be the case that she is using a homonym of “Holmes”. But if, after hearing the name and not having written a novel, she goes to a Holmes fan and says, “You know I heard about this great character called “Holmes”, who I think is a surgeon with a Ferrari...”, the reply would be, “You heard it wrong sister.” In this case, she was talking about the same character, but got the details wrong. The reason why it even looks like a clash of intuitions is because as I have argued above, there is no independent fact of the matter about their identities with respect to there being an intrinsic property of identity that the fictional character possesses (as said above in the discussion on co-identification). It is only that Hicks seems to be calculating probability that it is the same based on thinking about contexts like the one with the Holmes fan in mind and Pautz is doing so with contexts like her writing a different novel or that she has insufficient contextual cues to say that it is the same which is why she’s saying it is probably different, and therefore it is different. The contextual clues and conventions make it highly likely that it is a different character.

If this is so, then Pautz’s counterexample works, since despite being causally related to the initial pretense, and having at least one intended pretend belief right, she can be talking about a different character.

Garcia-Carpintero objects to Pautz saying that the causal relation needs to be of the “proper form”, that is, he asks “does the speaker with the wrong beliefs use “Sherlock Holmes” as *de jure* coreferential with the uses of the speakers from which she has taken it? I.e., does she intend to use it as it has been used by previous speakers?” (*ibid.*, 27). Given that this was her intention, Pautz is wrong that there is no co-identification. Caroline merely has mistaken beliefs about Holmes. Moreover, if she does not have the intention, then the character does not co-identify, even if there is a causal link, and in Garcia-Carpintero’s case above, a causal quasi-anaphoric link.

There are two problems with this argument. One is that it is circular in terms of the individuation of the intention. Let us make a distinction between a *de facto* intention to use it as used by previous speakers and a *de jure* intention. A *de jure* intention is one in which one intends to use it as previous speakers and which is right, whatever one thinks about it. A

*de facto* intention is one in which one thinks one is intending to use it as used by previous speakers but one might get it wrong. Like in Evans' Madagascar case, "Madagascar" was originally a part of the African Mainland, but Marco Polo mistakenly thought it was an island. For all you know, he intended to use it as previous speakers, he thought he was right, but got it wrong.

The question then is, how does one distinguish between a *de facto* intention where one gets it wrong and a *de jure* intention which is right, where thinking that it is right or wrong does not really matter to getting it right? What makes "x intends to use a name a y" true? Surely it cannot be a thought that "I intend to use "Holmes" as the same as her"? For I might still be wrong. It looks like, in this picture, the only way a *de jure* intention is right and one therefore talks about the same thing as the previous user is for it to stand in the right causal relation as previous speakers. This makes the whole account circular. Because the initial point was about what makes co-identification correct given causal links. The crucial factor was the correct intention. However, even to say that the correct intention was right, we need to appeal to the right causal link.

Perhaps that's not fair. One can say that one's intention is right because one has made a probabilistic inference, based on contextual cues and there is a high probability that one has got it right. And it is right because the right constitutive rules were appealed to which were relevant to context. If that is so, however, then there is no reason why a causal link should even be appealed to in judging the intentions to be right ones, nor even to determine that one is thinking of the same object as previous ones, since that is determined by contextual cues and constitutive rules, as I have argued above. Also, given that one is making such an inference, it also explains why one can get the intention wrong.

This brings me to another point. The causal links look even more magical given the experience of fiction. Normal physical causation involves causality in the physical world. The causal picture would include words uttered as sound waves. Unlike physical objects, there is no such physical effect of the meanings of words or word usage nor do visual or tactile effects of marks on paper or another medium by themselves contain meanings. Why should such impersonal causal features by themselves even be determinative of securing reference or co-identification? One might reply that other people's intentions and/or the meanings they mentally associate with sounds too feature in this causal link. But I have no causal contact with other people's intentions or other people's mental associations. How can they have a causal effect on me like physical features of the world? A New Age answer

would be that mental energy is transferred causally to me like magic coming forth from a Harry Potter wand. This obviously should not be taken seriously. If this were so, the causal link is magically transferred, preserving the identity of the object thought about. The other option is to say the intentions are very much present in the causal process, but only piggybacking on them indirectly, since one is simulating or making inferences about what is in other people's minds. But then it looks like what is determining sameness of intention is not the causal link of blind physical processes but my inferences about other people's intentions and the meanings they associate with words (of course, this can be considered physical too. But let's just say for the sake of discussion let's say we are talking about all other physical things apart from mental inferences). The causal links only allow me to make the inferences and possess contextual cues. If I possess the correct contextual cues and rules, I can say of objects of thought whether they be the same or not by making an inference. I can be in a world without causal links apart from me where I possess contextual cues and be reading a work of fiction which popped out of nowhere, and I can still understand it. Moreover, if one is in fact correctly inferring what meanings people mentally associate with words which are conventionally accepted, why is the causal link even relevant to securing reference or co-identification? My knowledge of social conventions, especially as to what are the correct associations of meanings with words, and contextual cues should be enough for me to be able to infer the correct sort of mental associations with meanings people have in their heads.

The reason why I say, in particular that this becomes even more salient in the case of fiction is that fictional characters do not form part of the ordinary causal links in the world, unlike concrete objects, which are linked to the vibration of sounds I produce.

One objection that can be put to me is that what happens in the case where Caroline thinks about the wrong character? The answer if that I say that she is thinking about the wrong character because given what constitutive rules I possess, and the context she is in, the criteria does not match. There is no need to invoke any sort of causal link here.

Now, Hicks acknowledges that Pautz has a stronger objection in that she argues that the best way to make sense of why Conan Doyle, Bridget and Caroline agree that "Sherlock Holmes" is about the same character is to understand the agreement as (Hicks 2010: 155):

(14)  $\exists p$  (Conan Doyle pretended to believe  $p$  and Caroline and Bridget pretend to believe  $p$  and  $p$  is the proposition that Sherlock Holmes is smart).

Thus, contra Walton, one now has to admit that there exists a proposition outside the scope of the pretense, and if one is Russellian about propositions, so does the abstract object “Sherlock Holmes” refer to.

To this Hicks replies that one merely pretends to learn and thus pretends to think from Conan Doyle that “Holmes is smart” for both Bridget and Caroline. In the way there is no actual learning, there is no actual thinking either. He bases the idea of pretending to learn on Evan’s view as given below:

A story-teller pretends to tell (inform) us about things. He pretends to be informed about these things. (That is, he pretends to be informationally related in some way or other to the events he relates; although the precise relation need not enter into the pretense.) We, hearing him, are prone to carry on the pretense [...] we pretend to have been told of these things (to know them by testimony). (Gareth Evans 1982 : 358, n. 31 as quoted by Hicks 2010 : 397-398).

What is important to realize is that our tendency to carry on the pretense, our engagement with the fiction, is our tendency to pretend things, not our tendency to actually think anything. It is, as I said above, a further step to get to the beliefs “about fictional characters” that are true or false. (Hicks 2010: 399)

Thus, after one pretends to learn that Holmes is smart, what happens is that “there is a way of “thinking,” here, pretending, shared between the three of them.” From this, Hicks says that if thereby they attempted to deceive us into thinking that Holmes exists, they won’t be creating a new entity by thereby doing so. Therefore, there is nothing new introduced in one’s ontology.

One reply that Hicks anticipates is that in doing so, there might still be “genuine thinking”, that transpires on the basis of the pretense that it is fictional that Holmes is smart. However, his reply is that even then, there is no such proposition as the quote from Evans indicates, and even the so called “proposition”, that is, what is said by “it is fictional that Holmes is smart” is understood in terms of pretense about Holmes.

There is a simple counterexample to what Evans’ view above which, as I interpret it, shows that the presence of pretense is a sufficient condition for there being no proposition and of not learning or genuinely thinking anything. When one is doing counterfactual history, one pretends counterfactual scenarios. Good counterfactual history examines evidence and considers real world conditions to make predictions about what would have been most likely

if some even had not happened. Rebecca Onion (2015) talks about *The Holocaust Averted: An Alternate History of American Jewry, 1938-1967* (2015) by the historian Jeffrey Gurock. In it, he compellingly argues that if the Allies defeated Hitler in 1938 itself and the Holocaust and the World War did not take place as a result, then given the evidence, it is most likely that the anti-Semitism in the United States prior to the war would have persisted the way it was. The author in this case pretends that the antecedent is true (even though it is false) and makes his case. Despite the pretense, one actually gets informed by him, there is no “pretend” informing and there is genuine thinking involved. There is also a false proposition in the antecedent. Even if the account would have even one fictional name for the sake of argument, say a Jew facing the brunt of anti-Semitism, we would still have been informed (For instance, “if there was a Jew named Mordecai, it is likely that Mordecai would have felt a strong pressure to assimilate and his political loyalties would have been suspected”). He might also have gone ahead and described what is most likely to have happened to Mordecai in such and such circumstances, and it would genuinely have been informative.

From the discussion above, it is evident that it is a false dichotomy to assume that pretending a false proposition to be true implies that there is no proposition expressed, since it assumes that either there is pretense or there is a proposition expressed.

There are two more adverse consequences of this picture. The first consequence is that it does not preserve logical form for paraphrases and relations of entailment. This is a criterion that van Inwagen insists that any theory of fiction must have. A realist picture fulfills this. For instance (15) entails (16) (Walton 1990: 416):

(15) There is a fictional character who, for every novel, either appears in that novel or is a model for a character who does.

(16) If no character appears in every novel, then some character is modeled on another character.

However, for Walton, both will be paraphrased on the lines of the form:

To engage in pretense of kind K is fictionally to speak truly in a game of such and such sort. (ibid., 416)

In Walton’s account, “what our paraphrases seek to capture is what speakers say in uttering the sentences cited, not what the sentences themselves mean or the propositions they



express, if any”. (Walton 1990: 417) Since this is so, then apparent logical form is not part of the content of what is said by the speaker. The quantificational structure then, is also different. Thus, when I utter (15) and (16) it is only fictional of me that there is entailment by virtue of logical form alone. Consider (*ibid.*, 418):

(17) Statement (15) entails (16) by virtue of logical form alone.

However, (17) is false literally, but true if it is “indicating a kind of pretense” and claiming that to pretend this way (kind K) is fictionally to speak truly. It is thus not the logical form alone that makes (15) entail (16) but also the principles of generation. In the following quote, Walton’s position is made clear ((20) and (21) in the quote are (15) and (16) here):

The likelihood of confusion is enhanced by the fact that what is asserted by (20) probably does entail what is asserted by (21), though not by virtue of logical form alone. The principles constituting the implied games of make-believe are likely to be such that it cannot be fictional that to assert (20) is to speak truly unless it is fictional that to assert (21) is to speak truly. The unofficial games would have to be rather exotic logical fantasies for this not to be so. And the principles are likely to be such that the circumstances required to make it fictional that one speaks truly in asserting (20) are ones that make it fictional that one speaks truly in asserting (21). No wonder it is difficult to conceive of (20) being true and (21) false. (*ibid.*, 419)

I have quoted the above to highlight how vague and confusing it is to say that something entails something but not by logical form alone. Logically speaking, either a conclusion follows from the premises by virtue of logical form alone or it does not follow at all. So, it remains to be interpreted what Walton wants to say.

Perhaps what Walton wants to say is that the principle of generation might say something on the lines of “Unless otherwise indicated, for every sentence in the game, inferences are valid by virtue of the logical form”. Thus, the sentences with their apparent logical form work like a real world prop, like in a game of make belief, where every stump is counted as a bear. So, the inferences are fictionally valid not by virtue of logical form alone. (Walton also says “it is only fictional, not true that (20) has a quantificational structure such as to entail (21).” (*ibid.*, 418))

If we look at the stump case, children’s games of pretense where bears are constrained by the laws of nature and the nature of the stumps. It is because the laws of nature hold independently that some things are fictional. Similarly, if we take the analogy, it looks more likely the game is constrained by a valid inference outside the game, which has to valid independently of the game. Of course, Walton might just reply that yes, the constraints are there, but what ultimately determine entailment are the principles of generation, since a valid argument can be constrained to be invalid in fiction.

However, in the counterfactual history case, the entailments would be based on logical form, but truth values can be pretended. So it is not necessary that the entailments are pretended even when there is pretense. Moreover, as Walton himself acknowledges, unless there are exotic logical fantasies, the default setting is of entailment by logical form, even by the principles of generation. What is the explanation for this?

The best possible explanation is that we need the counterfactuals to be valid by virtue of logical form to carry out our reasoning, even in pretense and the default setting in pretense is that entailment is by virtue of logical form. I will back this up by empirical work which links the use of counterfactuals with children's understanding of causality and in turn links counterfactuals with fiction and pretense. Gopnik (2009) elaborates on this. Contrary to what psychologists like Piaget thought, that pre-school children didn't understand causation, psychologists found that they did. Apart from other evidence, she points out the psychologist Henry Wellman found out via a database of children's conversations that children gave and wanted causal explanations dozens of times each day like "The teddy's arm fell off because you twisted it too far"; "Jenny had my chair because the other chair was brokened (*sic.*)" (physical explanations) "He needs more to eat because he is growing long arms"; "Mean hawks eat meat because meat is tasty for mean hawks." (biological explanations) "I didn't spill it last night because I'm a good girl"; "I not gone go up there because I frightened of her." (psychological explanations). (*ibid.*, 36) In other studies, children understood causes that are abstract as well as hidden.

Wellman further links counterfactual understanding with children's causal understanding, where children differentiate between what is or is not possible based on their causal knowledge. For instance, children could say that John could choose to raise his arm but not jump and stay in the air, he could not choose to gain in height or walk through a table.

This is more clearly seen in children's mental construction of causal maps, which are a map of "complex causal relations among events" (like neurologists map migraines with links between brain activity, pressure and pain), which they use to imagine what is possible and then make changes to the world. Gopnik and her team invented what was called a "Blicket machine" and which was a square box which lit up and played music whenever the right block among others (called a "blicket") was put on it. The children were told that blickets made the machine work and were asked to pick out the blickets. Children tried to figure out which one of the boxes were blickets by pressing, scratching and trying them out. Once they figured it, the experimenters added a nonblicket with a blicket, and placed them on a machine, which lit up. One four year old's response was the counterfactual, "But s'pose...you didn't put on the blicket that time, s'pose you only put on that one (pointing to

the nonblicket) then it wouldn't have gone." (*ibid.*, 44) This was again, an instance of children's knowledge of counterfactuals linked with causal knowledge, which they used to make predictions and imagine possibilities, like to stop the machine one must remove the blicket, without having seen a blicket removed.

The linking of causal and counterfactual knowledge also explains why children "have the parallel ability to generate counterfactuals and explore possible worlds. If children understand the way things work, they should be able to imagine alternate possibilities about them." (*ibid.*, 37) It also accounts for how children who do not understand how something works cannot think of possibilities for them, not because they do not have the ability to.

Thus:

Causation is what gives fantasy its logic. Think about the children in Paul Harris's studies who could work out precisely what the imaginary consequences would be if Teddy spilled the imaginary tea ("If the teddy spills tea you'll need a mop, but if he spills baby powder you'll need a broom"). A pretend game in which absolutely anything goes would just be a mess. Instead, pretend works by establishing imaginary premises ("I'm the mommy and you're the baby") and then working out the causal consequences of those premises quite strictly. Children can become quite passionate about whether the right causal rules are being followed: "You didn't get me with your ray gun 'cause I was behind the shield!" "You hafta drink your milk 'cause you're the baby!" (*ibid.*, 37)

We can now begin to explain why the default setting in pretense would be of entailment by virtue of logical form. When one reasons counterfactually about causality and works out "causal consequences" from premises, one needs to hold on to actual rules of entailment based on logical form. Not only that, if one is using causal knowledge in counterfactuals which further become the basis for imagining possibilities, which in turn are used to make predictions and act on the world, it must have entailment based on logical form as one basis. This is exactly why Gopnik would say above that "pretend works by establishing imaginary premises ("I'm the mommy and you're the baby") and then working out the causal consequences of those premises quite strictly." (*ibid.*, 37) It shows that the inferential processes are the same.

Moreover, Gopnik and her team showed to children how talking to a machine with a switch would make the block light up instead of turning on the switch (someone was orchestrating this while hiding). The children who thought previously that talking to the machine and starting it was impossible thought it was now possible. The conclusion is that it is new knowledge of causation that permits us to imagine novel possibilities. Possibilities we imagine are constrained by knowledge, which is why in *Blade Runner*, Harrison Ford can go to a pay phone with a video screen. Gopnik's conclusion is that work on causal maps shows that knowledge and imagination are not opposed to each other, and it is causal knowledge of

connections between events that allows us to produce counterfactuals which allow us not only make predictions but alter the way events are connected with each other and create novel connections. The normal use of counterfactuals and inference by means of them in causal knowledge and counterfactual knowledge is linked. Therefore, in its default setting, in order to make inferences efficiently, the counterfactuals cannot but have normal entailment by means of logical form, especially as we saw in the counterfactual history case.

Furthermore, Gopnik argues that like physical and biological counterfactuals, psychological counterfactuals as well as those which comprise imaginary companions like Teddy (like “if Teddy spills the tea, the floor gets wet”) have a link with our knowledge of other minds. She cites work by Marjorie Taylor (1997) that children who have had imaginary companions have a better theory of mind and can predict others’ thoughts, feelings and actions than others. They were also found to be more sociable. It was also seen by Taylor that this ability was also linked to the adult ability to conjure fictional worlds. Almost all of the fifty fiction authors she interviewed had had imaginary companions as children, and around half could recall details about them. On the contrary, only few high schoolers could recall them and most adults forget them. This also shows that fiction seems to stem from the same ability to have psychological counterfactuals from which we make causal maps of other minds as well as our own.

We propose therefore that children pretend so much because they are learning so much. Our causal maps allow us both to understand the existing physical and psychological world and to invent and realize new physical and psychological worlds. The line between a fiction and a close counterfactual is one of degree rather than kind: Fictions are counterfactuals that just happen to be further away from our real world than other possible worlds. It is possible that engaging in far counterfactuals allows children to fill in parts of their developing causal maps that are not accessible through real interventions and observations. (Walker and Gopnik 2013: 354)

Moreover, Gopnik says that fictions enable us to think of possible ways we can be and possible choices we make and can make and their consequences, just in the way we think about causal connections in the world and think about consequences without observing them.

The reason I bring this up is not only to show the question of logical form and entailment in which entailment through logical form helps us figure out possible consequences in the physical and the psychological realms. It also helps deal with a major picture given by Kripke (2013) to the question “Could there have been unicorn?” to which he answers in the negative, which is something that even Walton would have to accept. Like Walton, Kripke

also believes that there are no propositions that intrafictional statements containing fictional names express, which seems to be one basic assumption for such a denial apart from other arguments in his lectures on *Reference and Existence*, nor do names in fiction refer. The problem with their account would be the likelihood, without providing a further explanation, that Gopnik's empirical account of counterfactuals and the way we reason with them is false a priori. This is because the extensive use we make of them, given their foundation to make predictions and the like, would presuppose that we can reason with them by means of their logical form, and work out the consequences immediately, using our mental processes, without being dictated by conventions like rules of generation at every step. This is something I've mentioned above. What is most important is that, if that is so, then we have to allow them to express propositions to allow for logical entailment. It is also much more parsimonious to believe that under normal circumstances, (15) entails (16) only due to logical form, and the further addition of principles of generation for this seems like an additional burden merely to accommodate a particular theory of fiction. To accommodate logical fantasies in fiction, all one needs to say is that normal logical entailment applies, which allows one to make predictions about the real world, and logical laws are violated only by certain contextual indicators. This also explains why breaking logical laws in fiction always leads to surprise or some additional effect in a work of fiction. Of course, one might still say that the logical laws are valid in the world of the fiction per se, but that is not very different from us importing real world assumptions into fiction. This explanation is simpler than saying that something else too allows them to be valid in the fiction.

Thus, if we use counterfactuals for fiction in the same way we use counterfactuals for physical and psychological causation, unless otherwise directed by the fiction to do so, that fiction is constrained by causal knowledge, allows us to make predictions, enables us to learn without observation and think about possibilities which haven't occurred. An account which denies that there could be no unicorns would deny that counterfactuals are essential to fiction in the way that Gopnik depicts in her hypothesis, given that the foundation for reasoning in fiction would start with the considering the possibility of their being a unicorn or Holmes. All of this presupposes that on the face of it, at least that Holmes could have existed and done such and such things, on which the fiction is based. The burden is on Kripke and the denier of propositions for fictional statements to provide an additional account for how this phenomenon with counterfactuals is possible in his account.

### **5.1 Excursus into my account**

It is fitting here that I discuss how my own account fits in with the objections using counterfactual histories and Gopnik's hypothesis I have placed against Walton.

In my previous chapters, I have mentioned that all fictional statements about fictional characters presuppose statements of the form:

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

In fictional contexts, this presupposition (which is held in relevant metafictional contexts too) accounts for the truth of fictional statements and accounts for entailment by virtue of logical form.

(MET<sub>FIC</sub>) James Tiberius Kirk is fictional but is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

For instance, if (MET<sub>FIC</sub>) is presupposed, then the statement "James Tiberius Kirk is the captain of the starship USS Enterprise" (*ibid.*, 137) will come out to be true given conventions of fiction which are invoked by the "is fictional" predicate and the part about "all intents and purposes". Moreover, I have mentioned in my chapter on Thomasson (Chapter 3) that the intentional stance is at play in fictional contexts, which accounts for the aboutness of fictional characters as flesh and blood whereas the design stance is at play in seeing the abstract aspect, like when we see it as a fictional character in metafictional contexts.<sup>34</sup> It is pertinent here that the intentional stance involves one's theory of mind, the way in which Gopnik talks about the psychological origin of fictional characters in imaginary companions and our theory of mind, and it correlates with being able to better predict others' thoughts feelings and actions. My account of the relevance of the intentional stance in fictional contexts is therefore not only consistent with empirical work but is also bolstered by it.

(MET<sub>FIC</sub>)\* also dispenses with Walton's need to invoke props to generate fictional truths and thus to see novels as props. To recall from above, that props generate fictional truths accounts for the objectivity of how there are truths in fiction, which are objective, whether we want them to be true or not. If, however, we latch on to fictional conventions via

---

<sup>34</sup> To remind, "the intentional stance is the strategy of prediction and explanation that attributes beliefs, desires, and other "intentional" states to systems- living and non-living- and predicts future behaviour from what it would be rational for an agent to do." (Dennett 1988: 1) In it, an object or an artefact is treated as a rational agent. The design stance is the heuristic strategy of assuming that an object is designed for a purpose and that it will act according to its purpose

(MET<sub>FIC</sub>)\*, there is no need for this Waltonian picture. The presupposition and knowledge of conventions is enough to generate fictional truths, given the truth of the presupposition. The sort of objective conclusions that come out in counterfactual history are also accounted for by a presupposition like this, where instead we merely substitute a predicate like “is counterfactual” instead of “is fictional” for relevant intents and purposes.

It was also mentioned by me above that presuppositions like Garcia-Carpintero’s rf-presuppositions were required to account for being able to zero in on the right homonym in a particular context as well as being able to infer aboutness, say about London, given accurate content. (MET<sub>FIC</sub>)\* functions as such a presupposition, where triggering of contextual clues like “is fictional” as well as the discovery of contextual clues allows us to see relevant intents and purposes. This enables us to not only do the needful, but these parts of (MET<sub>FIC</sub>)\* reflect that the process of aboutness is the output of a particular inferential process involving contextual clues and conventions. Thus, my account includes the positive aspects of Garcia-Carpintero’s account, as well as accounting for the Alice/Alex/Adrian case as well the other cases he mentions. It is also stronger than his account, since I do not require possible worlds like him, but only require a unified mental representation quantified (with the content given by mind-dependent properties<sup>35</sup> encoded by each simpler mental representation the unified representation is composed of) over even as it is understood as an ontological metaphor<sup>36</sup> (reflected by “x is likened to a real person”). It is the presence of an ontological metaphor which makes Garcia-Carpintero think of it instead of as a dead conventionalized metaphor. But since I have pointed out that this does not account for the content, my account allows for content in terms of mind-dependent properties which falls in line with realist intuitions. The mind-dependent properties when it comes to fictional characters have been explained before in each chapter and will be explained further in Chapter 6.

## 6. Matravers’ Objections to Walton

---

<sup>35</sup> The mind-dependent properties I am talking about are, for instance, sensory, motor, emotional and verbal features encoded by representations as well as features of innate ideas. It helps to see the parallel with what Friend (2011) calls an idea, which is a mental representation of a property or relation, expressed as, for instance, *idea-of-being-the-chief-city-of-a-fascist-state*). We saw above, that is part of the content for fictional statements. My account is more fine grained as we saw in previous chapters and will see in Chapter 6.

<sup>36</sup> An ontological metaphor is a metaphor in which things like minds, (THE MIND IS A BRITTLE OBJECT presupposed by statements like “His ego is very fragile, “He was shattered” etc)) street corners are understood as concrete objects etc. for the purposes of quantification and reasoning. See Lakoff and Johnson (1980)

Matravers' (2014) project is to bring down the entire edifice of Walton's theory. In his book *Fiction and Narrative*, Matravers goes against Walton's foundations and attempts to solve philosophical problems arising out of fiction by means of a new theory. However, he does not try to solve the problem for the ontology of fiction, which I will try to do here.

In Walton's picture, truths about the actual world, whether it be by stipulation or convention, make things true in the game world. For instance, the stipulation "All tree stumps are bears" make it true of a stump on Eric's right that it is a bear. When it comes to fiction, the sentences of a work of fiction make them fictional and the arrangement of paint in a painting make it fictional that, for example, there is a couple strolling in a park in *La Grande Jatte*.

It is thus stipulations (stipulations are principles of generation) that take us from the actual world truths to fictional world truths and which are "prescriptions to imagine". Given the stipulation, we can imagine what is prescribed. As we have seen, fictional truth "consists in there being a prescription or mandate in some context to imagine something (Walton 1990: 39). Moreover, the prop mandates imaginings and thus generates fictional truths.

In the stump case, the tree stump is an object of imagination, since one is prescribed to imagine something of it. Thus, imagination allows us to move from truths about the actual world to truths about the fictional world. Since that is so, Matravers says, the stipulated principles of generation are not sufficient for us to go from actual world truths to fictional world truths. One could in principle derive truths by merely applying the stipulations, but the reason why that is not sufficient and the imagination is also required is because it explains our "psychological engagement" with the fiction. Using the imagination explains that we ourselves are immersed in fictions and games of make believe (this explains our perspective). For instance, Eric not only thinks of a tree stump that it is a bear but also imagines a bear facing him. A child does not merely understand a fictional truth as part of a list, but is immersed in the world of fiction. Walton calls this *de se* imagining. Another aspect of our psychological engagement is that the imagination makes our experiences vivid, that is, it makes us not think about the fact that there is no real bear in the tree stump case. It thus also strengthens our emotional engagement with fiction.

Matravers calls Walton's idea of something being fictional if the imagination transforms a true proposition in the actual world to another proposition which is true in the fictional world the "transformation criterion". It should be noted that this is with respect to propositions that are true in the actual world and not with respect to statements about



fictional characters, which do not involve propositions. He cites this passage (*ibid.*, 13) from Walton to support his formulation of the transformation criterion.

It is reasonably obvious, often enough, that a passage in a novel is to be construed as a more or less straightforward observation or pronouncement about the actual world, addressed by the author directly to the readers. The opening sentence of *Anna Karenina* ('Happy families are all alike; every unhappy family is unhappy in its own way') is frequently cited; there are discussion of love and other matters of real-world interest in Henry Fielding's novels; in footnotes to *Kiss of the Spider Woman* Manuel Puig presents a series of apparently straightforward essays recounting the views of Freud, Norman O. Brown, Herbert Marcuse, Wilhelm Reich, and others on sexuality. But it is rarely wholly clear that such passages do not also have the function of eliciting imaginings, of making it fictional, for instance, that someone—a character through whom the author speaks or even the author himself—is making those pronouncements. If, in setting down the opening lines of *Anna Karenina*, Tolstoy was claiming (with allowance for some exaggeration) that all happy families really are alike but that there are many different kinds of unhappy ones, his words may also make it fictional that someone—the narrator—utters them assertively. If this is their function, the passage is fiction in our sense. (Walton 1990: 90–91)

Furthermore, Matravers makes this comment on the passage to justify his formulation of the transformation criterion:

Walton allows that some passages of novels can be construed as non-fictional: 'more or less straightforward observation or pronouncement about the actual world'. However, with others, the pronouncement changes its identity: an 'un-owned' claim in the actual world is transformed into—for example—someone's utterance in the fictional world. If there is such a change the passage is fictional; if there is not such a change, the passage is not fictional. (*ibid.*, 13-14)

The counterexample Matravers provides is that in Raymond Chandler's novel *The Little Sister*, the same proposition is true in the actual as well as the fictional world for the sentence "Room 332 was at the back of the building near the door to the fire escape." (Chandler 1955: 48 as quoted in Matravers 2014: 14) Matravers argues that in this case, the application of the transformation criterion will yield no difference whatsoever, since there is no transformation that has taken place. The transformation criterion thus, "will not distinguish actual propositions from fiction propositions." (*ibid.*, 14) Though another account may be required of fictionality, it does not involve the transformation criterion.

One anticipatory response by Walton was that the transformation takes place because these sentences are uttered by a character or a narrator. This, Matravers points out is inadequate since all that is needed to understand the sentence is a basic grasp of English. One reply can be:

One might try to argue that there will always be some transformation because, at a minimum, in the actual world the proposition is false and in the fictional world the proposition is true. (*ibid.*, 15)

My understanding of this is that a proposition like “Room 332 was at the back of the building near the door to the fire escape” is false in the actual world, since it is a different proposition in the fiction. For this to happen, there must have been at least a transformation of the actual world proposition which was true into a fictional proposition. However, Matravers points out that the content of the proposition remains the same, even more so since there is no analogy between this case and the transformation in the tree stump to bear case, where the actual world truth “there is a stump in the thicket” transforms into the fictional “there is a bear in the thicket”. (*ibid.* 15)

Moreover, the transformatory step is redundant. If one can really grasp the content of the sentence above from *The Little Sister* and one recognizes that it is fictional, then one can hold that it is “true in the fiction”. This is sufficient to establish the fictionality of the proposition and there is therefore no reason why one needs to imagine it to be true to make it fictional.

One more objection by Walton that Matravers conceives of is that even if some cases of transformation are conceded by Matravers, then one can argue that the imagination is not redundant, since one can imagine the same propositions which are expressed by sentences in the work of fiction which are also expressed by sentences about the actual world. The problem with this is that the transformation criterion was what was supposed to make propositions fictional. So there is a circularity, since one already has to know which propositions are fictional to know what to imagine.

The next criterion Matravers goes after is the “engagement criterion”. That is, “something is fictional if it engages our imagination as only the imagination can account for facts concerning our engagement with fiction.” (*ibid.*, 16) To recall, engagement with fiction involved *de se* imagining. But a non-fiction book like the biography of Churchill can mandate *de se* imaginings too, like when one imagines seeing the House of Commons provoking Chamberlain’s resignation.

Matravers says that Walton does acknowledge however that nonfiction like “the New Journalism” involves mandating imagining as well as some histories which also do so, but that is only doing so since what is being said, like people’s mental states and the imaginative parts are outside the purview of their “sober epistemological grasp”. This is thus not a counterexample. The problem Matravers points out, however, is that Walton claims that, “any work with the function of serving as a prop in a game of make-believe, however minor or peripheral or instrumental this function might be, qualifies as “fiction”” (Walton 1990: 72 as quoted by Matravers 2014: 18).

According to Matravers then, this criterion becomes so broad that every narrative (without the transformation criterion) becomes fictional, since it mandates the use of our imagination. Matravers has also argued against the transformation criterion, so it is tough for Walton to escape this objection.

Matraver's own view consists of denying that the fiction/non-fiction distinction is primary and saying that what is in fact fundamental is what he calls the representation/confrontation distinction. The standard view, including Walton's, maps "pre-existing structures of belief" (*ibid.*, 76) in different ways to non-fictional and fictional content, with belief related to non-fiction and make-believe related to fiction. Matraver's view, however, makes the distinction between fiction and non-fiction not dependent on propositional attitudes. Fictional content in Matravers' theory belongs to a broader category "representations" which are "symbolic arrays like photographs which depict events and states of affairs outside our possible sphere of action because they are removed from us in time as in the case of histories) and/or remote from us in space." (Carroll 2016: 368) In contrast with confrontations, which are situations in it is possible to act, we cannot act on a representation. When we are in a confrontation relation with an object, an object is in our egocentric space and we can perceive it. In this situation, our mental states are online, and "caused by perceptual inputs from the objects of those states." (Matravers 2014: 50) On the contrary, when it comes to the representation relation, our mental states are offline and we do not act on the object of our mental state in immediate surrounding.

One of the strengths of this account is that it explains why we do not act in response to a documentary as well as a work of fiction. However, Carroll's (*ibid.*, 368-369) objections are as follows. Dolls and traffic signs are also representations. But we can act on them in our egocentric space. Moreover, it is arbitrary. Suppose one is looking at a surveillance camera when three robbers attack. He pushes the alarm below him. It's arbitrary that this would count as a representation and him looking six inches below at the robbers would be a confrontation.

The doll objection can be easily countered by saying that the doll qua a physical object is a confrontation and we are in a confrontation relation with it, which is why one can act on it. However, qua it being a doll, it is a representation, and we are in a representation relation with it. Matravers (2017) himself makes the distinction in his lecture "Those Pesky Categories of Fact and Fiction", where he says that R-properties are properties that belong to the representation and W-properties, "properties that are represented as belonging to the world". In a black and white photograph, for instance, the R properties are it being black and white and a Bloody Mary on a table in the photograph is a W property. Since the doll's

being a person is an R property and it being a physical object is a W property, the R property is what explains why we still do not act like we would when confronted with a person. Similarly with the traffic light. We are acting based on the R property which tells us to stop or go but not on the R property itself. We are acting in the world on objects we have a confrontation relationship with, like the brake pedal.

As for the surveillance camera instance, it represents a W-property but again, we are not acting at the representation and the objects represented in the camera itself. We use the representation as a reason to act on objects. Moreover, Matravers (2014) anticipates this sort of objection and counters it thus. When he says it is not possible to act on representations, he can merely reply that there is an ambiguity with “possible”. It might not be immediately possible to act on something one sees on the news, but given that it is non-fictional, one can in fact search for the events and act on them. However, when it comes to fictional representations, this is not possible and thus it is the fiction-and non-fiction distinction which is primary.

However, all this does is reinforce the point. If it is possible for us to act we are in a confrontation relation with the object of our action. All this shows is that it might be unclear at the outset whether a representation relation could be turned into a confrontation relation. Representation relations which cannot will include all fictional representations and also all non-fictional representations where we lack instrumental beliefs that would enable us to act: that is, all non-fictional representations that are out of our grasp because they are separated in time, in space, or both. (*ibid.*, 50-51)

In the surveillance case, it is in fact clear at “outset whether a representation relation” can “be turned into a confrontation relation” (*ibid.*, 50).

The reason I defend Matravers is that his account is not only the perfect foil to Walton, but also to the realist theories I have outlined in previous chapters, all of which invoke fictionality as being relevant to the semantics of fiction and therefore ontology. In the previous chapter on Salmon, I have argued just that. Even in this chapter, there were signs that showed, for instance, that the transformation criterion was irrelevant to entailment based on logical form and that it was incoherent to say that fictional statements did not say anything or they did not involve genuine thought, since they were part of the same counterfactual reasoning which was linked to causal knowledge and used in making predictions and simulating causal possibilities without actually experiencing them so we can navigate the world without constant trial and error. Most importantly, we also saw from Richards and Zemach that pretense had nothing to do with the semantics of fiction, since an understanding of English was all that was required to translate and understand fictional sentences.

Now one of the things that Matravers includes as part of our engagement with fiction is understanding it. (*ibid.*, 76) Matravers brings out evidence from psychology experiments to show that reading processes of narratives in fiction and non-fiction are not very different and what is common is narratives, which is separate from what happens in confrontations. People also understand literal sentences in biographies and in fiction in the same manner, so understanding these is only dependent on a grasp of English and is neutral between fiction and non-fiction. Matravers' differentiates between thin experiences of representations which merely convey information, which he calls thin representations, like reading his wife's note to cancel an order of milk (in which he just needs to understand English) and thick experiences of representations, which he calls thick representations, which involve engaging with representations by means of the imagination. Thick representations, Matravers says "transport' us into other worlds". (*ibid.*, 77) Thus, our experience of narrative transports us from our present and immediate surroundings, like when we get lost in a book, and the imagination is a mechanism that does this transportation (with verbal representations as input). Now the argument by Matravers is that this mechanism has to be neutral with respect to fictional and non-fictional inputs, since the imagination can transport us to the battle of Waterloo in a biography of the Duke of Wellington as well as to the Battle of Austerlitz in Tolstoy's *War and Peace*. Moreover, we can be ignorant of whether something is fiction or non-fiction and we can still engage with it. As Matravers says, "the experience of reading de Quincey's 'The Revolt of the Tartars' is the same whether we believe it is non-fictional, believe it is fictional, or (as is most likely) we are ignorant of whether it is non-fictional or fictional (in fact, it is a highly fictionalized account of actual events)." (*ibid.*, 78)

The above statement by Matravers' seems to be an overstatement and needs to be qualified since knowing whether a work is fictional or not can in fact make a difference in our experience. In fact, Matravers' does clarify this issue, where he seems to limit the sense of "engagement" with the representation and the experience of reading "The Revolt of the Tartars" which is neutral between fiction and non-fiction to "constructing a mental model of its content" (*ibid.*, 78), which is involved in understanding the representation. This is consistent with the fact that knowledge of whether a work is fictional or non-fiction does and can add to further experiences. Nevertheless, the experience Matravers' is talking about is a minimal experience common to both fiction and non-fiction, which is primary and forms the basis for further experiences.

Thus, in effect, "our engaging with representations is neutral between fiction and non-fiction." (*ibid.*, 76) All of this explains, as pointed above, when Richard says that "there will be nothing about your usage or linguistic behavior which would justify thinking that you (or

your “language module”) perceives the last sentence to differ in some semantically significant way from the first two” (Richard 2000: 211) about fictional and nonfictional sentences like the following (*ibid.*, 177):

“Snow is white” says that snow is white.

Mary said that the rain in Spain falls mainly on the ground.

In *Moby Dick*, it says that only Ishmael escaped the wreck of *Pequod*.

The reason why the consensus view links make believe to fiction is that they think that we are supposed to make believe the mental models we make when it comes to fictions and believe the ones from say, a biography. But this is not true, since a lot of fictions have a mixture of both propositions to be believed and make-belief.

A more plausible solution is to hold that, at the stage at which we are building a mental model of the content of a representation in our heads, our attitudes to the individual propositions is not one of either belief or make-belief. In as much as we have an attitude to them at all, it is merely one of them being part of the content of whatever particular representation we are reading or remembering. (Matravers 2014: 78-79)

What we then have is an attitude of make-believe that can be towards our representations, where the fiction/non-fiction divide comes in later. But as of now, the content does not seem to involve belief or make-believe. If this is so, then it again reinforces my view that the content and the semantics of fiction have nothing to do with make-believe or fictionality. Moreover, Matravers makes the fiction/non-fiction difference occur at and relevant to a different level. This is based on adherence to the fidelity constraint, the explanation of which by David Davies (2007) as follows:

To read a narrative as non-fiction is to assume that the selection and temporal ordering of all the events making up the narrative was constrained by a desire, on the narrator’s part, to be faithful to the manner in which actual events transpired. We assume that the author has included only events she believes to have occurred, narrated as occurring in the order in which she believes them to have occurred. We may term this the ‘fidelity constraint’. To read a narrative as fiction, on the other hand, is to assume that the choices made in generating the narrative were not governed in the first instance by this constraint, but by some more general purpose in story-telling. (Davies 2007: 46 as quoted by Matravers 2014: 99)

There are three objections to Matravers thesis above, two of which Matravers himself handles. One is that even when it comes to text processing, the processing is different because that which is incoherent in fiction is not incoherent in non-fiction. For instance, magical realism or contradictory fictions, or that things are incomplete in fictions, like there is no fact of the matter about the number of hairs on Holmes’ head, are coherent in fiction but incoherent in non-fiction. However, Matravers replies that our mental models while

processing verbal representations will have gaps both in fiction and non-fiction, which we fill by means of idealizations, such as that the wolf in a narrative is dangerous or not. Even in our idealizations of the mental models, the number of hairs that the wolf has is still indeterminate. Another difference that can be pointed out is that all the propositions about the world are true when taken together as background information. In this case too, Matravers says that this cannot be a ground for the distinction since even in a lot of fiction, the actual world gives the background information which is consistent with the content of the fiction, and which forms the basis for Walton's reality principle, that is:

If  $p_1, \dots, p_n$  are the propositions whose fictionality a representation generates directly, another proposition,  $q$ , is fictional in it if, and only if, were it the case that  $p_1, \dots, p_n$ , it would be the case that  $q$ . (Walton 1990: 145 as quoted by Matravers 2014: 83)

However, Matravers points out that Lamarque and Olsen note that the distinction can still be drawn, since even reference to the real world (by non-fictional names) does not occur in fiction, as substitution of co-referring singular terms does not hold *salva veritate*. So "Holmes returned to London" would be true in the fiction but not "Holmes returned to where Mrs. Thatcher lives". However, Matravers' points out that this failure of substitutivity even applies to non-fictional narrative in (18) and (19) (*ibid.*, 84):

(18) 'When Wellington rode to his official residence, No 10 Downing Street, on Copenhagen, the charger he had ridden at Waterloo, he was making a political point'

(19) 'When Wellington rode to his official residence, the house later to be occupied by Mrs. Thatcher, on Copenhagen, the charger he had ridden at Waterloo, he was making a political point.'

As we can see (19) involves as inappropriate a substitution of "Thatcher" in (18) as in the Holmes case. Moreover, they also say that substitution of a co-referring term changes the sense. Matravers' reply is that that simply shows that it involves sense as well as reference or pragmatics and semantics and says nothing about names in fiction not having the latter. Moreover, the substitution of the future abode of Mrs. Thatcher by itself does not show that senses are sufficient for meaning in fiction and that denotations are not compatible with fictions. However, Matravers points out that it is only inappropriate to do it in fiction and the reason why "Thatcher's future abode" is not substituted in both cases is that it is not done usually as it violates principles of conversational implicature and this in turn "would make it impossibly difficult for the listener to construct the kind of mental model that would enable him or her to engage with the narrative." (*ibid.*, 85)

In my view, this is a significant conclusion for the semantics too, since it also shows that fiction does not necessarily produce an intensional context. Operator approaches block fictional realism by positing intentional contexts prefaced by “according to the fiction”, where “Holmes is a detective” is not true simpliciter with “Holmes” referring to Holmes, but in an intentional context within the scope of the operator. This in turn prevents quantification over Holmes. This was Richard’s approach above. However, Matravers’ picture is strong enough to show for both Walton as well as the theories discussed in my thesis, that they all have it wrong since they assume that fictionality is somehow relevant to the content of a fictional statements and is relevant to our understanding them. If this is so, then Walton’s pretense theory by means of make believe does nothing to undermine ontological commitment to fictional characters. It also undermines any specific problem regarding fictional characters versus mythical characters or theoretical posits or imaginary companions.

However, it is instructive to look at a counterexample to the thesis that something being fiction or non-fiction is not relevant to understanding the content of a statement. Sainsbury gives the example that if we come across a scene of a movie on riots in Chechnya, then we can understand its content without knowing that it is fact or fiction. For instance, one may mistake it for the news. The counterexample Hicks provides is that if Toronto is being used as a film set to depict New York City, and if one watches it and thinks of it as Toronto without knowing that it is depicting New York City, then what one understands about the world is not a sufficient condition to understand the content of the film.

I think the first counterexample is weak, since one watches the movie and thinks of it as Toronto only because one hasn’t understood its content and relevant contextual cues, not because one needs to understand its fictionality or pretend anything. One just hasn’t understood it properly as per the narrative given. Moreover, the reason why one thinks of it as New York is because of a perspective shift, in which he is thinking in terms of how language’s relation to the external world rather than in terms of language’s relation to the representation per se. In effect, the person hasn’t understood the semantic content of the representation properly, which explains why someone walking in the middle can both veer between thinking it is the news or that it is about New York. Hick’s counterexample cannot explain why one can mistake it for being about the news. This also further strengthens my point about probabilistic inference from contextual cues and knowledge of conventions determining aboutness. In this case, a person who thinks it is New York is wrong because he or she hasn’t calculated it based on the contextual cues that the film provides.



The other counterexample is that suppose Jill does not know that a film is fictional, sees a character in a film whose name is “Smith” jump and says “Smith jumped”. Sainsbury would say that this is “ontologically committing” and would entail the belief that someone jumped and her belief is false since it also entails that someone is Smith. Similarly, if Tom Cruise was playing “Smith” and she recognized him, she would say “Cruise jumped”, and her belief that “someone is Cruise” would turn out to be true. The difference between this and the Jones case is that since she does not recognize Jones, she thinks about Smith since she is “engaging with the fiction, and Jones is not a character in the fiction. Smith is.” (*ibid.*, 1558)

This is a counterexample because the difference is explained only because she is engaging in the fiction, otherwise she would have thought of Jones, and the prediction of the literalist would in fact be that her thought is actually about Jones which is why she has the false belief that Jones’ name is “Smith”.

Furthermore, Hicks asks us to suppose that Jones jumps on the set, Jill now has two false beliefs. One that Smith jumped and the second that Jones is named “Smith”. But now suppose that the actor is named Smith and the actor jumps. Her belief is now true, that Smith jumped. But in this context, even if true, the appropriate belief which ought to have been right is that Smith the character jumped, and the reason she does not get to that is that she is not engaging with the fiction. It is thus necessary to engage with the fiction to understand its content and without it, her thoughts are world directed. If this is so, then there needs to be, as per Hicks, a separate category of fiction directed thoughts.

The problem with both the counterexamples is that they confuse fiction as intrinsically related to the content and the fiction being merely a contextual cue to infer content. This is because the problem is, that Jill says the things she does is either because she hasn’t really understood the content of the film because of inadequate focus on contextual cues or because she has a focus on the W-properties of the film, rather than the R-properties. What is more, due to inadequately picking up contextual cues, her beliefs that “Smith jumped” and someone is Smith is false with respect to the external world, but it is true with respect to the representation and that “Cruise jumped” is false. Matravers can reply that Hicks gets this wrong because he presupposes that make-believe is essential to fiction. If he does not do this, and one is neutral with respect to the fiction/non-fiction distinction here, then all it shows is that “Smith jumped” and “Someone is Smith” are in fact true with respect to the representation (even if it is a work of non-fiction, say it’s a documentary where Jones does not really jump) and false with respect to the either a confrontation relation (since there isn’t any) or with respect to the W-properties which are linked to a confrontation via the causal link between the original event, the camera and the film. The same applies to the case in

which Jones does jump, where her belief that “Smith jumped” is false with respect to the purported content of the representation where Jones is not in fact named “Smith”.

Matravers’ account also is in line with what I have mentioned above from Gopnik. Our capacity for fiction is not *sui generis* but is intricately linked to our capacity to think about past and future possibilities. In order for that to be so, the line between fiction, the content of our counterfactual statements and possibilities needs to be blurred so they are part of the same chain of reasoning, where sometimes, we might not know whether what we are thinking about is actual or merely possible. It also provides a unified account of our ability to think about what is not present in front of us and entertain thoughts in the absence of stimuli, where the fundamental distinction between representations and confrontations allows us to think about an object in its absence. This is in line with the intuition that we can in fact entertain a thought about an object with that object in mind and yet not having the object itself magically transferred into our heads. If this is possible, there is no reason why objects of fiction are part of the same category as the content of our thoughts when we think about objects in their absence.

As I elaborate in the next chapter, we are not only able to think of our mental representations with respect to their W-properties, but also with respect to their R-properties. This means, for instance that when we entertain a thought about an object in its absence, with respect to its W-properties of my mental representation, the object in the external world forms part of its content. But we can also think of the R-properties of the object. For instance, we can also have higher order thoughts on our experiences of shapes, colours, tastes, sounds etc. In that respect, our mental representations not only can represent the external world, but also themselves. The properties of our experiences and those encoded by our mental representations are mind-dependent and which we also might have abstracted from the world. It is these properties that can also form the content of our mental representations, especially when the environmental input is absent. It is with respect to R-properties of mental representations that fictional objects get their content, which is why they come to be of the same category as some of the content of our mental representations in representation.

### **6.1 How Matravers’ Account Fits in With Mine**

This brings me to my account and how it falls in line with what has been presented above. Let us take a case of representation of objects in mental time travel (“the capacity to relive past events and imagine possible future one” (Corballis 2013)). Imagine a case where we need to refer to the representation of an object while planning or in memory and then keep

track of it by means of attending to it (for instance when someone loses concentration while thinking about something, one cannot track it anymore).

Firstly, what would be common to cases of tracking in space as well as mental time travel, or even representations and confrontations, are not concrete objects, but mind-dependent properties like sensory, emotional, motor and verbal properties encoded by our mental representations (Thagard 2019) which also include our abstractions of properties from the external world. Unless humans have the ability to abstract features from concrete objects, mental time travel would not be possible, at least when it comes to thinking about the very same objects one finds in space. One reason would be to be able to recombine properties in thinking of various possibilities to allow for planning. Unless the recombination of properties by means of recombining mental representations is possible, objects can only appear as how they have been perceived and in no other way. And another reason, as mentioned above, would be the simple reason that it would be impossible to transfer a concrete/distal object into one's head and therefore memory! The third reason is that in both cases of spatial and mental time travel, one is not just pretend tracking but actually tracking features. Moreover, unless one cannot explain tracking in one's mental map without being able to pay attention to that object, in a way an eagle zeroes in on its prey. This does in fact sound like being able to refer to something, when we deictically pay attention to an object in referring to it. As a preliminary, this does not unsurprisingly look like referring to a unified mental representation, in which simpler mental representations that encode properties have been bound through neural binding. This is perhaps why realists do in fact come up with the theories they do of being able to refer to fictional characters. Furthermore, if tracking is not pretense, this mode of reference should not be pretense too. It would be highly implausible that the reference to mental representations is a form of pretend reference, even in mental time travel, even more so given the fact that the sphere of representation does not involve make believe.

One further advantage of this is that looking at the content of fictional objects in terms of mind-dependent properties is that this explains why other social objects can be seen in a similar fashion and unifies an account of how to think about things which are not really present in the external world, which do nevertheless exist, like money and national borders. The unification is because there is only one set of cognitive processes which accounts for both without any reference to make believe as well as one single account of content, which comes under the broader category of Matravets' representations. If we think of money too like fictional objects and bring it under the category of representations, then we merely project R-properties of money onto external objects like paper (which also sounds like an

animist thinking of a bird as a god, of looking at a film and thinking of Christopher Reeve as Superman). Money does not exist with respect to the external world and is really just paper, we still think of it as money without pretending to do so and we can also think of money in the abstract, in the absence of the paper. This is not unlike the stump-bear case, where the stump is seen as a bear in Walton's picture because of a prescription to imagine due to rules of generation. If Walton is right, these two phenomena will be seen as radically divergent, and arbitrary because in both cases we think that there is something which is not really there in the external world but we have pretense only in the case of the stump but not in the case of money. It is arbitrary because pretense will give a different output in terms of ontology for two phenomena which are starkly parallel, for which we somehow use different cognitive processes. Moreover, in what looks like a parallel to the money case in a "game of make believe", there can be cases in which children think of an imaginary line on the floor as a boundary and even after the game, they are scared to cross it, even though the game is over and no one is pretending anymore. Why should not be analogous to our thoughts about borders between countries and why should this be considered make believe and not borders or limited liability companies? If it is absurd to say that money does not exist merely because it involves the imagination, it is also absurd to say that fictional objects do not exist because we exercise our imagination.

## Chapter 6

### The Metaphysics of Fictional Entities

#### Bundles<sup>37</sup>

and it trees our forest,  
a bundle in our little aching heads.  
further and below, in the bottom here,  
in a bottom heart plinth a sear of pain, a death.

moving ahead a surging still, a surging surging still  
and reversal of past passive and lying wait  
and surging still of life and semi flamboyant wit  
and a sort of self sardonism.  
and yet beyond ascends the will  
above these plains  
dotted houses  
and cattlestock.

all brought together  
in one fell swoop  
and a wait for gladness

In the introduction, I have already outlined some of the motivations for why fictional objects are constituted by mental representations, the content of which are mind-dependent properties. The question that I suggested that ought to be pursued is not Thomasson's question "if there were fictional objects, what would they be that makes best sense of our literary practices?" but "if there were fictional objects, what would they be in the light of our cognitive mechanisms?".

---

<sup>37</sup> Original Poem (2011)

In Section 1 of this chapter, I outline Anthony Everett and Timothy Schroeder's (2015) view that fictional characters are ideas, which are mental representations. However, they face problems in their theory regarding the content of the ideas, since ideas themselves cannot possess properties like *being a detective* etc. that are attributed to fictional characters in fiction. First, to signify my departure from their account, I reiterate my master argument for the existence of fictional characters from Chapter 2, which is a modification of van Inwagen's (1977) account. Second, while still holding fictional characters to be ideas, I propose that the best candidates for the content of mental representations in the case of fictional characters are mind-dependent properties. This, I believe, is hinted at by Everett and Schroeder themselves, that is, "the sorts of properties one would expect of distributed systems of contentful mental states." (Everett & Schroeder 2015: 313) I also argue that a fine grained view of bundles of representations constituting fictional characters will further solve problems for Everett and Schroeder. In effect, I propose that my account of fictional characters that are in fact ideas, but are constituted by bundles of mental representations and that each bundle gives rise to a unified mental representation. This is a modification of Everett and Schroeder's account. In Section 2, I further argue that bundles of mind-dependent properties encoded by mental representations are good candidates for the content of mental representations, since it accounts for intuitions that fictional characters do not have an internal structure like substances. I also argue that if fictional characters do "flit in and out of existence"<sup>38</sup> (Thomasson 1999: 22), then the most likely explanation for that is that they are mental entities. In Sections 3, I elaborate on interest relative identity with respect to fictional characters and try to account for cases based on that. In Section 4, I examine how reference to bundles of mental representations is possible. I begin with Simon Garrod's (2011) exposition of referential processing with respect to anaphora. The hypothesis is that there are situation models, which are surrogate- representations of the world in which we map such referents. Using this, and Lauri Karttunen's (1976) work on discourse referents, I argue that reference to bundles of mental representations is possible. In Section 5, I put forth Stainton's (2006) exposition of Chomskyan theories of reference, arguing that idiolects are primitive and are properties of the brain and not dialects and languages as are understood by common sense as well as by some parts of mainstream philosophy of language. I propose that reference to mental representations is possible by means of our idiolect. Using this, as well as Ermanno Bencivenga's (1983) view on verbs like "refers (to)" as intentional (*sic*) transitive verbs which takes reference to not necessarily involve concrete objects but also entities in the mind, I propose that reference to mental

---

<sup>38</sup>A particularly apt description Thomasson (1999) uses but unfortunately rejects that fictional characters "flit in and out of existence".

representations is possible by means of our idiolect. In Section 6, I take up objections to my claims (which are modifications of Everett and Schroeder's account). The first key objection I take up is the phenomenological objection, namely that we do not think of mental representations nor of fictional characters as ideas when we think of fictional characters, we think of them as flesh and blood people and creatures. In the light of this, I reiterate what I argued in Chapter 4 that facts about reference and facts on aboutness diverge, and aboutness is explained in terms of more basic facts and relations. I also take up a Fregean objection to my account, the objection being that meanings will turn out to be subjective. Using Ray Jackendoff's (2002) criticism of Frege, I argue that public and shared meanings are possible. In Section 7, I lay out one key advantage of my view where I try to answer Anthony Everett's (2005) crucial objection against artefactualism from the metaphysical indeterminacy of abstract fictional characters.

In the course of my thesis, I have spoken about fictional characters as constituted by tokens of bundles of mental representations. I have also spoken that the best candidates to account for their content are mind-dependent properties. To recapitulate, apart from agreeing with Everett and Schroeder that fictional characters are ideas, my key claims are :

Tenet 1. The content of mental representations with respect to fictional characters is given by bundles of properties of mental representations. The properties are properties such as *being a thought of being red, being an appearance of being red, being a conception of thought of being a detective, being an experience of something's looking square.*

Tenet 2. A fictional object can be multiply realized as constituted by different mental representations, relative to time  $t$ , functional mind  $m$  and contextual convention  $c$ , a fictional character is constituted by particular token of a bundle of mental representations that form a unified representation.

Tenet 3. A bundle of representations is an instance of representations linked by neural binding. The bundle itself forms a unified representation. Neural bindings are brain processes and insofar as mental states are types of events that are multiply realizable in a functional mind, they are types of events. Each instance of a neural binding is a token event and so is a bundle of mental representations a token event in the brain of someone thinking of the character.

## **1. Fictional Characters as Ideas**

Everett and Schroeder (2015) propose a realist view that fictional characters are ideas (mental representations), which are mental states that have content. For them, an idea is a

concrete entity, since it has its origin in a mind in a particular place and time, spreads from one mind to another and even that a single mind “carries” an idea from place to place. They say: “Ideas are not the contents of beliefs or desires: not senses, not propositions, not possible states of affairs, not facts. Rather, ideas are mental states bearing contents.” (*ibid.*, 276) Thus what they say is ideas are mental states that bear content, not the content themselves. Darwin invented the idea of evolution by natural selection, but not the content of the idea itself, because that is the fact of natural selection.

Ideas can not only be created but also get lost. For instance, as per Plato, there was an idea to make Socrates escape Athens which specified how it would happen, so that he could escape his death sentence. If there was such an idea, it has been lost to us and ceased to exist. Thus:

There are ideas. They are invented, get spread, get forgotten, and eventually become lost. They have the sorts of properties one would expect of distributed systems of contentful mental states. And particular ideas will have particular properties that are specific to them, properties that often depend upon the content, or context, or influence of the idea. But, of course, we hope that ideas are more than an uncontroversial part of any sensible ontology. (*ibid.*, 278)

Thus, the reason why they take fictional characters to be ideas is that ideas have exactly the same sorts of properties that fictional realists take fictional characters to have. For instance, ideas are created, they cease to exist, they can be more original or less original, they can be inspired from other ideas, real people, objects or events. Like characters who are one dimensional and many dimensional, ideas can be simple or detailed. Thus, more specifically, for Everett and Schroeder, fictional characters are fictional person ideas, that is, “ideas for telling a story as if there were a particular person with certain features.” (*ibid.*, 279) In the light of this, like artefactualists, they make a distinction between character-properties attributed to the idea, like *being one dimensional, created, original, inspired by a real person* (Dr. Joseph Bell in the case of Conan Doyle) which they possess independently of the context of a story, and person properties, which Holmes has like *living on 221B Baker Street, being a detective* etc. attributed within the context of a story.

For Everett and Schroeder, ideas are “spatially discontinuous individuals”<sup>39</sup> (*ibid.*, 276) like species (if species are taken as individuals). An idea begins when a mental state is tokened in a particular mind. At this stage, the mental state bears part of the content of the idea or the entire content and it is a single mental state that constitutes the idea. When the idea spreads

---

<sup>39</sup> I worry about ideas being called “individuals” since I don’t know how much of this usage is metaphorical. It’s like the brain-in-a-vat problem or Descartes’ Demon but in this case I don’t know if the demon of metaphor is deceiving me. Perhaps I’m being paranoid. But I’d rather say “entities” and be more specific and call them processes. Folk psychology is deceptive.



from one person to another, there are two mental states, the first and the second person's, which constitute the idea. Furthermore, while they are open to the view that a whole idea is fully located at a particular place where tokened, their preferred view is that an idea is composed of parts, where each part is part of a network of attitudes. These attitudes are joined by communicative chains that have the same content (I prefer the former view, since it allows us to say things like "my idea of x at time t, my idea of x at time t<sub>1</sub>. It feels odd to say here that I am thinking of only a part of an idea.) Moreover, in line with their view, it seems to be that it suggests that failed theoretical posits like phlogiston and witches can also be considered as ideas, something they themselves hint at when they talk about the perpetual motion machine as a mere idea. (*ibid.*, 288)

The individuation of ideas, according to Everett and Schroeder, is interest relative and context sensitive and we even sometimes misindividuate them. In some contexts, where people come up with ideas with the same content, like Newton and Leibniz both inventing ideas of calculus, the ideas are individuated by origin and are "counted as different ideas". (*ibid.*, 276) So there would be a distinction between Newton's idea of Calculus and Leibniz's idea of calculus. In other contexts, ideas are individuated by content, and the same way in which we individuate many other things, we individuate these ideas in an interest relative manner. For example, Everett and Schroeder say that we count two people reading two copies of *The Phantom Tollbooth* in one context as the same book because we are interested in it with respect to its content. In other contexts, we count it as reading two different books with respect to individuating it as two copies (Everett and Schroeder also propose that stories are ideas). In the light of this, they say: "Likewise, in some contexts, where we are really only interested in the content of an idea, we may happily bundle together distinct ideas that have the same content and count them as one." (*ibid.*, 277)

Everett and Schroeder go on to account for varying intuitions in the Menard case based on interest relative identity. For instance, as we saw before in the Menard case where there are two word for word identical novels, one by Menard and one by Cervantes. Some say that there are two characters and some say that the characters are identical. This is explained by Everett and Schroeder as a case of interest relative identity, where in the former intuition, the fictional character is individuated by means of its origin and in the latter, by means of the content.

Another respect in which Everett and Schroeder find a parallel between fictional characters and ideas is as follows. In the case of the idea of making Martinis, that is, the idea of making a cocktail by combining gin and vermouth has been modified through time. There are cases where it has been modified to make martinis not as sweet, but it still is counted to be a

version of the same idea, whereas the idea of making what is named a “Chocolate Martini” which involves mixing vodka with cream and chocolate liqueur is counted to be a different idea. In between clear cases and the Chocolate Martini case, there are borderline cases like the idea of a basil Martini where it is not obvious whether it should be counted as the same idea or a different one. So ideas that exist are taken and modified. If the content is modified radically, then it is more likely that the idea will be counted as a different one. Furthermore, there are cases where it is not obvious where it is the same or a different idea. Likewise, when it comes to fictional characters, we see the same sort of phenomenon. An author may write of a detective in the 21<sup>st</sup> Century with a friend named Watson, and shares a lot of Holmes’ characteristics, we will most likely count it as the same character. On the other hand, if someone writes of a farmboy who is a meek simpleton named “Sherlock Holmes” it is likely to be counted as a different character. There are also modifications of the Holmes’ character where it is not obvious whether it is the same character or not.

Everett and Schroeder outline that their view has distinct advantages over artefactualism. While it is unclear how an abstract object is created within artefactualism, everyone would agree that an idea is in fact created. Furthermore, their view accounts for the fact that we can create a character without having an idea for the story in which the character is included. For instance, an author can have a fictional person idea without yet having the idea of the story in mind, or be writing many stories and then the author deliberates on which story to include the character in.

There are two key objections to their view that Everett and Schroeder bring up. The first objection is that fictional characters are like institutional entities like nations, universities and the law, which are public. Thus, the idea of law cannot by itself be the law. Similarly, a fictional character is distinct from the idea of one. Their reply is that we do not think and talk of fictional characters as public institutional entities so as to allow for even one person having created it. For instance, an author thinks of two characters a and b who have a duel over an argument but does not write the story. She might base the character of real people she knows or base future characters she writes in stories on a and b. Nevertheless, in this case, the characters remain private (I shall address the worry that even then we typically think of a fictional character as a flesh and blood person, not an idea in my section on aboutness and reference).

The second major objection is that ideas simply are not the kind of things that can have person properties like being a detective or being cold hearted. To this, Everett and Schroeder have a *tu quoque* reply. They point out that realists too face the same problem and come up with different strategies to deal with it. Platonists face the problem as to how abstract

Platonic objects can have properties like these. To deal with this, Neo-Meinongians like Zalta (1983) make a distinction between two modes in which properties are predicated, where properties like being a detective are encoded by fictional characters but properties like nonexistence and being a fictional character are exemplified by them. Parsons (1980), on the other hand makes a distinction between two types of properties, nuclear and non-nuclear, where Hamlet has a nuclear property of being a prince and the extranuclear property of *being non-existent*, *being a fictional character*, *being possible* etc. Of course though, unlike Neo-Meinongians, they would not say that the ideas exemplify the property of *being non-existent*. Artefactualists too, as we have seen before, give an account of these person properties by saying that talk in fictional contexts occurs is in the scope of pretense. Everett and Schroeder suggest that one may take up some of the solutions in the literature. Nevertheless, they say the following:

We recognize this is a potential problem for our account. But we emphasize that our account is no worse off in this respect than that of the artefactualist. If anything, there is particular hope for our account, since the canonical person properties, such as being a detective, that are attributed to fictional characters are properties that are part of the representational contents of the corresponding ideas. And we note that people are sometimes prone to run together talk of ideas and talk of their contents, as when they talk of Darwin inventing natural selection and Einstein inventing general relativity. (ibid., 289)

In Chapter 5, I have already shown issues with the providing an account of the content of pretense. If these properties are within the scope of pretense, then they are not really properties, but pretended properties, the way Walton talks about pretended propositions. Again, there is no satisfactory account of the content in Everett and Schroeder's account, nor within the existing artefactualist camp.

Furthermore, Thomasson (1999) accuses both Parsons and Zalta of making *ad hoc* moves to adjust their theory to their Neo-Meinongian ontology. One reason why this is so is because there might be a metafictional novel which talks about a non-existent golden mountain versus a novel that talks about an existent one. One way to account for this difference would be to say that the golden mountain attributed with existence has a "watered down" nuclear property of existence and the extranuclear property of non-existence. This move is necessary, since existence is also an extranuclear property of things and there thus have to be a different version of it as a nuclear property. This is the kind of move that Thomasson finds *ad hoc*. Zalta too faces a similar problem, where to account for the golden mountain attributed existence in a novel versus one attributed nonexistence, he would say that it encodes existence but does not exemplify it. But then, Zalta ends up postulating not just two modes of predication but two kinds of properties, especially, as Thomasson points out,

“given that abstract objects can only exemplify extranuclear properties.” (*ibid.*, 103) Again, this is *ad hoc*.

This becomes even more troublesome with Everett and Schroeder’s account. Suppose, if it takes Parsons’ or Zalta’s picture, in the case of metafiction, where an author explicitly talks about the fictional character as a fictional character, it will have to take a property like *being a fictional character* to be a nuclear as well as an extranuclear property, and in Zalta’s picture, exemplified by the idea as well as encoded by it. This will face the same objection of the move being *ad hoc* from what we saw from Thomasson apart from the argument that there would be no criterion to distinguish between the two kinds of properties in metafiction. (Kroon and Voltolini 2018) Furthermore, Kroon and Voltolini point out that Zalta might make the distinction between internal predication of *being a fictional character* from within the fiction and external predication. Nevertheless, as we saw with Thomasson’s objection, since Zalta ends up with two kinds of properties too, he will have to face the same issues as Parsons.

In what follows, I shall try to give a naturalistic account of the content of our mental representations with respect to fictional characters and attempt to solve the problem that Everett and Schroeder face. It looks to me that the reason why Everett and Schroeder have trouble with their account is for two reasons. One is that their notion of ideas as constituted by mental representations is not fine grained enough. And the second is that they do not take into account properties encoded by mental representations that can function as the content of ideas when it comes to fictional characters and account for the properties predicated in fiction.

### **1.1 Master Argument for Fictional Characters**

Before attempting to provide an account for the content of ideas, I shall lay out my master argument for the existence of fictional characters to show where exactly my account diverges from Everett and Schroeder. I argue that fictional characters exist because all statements regarding fictional characters presuppose (MET<sub>FIC</sub>)\*. I shall rehearse the argument I gave in Chapter 2 below for why this is so.

(MET<sub>FIC</sub>)\* There is an *x* such that *x* is fictional and *x* is likened to a real person and *x* is and ought to be treated/counted as a real person for all relevant intents and purposes.

What we saw with van Inwagen (1977) was that literary critical statements like (1), when taken at face value entail (2) (*ibid.*, 302):

(1) There are characters in some 19th-century novels who are presented with a greater wealth of physical detail than is any character in any 18th-century novel.

(2) There are characters.

Thus van Inwagen postulates fictional characters to be theoretical entities of literary criticism. In the context of Yagisawa's (2001) objection to van Inwagen, I modified van Inwagen's argument to come up with my account. Thus, I will begin with Yagisawa's argument:

Yagisawa argues that if literary critical statements are taken at "face value", then a sentence like (3) would entail (3') (*ibid.*, 164):

(3) Mrs. Gamp is the most fully developed of the masculine anti-women visible in all Dickens's novels.

will entail

(3') Mrs. Gamp is a masculine anti woman.

Yagisawa argued that, however, taken at face value, (3') is false, since predicating "masculine anti woman" literally of an abstract object as masculine is a category mistake. This extends to other predicates like fat, female and gin drinking. So while at face value, the quantification is supposed to be over an abstract object, the quantification is over a flesh and blood person.

Jeffrey Goodman's (2004) response on behalf of van Inwagen's view to Yagisawa was by reconstructing van Inwagen's argument as replacing all sentences of literary criticism with "some sentences" (*ibid.*, 137):

P1 There are some sentences of kind  $\phi$  that are truths of literary criticism.

P2 Therefore, some sentences of kind  $\phi$  are true.

P3 All true sentences of kind  $\phi$  entail that entities of kind  $x$  exist.

C Entities of kind x exist.

Thus, he said that only some statements of literary criticism are to be taken at face value like (*ibid.*, 138):

(4) James Tiberius Kirk is a fictional individual

If (4) is taken at face value it will quantify over an abstract object, given that it is true simpliciter.

My response to Goodman was that he misses a general principle like (CA), which is violated when one predicates fat, masculine etc. of an abstract object.

(CA) A predicate that correctly ascribes a property that can belong only to a concrete individual can be literally predicated only of a concrete individual.

The reason for this was, that even with a sentence like (4), the word “individual” cannot arbitrarily taken in the philosophical sense, since the predicate “is a fictional individual” is applied by a literary critic, who uses the sense of individual in this context as per the dictionary definition of person “a single human being as distinct from a group” (Oxford Dictionary of English (2010)). This implies that individual in this sense can only be correctly applied to human person. I made the case that even “James Kirk is a fictional character” cannot be taken literally either, since the meaning of “fictional creature” (and “fictional person” in this case) will be presupposed by “fictional character” and “creature” and “person” cannot literally be applied to abstract objects (since Holmes is a fictional character only if Holmes if a fictional person, Godzilla is a fictional character only if Godzilla is a fictional creature). This presupposition explains why, at face value we tend to say “He is a fictional character” not “it (the abstract entity) is a fictional character”. I had argued, that the only way out of this mess for van Inwagen was to take only the statement “James Tiberius Kirk is fictional” as literally true, and which would also be true in a true/false literature exam. If that was so, then the only way to salvage his theory was to consider statements like “James Tiberius Kirk is a fictional individual” as analogous to “Companies are fictional persons” which presupposed a conceptual metaphor:

(MET<sub>COMP</sub>) A company is artificial but is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

And statements about fictional characters, like (4) would presuppose:

(MET<sub>FIC</sub>) James Tiberius Kirk is fictional but is likened to a real person and is and ought to be treated/counted as such, for all relevant intents and purposes.

Thus, in general, statements regarding fictional characters will presuppose the following:

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

Now when we say that A, which is a truth-evaluative statement, presupposes B, it means that the truth of B is necessary for the truth or falsity of A. It is also necessary for A being a truth-evaluative statement. For instance, presupposing the truth of “The Kind of France exists” is necessary for the truth or falsity of “The King of France is bald”. We find statements in fiction that are true and false and thus truth-evaluative. For these truth-evaluative statements to be true or false, they have to presuppose the literal truth of (MET<sub>FIC</sub>)\*. Thus, despite involving a metaphor, (MET<sub>FIC</sub>)\* is literally true. So if (MET<sub>FIC</sub>)\* is literally true and says that there is an entity that is fictional and likened to persons, and this is presupposed for all statements about fictional character, then we are led to the conclusion that fictional entities exist. Also, when we combine the conventions of fiction with the truth of (MET<sub>FIC</sub>)\*, it would imply that statements in the fiction would be true according to the fiction. This accounts for why fictional statements can be literally true.

What follows from this is that mine is a uniform account of fictional statements, where unlike van Inwagen, who says that while it is predication that takes place in literary critical contexts, predication does not happen within fiction, but ascription. Moreover, while in literary critical contexts, fictional characters *have* the properties predicated of them, in fiction, they *hold* the properties ascribed to them. My account has no such distinction as all statements about fictional characters uniformly presuppose (MET<sub>FIC</sub>)\* and if this is so, there is only predication, but no ascription.

There are two points that are pertinent here. Firstly, that it is in place to ask what constitutes x in (MET<sub>FIC</sub>)\*. My claim is that x is constituted at t, when a subject y is thinking of x, by a token mental state, which is a mental representation. This is on the lines of Everett and Schroeder’s account. Furthermore, the same type of mental representation can be tokened in another person’s mind at time t, and can constitute the same character. In asking this question, we are asking, what are the constituents of the world that comprise it? This is similar to asking what is a bank note composed of? The answer to which is: paper.

At another level of explanation, presupposing (MET<sub>FIC</sub>)\*, with respect to our interests and purposes, we can ask: What makes a fictional character that particular fictional character and not another fictional character? In this sense, the identity is interest relative in that it is not an intrinsic property of the mental representation that defines the identity of the fictional character (it may guide and constrain us, however), but it is our interests that determine, with respect to the content, whether the mental representation is to be counted as is a particular character or not. In other words, fictional characters are like games of chess in that the conventions (what Searle (2006) calls constitutive rules like “x is counted as y in context c”) determine their identity. So in context c, a such and such mental representation x will be counted as y.

Moreover, I said above that even saying “x is a fictional character” presupposes that “x is a fictional person”. Without the presupposition, applying the predicate “is a fictional character” literally is a category mistake. So one cannot ask or say “Why are we then thinking about the *same* character?” without presupposing (MET<sub>FIC</sub>)\*.

## 1.2 Mind-dependent Properties

As we saw in Everett and Schroeder’s account, thoughts and mental representations cannot literally bear properties that literally belong to concrete objects. For instance, my thought cannot have the property *being round*, because objects are round. However, I argue that a thought can represent that property and when it does so, the property of the thought is *being a thought of being round*. Literary characters can never literally possess properties like *being round*. So, the question is, what are the properties we think of when we think of literary characters that accounts for the content of the mental representations that constitute fictional characters?

Furthermore, these properties can be properties encoded by any mental representation. So, when an apple looks round to us, the representation can have the property *being an experience of being round*. I will use “feature” interchangeably with “property”, which is the way something is. In the case of fiction, my claim, more specifically, is that they are bundles of properties which are mind-dependent, causally linked to the external world (which includes our bodies and brains) and carry information about it.

If we take a fine grained view of mental representations, we find that ideas are constituted by bundles of mental representation, which in turn form a unified representation in a token mental state. Mental representations are defined by Thagard (2019) as “not things or



properties but rather processes consisting of patterns of firing that result from the synaptic connections among neurons. These learned or innate connections enable neural groups to produce useful patterns of firing in response to inputs from the environment or other neural groups<sup>40</sup>. (Thagard 2019: 24)

Thagard uses binding of representations and binding of sensory, emotional, motor and verbal features interchangeably, but there is more to it than meets the eye. So Thagard (2019) talks about semantic pointers, which are constructed by repeated binding of mental representations<sup>41</sup> which are formed when representations like sensory, motor, emotional and verbal representations are bound. Thagard says:

Semantic pointers combine the flexible capabilities of neural representations understood as firing patterns with the symbolic capacity that enables people to accomplish language and reasoning.” (ibid., 29)

For example, the concept of a cat can be viewed as a semantic pointer that binds sensory features such as what cats look and sound like<sup>42</sup>, motor features such as how it feels to pick up a cat, emotional features such as how much you like cats, and verbal features such as that cats are a kind of mammal. (ibid., 30)

In effect, what binding does is to take information in neural patterns and compresses it in a way music is compressed by a computer, and yet keeps the original songs intact. So I can talk about bundles of mental representations, and the information encoded in the patterns are these features, which in turn represent features of objects. It also looks like these features are essential to a particular representation.

What binding does is to unify features of objects that groups of neurons encode- like yellow, long and round for bananas. This explains why we do not see bananas with disconnected and disparate features but as unified representations. Binding takes place with other sense

---

<sup>40</sup> On representation, Thagard also says:

Neural groups provide much richer representations and processing than is possible with individual neurons. Brains represent the world by having groups of neurons that collectively encode inputs from the world through the excitatory and inhibitory links between neurons. These connections produce a form of representation different from the words, pictures, and other images that are familiar in everyday experience. In neural groups, the representation of something is distributed across all the neurons by virtue of their interconnections that produce patterns of firing. What matters is not just the pattern or rate of firing of each individual neuron but rather the pattern of firing across all the neurons in the group. The resulting patterns are both temporal and collective, like the performance of a band that requires the coordinated activity of many musicians and their instruments. (ibid., 24)

<sup>41</sup> On Semantic Pointers, Thagard says:

The term “pointer” in “semantic pointer” was inspired by the use in some computer programming languages such as C of objects that refer (point) to values stored elsewhere in memory. This analogy is limited because in semantic pointers both the pointing and what is pointed to are dynamic processes of neural firing, not just objects and values. A more dynamic comparison would be to the processes when a nod from one musician serves as a pointer to another musician’s playing. The appendix to this chapter provides comparisons of semantic pointers with more familiar ideas about mental representation. (ibid., 34)

<sup>42</sup> Italics are mine. These are the mind-dependent properties I talk about.

modalities too, and explains how we hear a unified sound rather than separate instruments disjointedly when we hear a band. The same happens with taste, when we taste a sandwich's unified flavour rather than individual flavours of cheese, bread and tomato.

One of the explanations for binding that Thagard prefers is Chris Eliasmith's. Suppose the pattern of firing of a group of neurons represents a particular feature of a banana. A new neural group, which may overlap with this then performs the binding by computing a mathematical function called "convolution". Thagard explains convolution by means of the metaphor of hair being braided, when "previous representations are twisted together in a way that preserves much of the original representations so that they can subsequently be taken apart. (*ibid.*, 27) To take the metaphor in the way of explanation, strands are brought together but also can be taken apart to retain what they were before. So convolution brings together representations so that new firing patterns are formed and which can again be unbraided back to what they were. Furthermore, these representations can be manipulated to form an image of a monkey eating a banana as well as form inferences that the banana is sweet. The brain can also produce bindings of bindings, such that binding is recursive and even come up with "whole stories that introduce additional complex relations." (*ibid.*, 28)

Thus, mental representations are patterns of firing of neural groups. The properties are merely information encoded in them, which forms their content. Thagard (2019) says "Representation in brains is therefore an emergent process resulting from the interactions of many neurons, not just the sum of the limited representational capability of each neuron." (*ibid.*, 24-25)

In the light of the above, this is where a more fine grained view of mental representations aids Everett and Schroeder's account, especially when it comes to the content of the representations when it comes to fictional characters. If we look minutely at how mental representations are constituted, we will find that more complex representations consist of more simple mental representations bound by neural binding. Each simple representation encodes a feature, as we saw with Thagard's account. If that is so, then as the properties encoded by simpler mental representations keep changing (for instance, when an author keeps changing his mind about what properties his character should have), there are different bundles composed of simpler mental representations. Each bundle forms a unified representation at time  $t$ .

Five important points ought to be noted here.

i) If there are for instance, sensory features, they by definition are essentially sensory features. If they are essentially sensory features in the brain, they cannot by definition exist outside of one's brain. This applies to each and every feature encoded in mental representations. If they are essentially features in the brain, then they cannot exist outside the brain. The question of Platonism with respect to these is worthless. My preferred view is that these are token mental properties, which are in line with the view that token mental properties are tropes. (Maurin 2018) Thus, two tokens of the same property *being the appearance of being red* can be identified in terms of exact resemblances. Two tokens of the same property will exactly resemble each other but will not be numerically identical<sup>43</sup>.

ii) It is also implausible to think that our predicates only predicate properties of the world as described by physics and do not refer to any properties in the rich complexity of our mental life, even more so since it is standard in cognitive science to think that our minds do not mirror the world but reconstruct it. (Jackendoff 2002, 2006) The retina has no objects and no external location in its "ontology", but only registers distinctions like "dark point in bright surround at such-and-such a location on retina." (Jackendoff 2006: 226). The brain then is given information by the retina, and the parts to which the retina does so directly merely have line and edge detectors, "all in retinotopic format (Hubel and Weisel 1968) – but still no objects, no external world. And this is all the contact the brain has with the outside world; inboard from here it's all computation." (Jackendoff 2006: 226-227) The brain then constructs a percept, which is non-linguistic and is tracked by us, in which individuals have been individuated. Mental representations like percepts have *sui generis* properties over and above what the external world has, properties which we can identify, reflect on and talk about. For instance, the unified flavour of the sandwich which is an experiential property is such a *sui generis* property.

iii) It is even more implausible to think of these mind-dependent features as eternal since they will then be nomological danglers in the physical world, that is, they won't be part of the causal laws of the world. Moreover, in our process of language acquisition, we come to

---

<sup>43</sup> On exact resemblance, Miller (2019) says:

Most bundle theorists take (exact) resemblance to be an internal relation, and hold that resemblance between tropes is determined by their primitive intrinsic nature. Following Maurin, we can take exact resemblance 'to be formally characterized as an equivalence relation, i.e., as a relation that is symmetrical, reflexive, and transitive. As such, exact resemblance partitions the set of tropes into mutually excluding and non-overlapping classes; classes functioning more or less as the traditional universal does' (Maurin 2018) (*ibid.*, 6-7)

learn and know the meanings of a lot of these predicates, especially when it comes to fiction, easily and rapidly without the kind of complex inferences we need to make for entities that are not directly perceivable, like quarks. My argument is not from the Eleatic principle which Cowling (2015) formulates as: “Necessarily, some entity  $x$  exists if and only if  $x$  is causally active” (Cowling 2015: 307). The point is that if they are of a different sort than the physical world like ghosts or mental substances, it would be less parsimonious (than to say they are essentially sensory, motor etc. features like I said in point (i) and therefore not Platonic or mind independent) to explain how we acquire knowledge of them so easily in the process of language acquisition, given our language faculty. There would also be no explanation (Field 1989) of what mental faculty is responsible for gaining this knowledge so quickly in childhood. We cannot come to know of, or at least know of so easily, what we cannot in principle causally interact with. But we do in fact do so, which makes my explanation more parsimonious. Moreover, if some of the mind dependent features are known this way and are not Platonic, there is no further reason why other more complex ones which we learn of later in life are Platonic. There is also the issue that Thomasson and van Inwagen face in the literature, namely, how can fictional objects be abstract causally inert entities? It’s a serious problem. Analogously, to think of mind-dependent features as independent entities is a serious problem too.

iv) Thagard does talk about repeated bindings forming concepts, including that of a unicorn. However, not all mental representations are concepts. Moreover, suppose we take a prototype theory of concepts, where a concept is represented in terms of typical features of an object, where members of a category are to be more typical of the concept if they possess typical features. In that case, we will use the concept of a unicorn to identify, based on the content, which mental representations are typical instances. This is one reason why I do not think that mental representations that constitute fictional characters are just concepts, since concepts are what are used to categorize instances of fictional characters.

v) It should be made clear at the outset that I do not take mind-dependent properties or for that matter, properties of any kind to be functions from objects to propositions, where propositions are sets of possible worlds or equivalently, properties as functions from worlds to sets of objects. The reason for this is that if like modal fictionalism, an account of possible worlds can be given by means of my account of fiction using conceptual metaphor, in which one understands abstract concepts in terms of more concrete ones, then one cannot appeal to possible worlds for an account of fiction. For instance, for modal fictionalism,

statements like (5) can be analyzed as (6) where PW is a “possible worlds fiction” (Rosen 1990: 335):

(5) There might have been blue swans.

(6) According to PW, there is a universe containing blue swans.

Now if my analysis is correct, that statements about fictional objects presuppose statements like “There is an  $x$  such that  $x$  is likened to a concrete object”, since statements about fictional characters involve ontological metaphor, one can analyze statements about possible worlds as

(7) There is an  $x$  such that  $x$  is likened to a concrete universe containing blue swans.

I have already pointed out in Chapter 4 that statements about possible circumstances cannot be taken literally, since literally, circumstances happen and transpire concretely. So if an analysis of possible worlds is given based on an account of fiction like mine and if my analysis is in fact correct, then I cannot give an account of properties based on possible worlds, metaphysically speaking.

### **1.3 What are Mind-Dependent Properties?**

In what follows, I shall explain further what the mind-dependent properties are that I am talking about. Properties are the way something is. But there are also ways in which an object looks or seems to us, typically in our percepts (to this it can be added how it tastes, feels, sounds, smells and other qualia etc.) In Thagard’s account of the binding of mental representations, we saw that mind-dependent features like *what cats look and sound like* (sensory feature), *how it feels to pick up a cat* (motor feature), *how much you like cats* (emotional feature) and verbal features like cats being a kind of mammal are encoded in mental representations and bound by neural binding. Furthermore, there are ways in which we think of things, conceive of things, ways in which we imagine things. For instance, as we saw, there is a way in which an object looks red to us, or there is a way in which one line seems longer than another in the Muller-Lyer illusion even though both are of the same length. That is, from “there are ways objects look and there are ways that objects seem to us”, it can be inferred that “there is a way an object looks and there is a way that an object seems to us” and that these ways exist. From the general statement, it can also be inferred that there are ways in which an object looks or seems to me.

Of course, if an object looks red to me or to us, it does not follow that it is not red, since the way it looks to me can carry information about the external world. Similarly, if there are ways in which we think of somebody as a detective, it doesn't imply that he is not a detective, even though there is also a thought that carries the information *being a detective* (which has the derivative property of *being the thought of being a detective* like an appearance of red has the derivative property of *being the appearance of being red*, the Muller-Lyer illusion having the property of *being an appearance of being square*) that indicates that the other person is a detective and that is also the way we think of the person.

We can also think of the person with the property *being a clown* (with that thought bearing a property of *being a thought of being a clown*), even if he is not, since that is a way we can imagine him. It is not the way that person is, but a way we think of or imagine him. Our thoughts and sensations, in effect might gain their content from the world by being causally linked to it, but the thoughts and ideas themselves seem to have magically, as it were, gained these properties. These properties are properties of our minds and taking physicalism as true, our individual bodies and yet can be partly individuated by properties they are causally linked to in the external world (like *being the appearance of being red* is linked to *being red*). This also explains why when we think of the moon as made of green cheese, the statement "The moon is made of green cheese" does not entail that "Moon has the property of green cheese" (Edwards 2014<sup>44</sup>), but it does entail "The moon has the property of being thought of being green cheese by x", which in this case is a relational property of the moon and x.

This becomes clearer in cases like that of money. In the external world, there is nothing over and above paper, but we think of it as being a certain way (*being a thought of being a five dollar bill*), that is, as *being a five dollar bill* that we represent to ourselves due to constitutive rules that allow us to think that way. One can, moreover, refer to these deictically, as this way, and that way. It is these sorts of ways that we can refer to by means of deictics like this and that I call *mind-dependent properties*. There is, as it were, an inner environment that indicates properties in the world as well as inside us<sup>45</sup>. There can also be

---

<sup>44</sup> Edwards (2014: 89) points out this absence of entailment.

<sup>45</sup> On the inner environment, Dennett says:

Mutations equipped with such benign inner environments would have a distinct survival advantage over merely Skinnerian creatures in any exiguous environment, since they could learn faster and more safely (for trial and error learning is not only tedious; it can be dangerous). The advantage provided by such a benign inner environment has been elegantly expressed in a phrase of Karl Popper's: it "permits our hypotheses to die in our stead." (Dennett 1981: 77)

ways of ways of looking and seeming and thinking which are not imagistic showing the capacity of the mind to abstract at a higher level. In fact, this can be extrapolated for more abstract properties too, since there are ways abstract thoughts about time or love are conceived or thought of.

This is made clear by Frank Jackson, who also seems to talk about these sorts of properties, which “*carry traces of the kinds of things*” (Jackson 2015: 1) at least the ones which are causally linked to the external world and are about its properties:

There is, however, a way of thinking about the famous examples that makes no mention of intensions, senses, meanings, etc. and is in terms of reference understood in causal terms. It insists that we can have thoughts about properties as well as about things, and the explanation of how this is possible goes back to the way our brains, and what happens inside us more generally, carry traces of the kinds of things to be found around us. Indeed, one way of thinking about a causal theory of reference is as a theory very roughly based on the idea that what happens inside us carries information about the properties instantiated in and around us. To use the old gas gauge example: the position of the pointer carries information about the level of gas left in the tank by virtue of its causal co-variance with that level. (*italics mine*, Jackson 2015: 1)

Jackson also talks about how the distinction between the predicates “creature with a liver” and “creature with a kidney” can be explained in referential terms. Thus, even though the predicates are co-extensive, both terms refer to different properties. This account also, in my view, partly explains what mental representations like *being a thought of being red* are about. Our thoughts about properties are partly individuated by their causal relation to the external world. So the mental representation is about the colour red by virtue of it being causally related to the property in the world either to a red object in my egocentric space or via a causal link from the past. Nevertheless, things aren’t so simple. Since the mind holds colours as constant to compensate for variances in light (like the colour of a banana will be constant amid different light wavelengths) and that two different reflectances can appear the same shade of red, colour is neither a reflectance property nor a property of wavelength, even though these form conditions for our perception of colour. (Lakoff and Johnson 1999: 20-21) Furthermore, like we saw above, our retina takes in very little information from the world and our brain actually goes to reconstruct the inner visual world through computation and percepts are radically different from how the world is, with clear boundaries and shape properties which are not present in the external world. (Jackendoff 2006) In the light of the science, the causal story becomes more complicated, since part of the causal link are the mechanisms in the mind which make an object look red to us. The content of the mental state then, is not just determined by reflectance of the external object, but also *sui generis*

properties that have been computationally generated in the mind, that makes our thoughts thoughts about the property red. But if we ask what the representation is a representation of, like in the case of unified taste of the sandwich that we saw, the representation is not only of properties in the world, but also intrinsic properties of mental states. In effect, the representations not only represent the world but also themselves, that is, they are also reflexive. This accounts for how there can be higher order reflection and thought of our experience of colours, tastes, shapes and the like, especially when there is no environmental input. It might look to us that the representations are only of and about the world, but it is a dispositional property of the representation that it is also about itself, so that when higher order thinking on it takes place, it can be about itself. It must be noted that a child can think about fictional characters despite not possessing higher order thinking, since a property like *being a thought of being red* is the property encoded by a first-order thought. It doesn't imply a thought about a thought, but has the dispositional property that one can think about the thought. And of course, children do have first order thoughts<sup>46</sup>.

Furthermore, one might say that if there are ways objects look and seem to us, then that entails that there is a concrete external object of which the way of looking or seeming is. But this isn't correct. This can merely entail that there is a way in which the experience occurs to us. And that the mental representation exists which happens to represent these properties as their content. That is because we can have a second order contemplation of the mental state. So we can think of how the object looks to us. But we can also contemplate what it is like to experience the object in the presence of the object as well as when the object is absent in our immediate surroundings, when we recall our memories of it. So, in this case, "something" in the way *something* looks like to us is ambiguous between a mental representation (we might think it is a concrete thing but it is just an experience) and the concrete object in the external world. There can in fact be such a way which seems like and is about a concrete object without itself being a concrete object. When we imagine or hallucinate something,

---

<sup>46</sup> Everett and Schroeder very clearly write that ideas are constituted by token mental states and are private. They're instantiated in the brain. One might mistake them to be public in the way concrete objects in the external world are, but there is a problem. What does public mean? But surely, ideas are not concrete objects but are constituted by mental processes. Furthermore, surely ideas are not floating around outside people's heads. Perhaps one might interpret what ideas are as a logical behaviorist would interpret them to be? But that is incredible. Plus Everett and Schroeder are providing a more credible alternative to non-spatiotemporal views that artefactualism holds to satisfy anti-realists about fictional characters and in fact say ideas are spatiotemporal. So they're not Platonic entities either. I have accounted for their public individuation in terms of causal links too. Moreover, my account is distinctly better than accounts consistent with methodological solipsism, since I talk about part individuation by means of causal links.



there is a way that my experience of the hallucination seems like a dagger. In fact, after I have a hallucination, I can reflect on my experience of the hallucination and say, “My experience of the dagger looked to me a certain way” and go on to describe it. And the reason why my experience of the hallucination is about a dagger is because, in my view, this way, that does exist and is encoded by a mental representation “is likened to” a concrete object via conceptual metaphor. This explains how the mental representation can represent these properties, that is, via a conceptual metaphor and ‘the way’ obtains its representational properties in virtue of being likened to a concrete object. Since there is a conceptual visual metaphor, it is not literally a dagger, but the mental representation presents to us the experience as that of a dagger. There can also be other ways in which mental representations are experienced by us. In effect, this fits the definition of properties, that they are ways something is. The representations are experienced as a “something” because of an ontological metaphor.

It also looks like when new unified mental representations are formed from simpler representations, these mind-dependent properties are recombined, since they are supervenient on the representations. For instance, the recombination of properties can form and be properties of imaginary representations (like a unicorn being a combination of the property of *being thought of being white*, *being thought of being a horse* and property of *being thought of being a creature with a horn*). Again, the unicorn does not appear to us as a thought but as a unicorn (as a concrete creature) because the thought is understood or conceived in terms of an ontological metaphor, with respect to being “likened to” the concept of a concrete substance. As we saw, it does not also follow that the mental representation itself does not exist, since we saw that there was a conceptual metaphor involved. So while it may be false to say literally that the imaginary unicorn is white, it is true to say that there are ways that the representation seems that are likened to a concrete object and that is why we presuppose an ontological metaphor and say that the mental representation that looks like an object, is sharp (say in the case of an imaginary dagger). These acts of the imagination seem possible because the “decoupling” of the content is possible:

...the content of Sally’s pretense representation (e.g., THIS BANANA “IT IS A TELEPHONE”) must be decoupled and quarantined from Sally’s representations of the world. If this representation were not decoupled then Sally would represent THIS BANANA IS A TELEPHONE as just another fact about the world, and hence would really believe that the banana is a telephone. If children did not quarantine pretense representations then every instance of pretense would be an instance of confusion and ultimately pretending would alter the meaning (for

example, the truth conditions) of representations in arbitrary ways. (Friedman & Leslie 2007: 108)<sup>47</sup>

That these ways exist without the need for an external object corresponding to it or it conveying any information about a concrete object can be more clearly seen when we look at mental representations in cognitive science. For instance, Chomsky in his interview with Ludlow (2011) talks about a cube as an internal representation being present in the mind that represents properties of cubes without there being a cube out there in the world, merely computed as a cube representation due to a tachitoscopic presentation by a perceptual psychologist and calls it an “internal event”. This also explains why the Muller-Lyer illusion looks the way it does. Even though they are not literally cuboidal (for instance, the mental representation of a round apple is not itself literally round), they can be said to have the properties *being an experience of being cuboidal* and *being an experience of being square*.

My claim is that properties like the way things appear to us, the way things look to us, the way we conceive and in fact properties derived from any sense modality and derived from amodal properties (properties not derived from the senses, for instance properties of innate ideas) are encoded by the mental representations that constitute fictional characters. We have already seen in the previous chapter from Matravers that propositional attitudes do not determine content in fiction (as well as the deluded Doyle argument which I gave before in Chapter 2 and which I shall rehearse below). I would add that it might influence it but would not determine it (I shall explain this again below). So, we can hallucinate Holmes, we can be deluded that we are Holmes, we can dream of Holmes, we can imagine Holmes, we can conceive of Holmes in particular ways and there might be mental imagery, there may not be mental imagery. We can derive properties of Holmes from visual properties, like in film, opera or theatre, in which there is a way the goings on the screen or stage appear to us, in

---

<sup>47</sup> I do hold that mental representations represent both mind-dependent properties and other properties as is evident in the passage from Friedman and Leslie. It’s just that the content of the mental representation when it comes to fictional characters is the mind-dependent property because the content of these representations, when it comes to fiction, is decoupled from representations of the world. So while these representations are representations of properties out there, the content expressed by predicates in fiction are the properties encoded by them (mind-dependent properties). This is not to say that the content (*being a thought of being red*) isn’t individuated because of a causal link to a property in the world. So being a thought of being red is individuated, that is, it’s a thought of the property red either due to a causal link to a red object immediately before me or in the past. The representation then carries information about properties in the world which it encodes as mind-dependent properties. It’s just that the representation is not directed at the world now.

which case, Holmes will be more vivid than what we think of him when we read a literary text. We hear random sounds (even in our heads) of fictional characters. They are visually richer than we can ever describe them, like in film, where new watching makes us discover more properties of the character, including visual properties (“Was Holmes wearing a watch? I missed it on the first viewing. Let me check.”). In the last chapter, we also saw the case of Toronto being used as a film set to depict New York. In that case, we might perhaps notice more properties of the fictional New York based on properties of Toronto than when we read about it, where the visual features of Toronto would be absent and we only think of New York. These are also the reasons why fictional characters and fictional objects turn out to be more than what we create of them.

Of course, to us it appears that it is Holmes the person who has these properties because of ontological metaphor, but in fact they are just properties encoded by mental representations. There can also be contradictory properties of fictional characters, like the round square. Since I am not talking of the properties of being round and being square being literally applied to an object, it is merely the mental representations encoding the properties *being a thought of being round* and *being a thought of being squared* that are bundled. We in turn think of that as a fictional character due to ontological metaphor.

It is pertinent to note that this does not lead to Meinongianism. The presupposition (and this is a conceptual presupposition like the metaphor MIND IS A BRITTLE OBJECT that we’ve seen before), for statements like “there is a dagger before me” in the case of a Macbeth’s hallucination can be expressed as (MET)<sub>HAL</sub>. In (MET)<sub>HAL</sub>, x is constituted by a bundle of mental representations that encode mind-dependent properties conveyed by the dagger-representation. Of course, he’s not aware of presupposition, like we are not ordinarily aware of conceptual metaphor.

(MET)<sub>HAL</sub> There is an x such that x is likened to a real dagger”

When he comes out of his hallucination (observe the container metaphor that was presupposed being expresses here, that he *comes out* of it, as if he could *go inside* it), what Macbeth means when he says that the “dagger does not exist” is that was not a concrete object in the external world but it was merely a thing in the mind. Of course, he calls it a thing, since there is a conceptual metaphor, whereas it is actually a bundle of mental representations and ontological metaphor doing their work. I also mentioned in Chapter 2 that when there is a quantification over such things like “there is a fictional character”, it is

underdetermined what the nature of the entity is, that it can be an abstract eternal object or whether it is constituted by a bundle of mental representations. In this case the entity quantified over is constituted by mental representations. (I rehearse this argument below. See also discussion on negative existentials in Chapter 4).

## **2. Arguments for Fictional Content**

In what follows, I shall lay out an argument for why bundles of properties encoded by mental representations are the best candidates for the contents of fictional characters. First, if fictional entities are eternal entities like for van Inwagen, the problem is that no substance will be found that “holds” properties together that explains why it persists through change. Second, if we look for fictional characters in the external world like in Kripke’s argument that fictional characters cannot have internal structure, then we realize that there is no substance to be found in the external world either that persists through change. In what follows, I attempt to explain why van Inwagen ends up with the problems he does and Kripke has the intuitions he has and in doing so, claim that a likely explanation is that the content of their mental representations is given by mind-dependent properties.

### **2.1 Explanation for van Inwagen’s Mistake**

In chapter 1, I argued against Van Inwagen’s theory in which he took fictional objects to be necessary existents that are not dependent on human minds, the causal nexus of the world and on language. Van Inwagen postulated them as eternal and necessary existents to explain why fictional characters are not part of the causal nexus of the physical world in response to Jung and Pellet’s (2018) worry that fictional characters are created and dependent on human minds and form no part of the causal nexus like theoretical entities of the sciences. What I argued against van Inwagen was that statements like “there are characters in some 19th-century novels who are presented with a greater wealth of physical detail than is any character in any 18th-century novel” are merely contingently true, because no author might have written such a novel and there might not have been such characters. In the case of “Mrs. Gamp is a masculine anti woman”, Dickens could have made her feminine or might not have even conceived of her. Van Inwagen’s reply, as found in the literature, was to say that they exist uninstantiated, since the author did not make them “accessible to readers” (van Inwagen 2018: 228). But the problem still remained, that properties like *being a masculine anti woman* were not necessary of Mrs. Gamp, but contingent, since she might have been feminine and that van Inwagen needs to explain why the property is contingent.

My explanation was a simple one, that the property was contingent, because it was contingently written that the character has a certain property. A possible explanation by van Inwagen might be that the properties are contingent because they were not intrinsic. But, as we saw in my argument above against Jeffrey Goodman (2004), the only properties that a fictional character literally had in van Inwagen's picture was *being fictional* (which was not an intrinsic property) and *being an abstract object*.

Thus, in this case, the only properties the fictional character literally had was *being fictional* and *being an abstract object*. Furthermore even the property *being fictional* is not intrinsic. Let me recapitulate an argument from Chapter 3, in which I showed that whether a work is fiction or non-fiction does not affect semantics and the nature of the fictional object. Furthermore, even pretense is not directly relevant to the intrinsic nature of what we refer to.

Zemach (2003) had an objection against Thomasson's distinction between fictional and real contexts. Her distinction was to accommodate why statements from literary criticism can be read literally and be about abstract objects, like "Hamlet is a round character" (because it is in a real context) and why statements are not literally true and can be about flesh and blood people (because they are in fictional contexts). Zemach's objection was that there is regular switching between that which is fictional and non-fictional in biographies which are embellished and books like the Bible, and yet we understand them merely due to a grasp of English, without it being determined whether the sentences are fictional or non-fictional. To make understanding them dependent on the distinction between the two contexts would make understanding dependent on the classification by librarians. I had also briefly entertained a possible reply by Thomasson, that someone can later on come and correct the reader who mistook a fictional person for a real person that what is really meant is that it is a fictional person. However, my objection to this was as follows. Suppose Conan Doyle writes the Holmes Stories as an account of his delusional state and then dies the moment he finishes. People who find his manuscript either think that it is a true account of events or it is a work of fiction. After reading it and understanding it, they classify it as fiction or non-fiction. In this case, understanding it did not depend on whether it was fiction or non-fiction. Suppose people start reading it as fiction and someone discovers that it was an account of delusion. At best people's interest gets piqued by it or psychiatrists take interest in it but nothing stops most people from reading it as fiction. The point is, that everyone understands the manuscript about the exploits of a flesh and blood person because they understand English, which serves as a condition for classification. Additionally, after Doyle dies and the work is construed as fiction (and it is never found out what whether Doyle had a delusion),

since pretense was not involved but belief was (in the delusion), the nature of the semantic object has nothing to do with whether he believed in it or there was pretense involved.

Thus, the only intrinsic property left to attribute the fictional character is *being an abstract object*. So what van Inwagen ends up with is a mere abstract entity which is akin to a bare particular, which are tough to accept in one's ontology. However, if he rejects a bare substratum as theoretically redundant and thinks of the fictional character constituted by a bundle of properties without a substance holding them together, then perhaps he can save his view.

But the question is not only for how eternal properties in a Platonic heaven might even occur in a delusion in the Deluded Doyle case, the question remains is what holds them together. Objects as bundles of properties are difficult to accept in the external world, without any substance to "hold them together". There are constant philosophical issues with relations like compresence. If they are not necessary existents, then van Inwagen has to concede to Jung and Pellet's objection and say that they are dependent on human minds, language and the causal nexus. A parsimonious reply to give then to explain the apparent quantification over bundles of properties is to say that they exist in individual human minds, unless one is in the mood to postulate group minds, which is a very contentious view in the philosophical literature on collective intentionality. This explanation becomes even more pressing, since if they were objects over and above human minds, and additional entities in the world, it would be tough to explain their causal role in the physical world. Furthermore, van Inwagen's intuition that the fictional entity is an abstract object which instantiates and has intrinsic properties can be explained and deflated by the explanation that the most likely reason why he sees it as a substance is because of the ontological metaphor he does not realize is there (assuming my master argument is correct). That is also why he has all the problems on them "having" the properties as literal, since in fact only concrete persons literally have properties like *being masculine* and *being feminine*. My alternative explanation about the reason why he has gone wrong is precisely that he is mistaking attributes encoded by our thoughts or other mental representations like *being a thought of being masculine* which indicate properties like *being masculine* that belong to concrete objects.

The deluded Doyle argument also makes a further point against Everett and Schroeder's account. In their account, a fictional character is an idea for telling a story as if there was a real person. However "as if" seems to imply either pretense or the use of suppositional

imagination. It seems odd to say that while Doyle is deluded there is either pretense or the suppositional imagination is at work. And since the nature of the semantic object is the same in both the deluded state and after it, it seems that there is a stronger unconscious process at work, which is the phenomenon of conceptual metaphor which as per Lakoff and Johnson pervades our thought. This is why the “likened to” relation has a distinct advantage over their account.

## 2.2 Explanation for Kripke not Finding an “Internal Structure”

In this section, I try to show that bundles of mind-dependent properties account for the content of mental representations with respect to fictional characters because fictional entities do not have an internal structure the way substances have in the external world and that there is nothing within two qualitatively identical bundles that can distinguish them.

Kripke (2013) asks a metaphysical question, that even if there are no unicorns, “could there have been unicorns?” The answer, according to Kripke, is no. If we look at two possible worlds, one with a fake unicorn (like a ringer’s tiger) and one with a supposedly real unicorn, there is no way to tell which of them is real. This is because the way we can differentiate a real tiger from a ringer’s tiger is by saying that a real tiger has the essence of a tiger (*tiger* being a natural kind) but a ringer’s tiger does not. But in the case of the two things that look like unicorns, one cannot say that one has the internal structure of a unicorn whereas the other does not, because, he says: “the story just doesn’t tell us what the internal structure of the unicorn is supposed to be, and therefore it hasn’t told us which hypothetical animal to look for in another possible world.” (*ibid.*, 2013: 47)

So qua external world, he finds that no substance can have such a structure. Moreover, if we look at our mental representations too, it would be a category mistake to suggest that our mental representations can literally possess the property of *being white* that supposedly belongs to the unicorn, since there is no such entity or substance there to which the predicate “is white” can be literally applied. Mental representations aren’t the sort of things that can be the bearers of such properties and nor is there a substance with an internal structure to which the properties belong. Kripke might say that these are pretend properties and not real ones, since real ones belong to objects in the world. If one is looking for properties of a unicorn, one will not find any, since there is no such substance in the external world which bears these properties. In that respect Kripke is right. So when it comes to the external

world, fictional objects are not the sort of things (substances) that can have an internal structure.

Now if one asks Kripke what he was doing when he was looking at the possible worlds with supposed qualities of fake unicorns and asked if he identified anything, he might say he was pretending to identify the imagined properties. But at best, *being white* is just a pretend property if one is thinking that it is the property *being white* that belongs to the imagined unicorn. So, if you asked Kripke, whether he was identifying such a property when he looked at the possible world in which found the qualities in the unicorn, he'd say no. He's pretending to identify them but not *really* identifying them, since there could not have been unicorns and they do not belong to unicorns. But what he missed out is that there is a way his imagination appears to him (he cannot deny that there is no such thing as the imagination). In fact, given what I have said before, it looks like he is looking at properties of the ways his imagination appears to him and identifying them. In fact, that's what people like Sartre and Colin McGinn do when they do a phenomenology of the imagination; they look for properties of imaginary experiences. An explanation for what he is looking for, when he is looking into the qualitative possible worlds, since there is no such substantial object in his imagination, is that he is looking at two qualitative aspects of his imagination both having the same property (different tokens): *being an imagination of being white*. The quality of the imagination, of course, is partly individuated by the external property *being white*, of which it is an indicator. On the face of it, such a property exists. In fact, the reason why he cannot find an internal structure is that they are two qualitatively identical bundles of properties encoded by bundles of mental representations and not a substance in the external world. There is no such substance with an internal nature in the external world by which to distinguish between them phenomenologically, except for the fact that they are two tokens of his mental states, that is two unified bundles of mental representations.

### **2.3 Persistence Conditions for Fictional Characters**

In what follows, I shall lay out persistence conditions for fictional characters, consistent with Everett and Schroeder's account, in contrast with the alternative that perhaps fictional characters can exist as part of our social practices and be dependent on literary works and not in individual minds and are not spatiotemporal. In doing so, I shall provide a further argument that since fictional characters are constituted by mental representations, they can in fact "flit in and out of existence". We shall turn to this option now.



### 2.3.1 Thomasson's Continued Dependence Condition

In Chapter 3, I brought up Reicher's (1999) argument against Thomasson's (1999) condition for the continued existence of fictional characters. To recall, Thomasson called the continued dependence of a fictional character on a literary work generic constant dependence. This condition made the continued existence of a fictional character dependent either on:

- (a) A copy *and* the existence of at least one person who can understand the copy (that is a reader who can understand it's language it and has contextual knowledge to interpret it).
- (b) The memory of a literary work.

The fictional character would be destroyed if every copy was destroyed and the last person could understand it is dead or every memory of it is gone. Reicher's counterexample, which I think is a definite counterexample was that, suppose a copy of a literary work remains but all competent readership with background knowledge to understand it is dead. According to condition (1) above, the literary work and therefore the literary character is destroyed. Now suppose scientists revive the ancient language and knowledge of the culture by reconstructing it. Then on Thomasson's very condition, it is not that the work and fictional character is reconstructed but comes alive again. However Thomasson holds that works and characters cannot "flit in and out of existence" (Thomasson 1999:22) but in the counterexample, they do.

One way in which Thomasson can salvage any semblance of artefactualism is to say that the fictional characters do in fact flit in and out of existence since they are in the mind (as ideas) and they cease to exist when no one thinks of them. Furthermore, fictional characters cannot be concrete objects in the external world (if they do indeed "flit in and out of existence") since, as Kripke (1980) argues that there might have been many people who did the same actions as Holmes, but then one cannot say that Holmes is one of them. Even Darwin could have performed all the actions of Holmes. Thus, it seems apt to give an explanation of Thomasson's mistake, like van Inwagen, that it is likely that she simply missed out on the ontological metaphor and thought of it as a persistent substance as the bearer of properties because of which she thought they do not flit in and out of existence but remain constant through change.

Moreover, to revise this in my view, the persistence conditions will have to change to:

A fictional character continues to exist iff:

- (a) At least one person existing can think about the character.
- (b) The conventions of a particular medium, visual, aural, written or digital are present or only the cognitive ability to think about the character without the existence of conventions of a particular medium is present or both the cognitive ability to think about the character and the conventions of a particular medium.

The first condition accounts for why the fictional character goes out of existence, since there is no person existing in all that time who can think about the character. The second condition accounts for why the scientist can think about the character since the conventions of media are present. However, the reason why I have added the part about only the cognitive ability without the conventions of a particular medium is because it will account for cases where I have just thought of the character before putting it into words or even thinking about it in words, as if there was merely a mental picture of it. Also, I do not want to rule out a priori that one cannot in cases of aphasia conjure up fictional characters through their imagination.

We also saw in Chapter 3 that one of the reasons that Thomasson brought into saying that fictional characters are non-spatiotemporal is that we cannot bump into them. The problem that I saw with that was it had the assumption: if something is not a physical object, then it is not spatiotemporal. This I said was based on a false dichotomy, that either something is a physical object or it is non-spatiotemporal. My argument was an argument from analogy with Cartesian dualism, considering that this might be one major reason given for it:

P1) The mind (or a mental event) is not a physical object.

P2) Anything that is not a physical object is nonspatiotemporal.

C) Therefore, the mind is nonspatiotemporal.

One possible reason given for this might be of the same sort given in Thomasson's argument regarding fictional characters, that since one cannot bump into minds, minds are therefore non-spatiotemporal. However, this is fallacious, since the reason why we don't bump into our minds is because it is a person with a mind that comes into contact with physical objects (including their own bodies) and that the mind is presupposed when one is on the street and bumps into objects. The plausible conclusion from this is merely that the mind is spatiotemporal in a particular way. Analogously, the safe conclusion to draw for fictional objects is that if they are not non-spatiotemporal for the reason that we do not come into

contact with them in the external world, then they can be spatiotemporal in a particular way (inference to the best explanation).

Of course, it does not follow from this that fictional characters exist only in human minds, merely that they can be spatiotemporal. But consider the condition for continuance of a fictional character that I provided above. The conventions of a particular medium were not a necessary condition for the persistence of a fictional character. Thus, a character exists only if there is one person to think about it and the person has the ability to think about it (which is also implied in the first condition). So from this, at least, the conclusion can be drawn that even without medium specific conventions regarding the character, it most likely exists in human minds, since it is dependent on thought.

The further question is of the existence of conventions in different mediums on the basis of which one can reply to me that fictional objects are spatiotemporal because the copies of the work existed and the characters were just dependent on the meaning of words contained in them (it might be a weak objection, but let's consider it). One might also object to me by saying that since the conventions that were retrieved by the scientists existed in the archaeological evidence, the fictional characters existed, since they were dependent on the conventions. Furthermore, the reason why the character seemed to cease to exist until the scientists reconstructed the language was that during that time, since there was no one to interpret and reconstruct the language, the words were mere squiggles on paper or whichever medium it was in.

However, my reply is that in that time, the language died too, since squiggles on paper by themselves are mere squiggles and have derived intentionality from the ability of human beings to interpret them. The squiggles themselves are not about something, even more so since the same squiggle can be accidentally formed on a beach and a person can read it as a word. In that respect, the language died too, since all that was left in the world were squiggles. There is no intrinsic property in them that they "contain" (another container metaphor) the meaning of words.

Human beings, who rediscovered, reconstructed and revived the language could only do so because they gained relevant knowledge. They could do this only because they knew how orthographic and language systems work and because those patterns resemble how the language of the original authors worked, not because of mere intrinsic properties of the squiggles. So it can be said that it was interpreted based on relevant knowledge of their own conventions, them matching the squiggles and found artefacts with patterns from the

previous culture and inferring based on that knowledge that this is the most likely interpretation.

It must be observed that each person interpreting it is doing so only because each person has some knowledge of conventions. Each one who interprets it based on that knowledge, individually thinks of the character in his or her mind, even if everyone interprets it together. Suppose there is a case in which each scientist interprets one predicate like “lives on 221B Baker Street” without knowing it’s about the character and another one does part of this and this goes on till the last scientist gets a fuller picture and says something like “Ha! Sherlock Holmes is a fictional man who lives of 221B Baker Street”. And then everyone thinks about Holmes after that. What are the relevant components of this situation? We can find knowledge of conventions and inferences that were made on patterns adding to the knowledge. At the end of which, unless one wants to believe in group minds there is each person with a thought about the character in his head (we’ve already considered whether they’re outside the causal nexus and rejected it in Jung and Pellet’s objection, I won’t consider that they’re thinking of something outside the causal nexus). It would also be *ad hoc* to postulate group minds here just for the sake of this. Thus, we can also say metaphorically that they thought about it together, but they actually thought about it in the head.

So what do we have at the end of this? There is knowledge, there are inferences each scientist makes, but at the end of the day, they think about what’s in their head, the character. So where is the character? Most likely in their heads (also given arguments against van Inwagen regarding causal issues). This is merely an example case but this can be extended to every case with social practices. Unless one wants to denounce methodological individualism and accept group minds, all one can say is that each person possessed knowledge, based on which they individually had a thought, even if we say they thought about it together. So at most, when one says the character is embedded or exists as part of our social practices, if one spells it out, it is a metaphor. There is only knowledge of conventions, inferred patterns, squiggles and a thought about a character. The character is in fact in the head, and given what I have said about the mental representation and content before, it is at best an idea constituted by a bundle of mental representations in the head whose content are the properties each representations encodes. Individuals thus think about them and identify the characters based on knowledge of conventions.

A story can be told on this too. As each scientist interpreted a predicate, they thought about a property and their mental representation encoded a property like *being a thought of being*

*a detective*. Of course, the character is in fact part of our social practices, but the character is in the mind. It is odd to say the character is all the conventions put together especially since the situation can be broken into its parts and it is seen that each individual thinks of the character by means of a knowledge of conventions. And it is a category mistake to think that characters are part of social practices, that form the criteria for identifying the character.

My account further explains why statements like “Goethe’s Werther caused many German youth to commit suicide” are true. If the properties are mind-dependent, then since the bundle of properties encoded by a mental representation was likened to a person and counted to be such for fictional purposes, German youth were caused by it to commit suicide. Thomasson’s non spatiotemporal abstract artefacts will not account for this, this it will have trouble allowing for causal relations. Furthermore, we can also account for why we do not bump into Holmes. That is because Holmes is constituted by a mental representations that encode a bundle of mind-dependent properties counted as a person and not a real person. And yet, one can interact with mind-dependent properties encoded by mental representations that are abstracted from the real world<sup>48</sup>.

### 3. Interest Relative Identity

The main problem that we saw in Chapter 1 was that there are no necessary and sufficient

---

<sup>48</sup> Perhaps the artefactualist can escape my objection by claiming that fictional characters are types or even created types. Enrico Terrone (2017) has a plausible view of fictional characters as created types, where he borrows the definition of types from Strawson as a “historically established principles of construction of like particulars, which are called the token’s types” (ibid., 2017: 163). However, for such a view to be more plausible, it has to explain what is happening in the scientific reconstruction case, where I can say that such principles are part of the information states of interpreters and they end up thinking of the character. I leave it open to compare advantages of mine and Terrone’s view for further literature on fictional characters.

Nevertheless, with respect to other views which see characters as types, it is ontologically parsimonious to say they are in the head, since we do not need to postulate an extra entity over and above the physical world. It is also ideologically parsimonious since the burden is on type views like Wolterstorff’s to answer objections to counterexamples against them possessing core properties, which we saw in Chapter 1. Furthermore, worrying about type views, Everett writes:

I am perfectly willing to accept the existence of types-of-people. And I grant that they may sometimes play an important role in the way we think and talk about fiction. But they are not fictional characters; *they are not* what we talk and think about when we think about the character of Holmes or talk about Austen creating Emma. (Everett 2013: 143 as quoted by Terrone 2017: 176)

conditions for the individuation of fictional characters and that every property of a fictional character is contingent. We also saw above from Kripke's argument that no substance can be found in the external world to hold the properties together through change and from the argument against van Inwagen that it cannot be found in an eternal realm either. One of the reasons a substance is postulated is to account for how two objects can have the same properties and yet fail to be identical despite the properties being the same, Miller 2019). It is because there are two substances (with an internal structure), each possessing the same properties.

However, I argued that fictional characters are constituted by token mental representations which encode properties. If we take the science on what brings properties encoded by mental states together, then they are most likely linked to each other by mental representations whose properties they are, which are in turn linked by neural binding.

In what follows, I shall borrow part of the discussion as aligning with James Miller's (2019) discussion where he claims that words are bundles of properties/features, which has a lot of similarities with my account. This, apart from the examples from Everett and Schroeder above, will explain how interest relative identity works for particular tokens of mental representations, especially when all the properties of a fictional character change between time  $t_1$  and time  $t_2$ . It must be said at the outset that this is not to say that the individuation of a particular mental representation as that mental representation is interest relative. This is precisely why I say that a fictional character is not identical to a particular mental representation but is constituted by it. The question is whether the same mental representation counts as one fictional character or another. This shall be clarified below.

Miller takes up his account from generative linguistics, where words are typically seen as bundles of features which "are the building blocks of language, and are taken to be the elements stored within our mental lexicon." (*ibid.*, 13) For instance, the word "airplane" will have the features "phonetic [begins with vowel]; semantic [artefact]; and formal [+nominal]" (*ibid.*, 13). The reason why I pick this account is to see a similar case of interest relative identity as my account, where features in the mind (since Miller understands the features to be "stored in the mental lexicon") and not just in the external world, are individuated in an interest relative manner. This is also consistent with standard Chomskyan mentalism, where mentalism is Chomsky's view that the object of the scientific study of language is properties of the mind/brain of a language user and language is in fact in the brain. This is in contrast with other notions of the study of language as an abstraction or as

texts or in a community. The Chomskyan view entails that a community uses “a language by virtue of each speaker having “essentially the same linguistic system in their mind/brains.” (Jackendoff 2004: 651) The mentalistic enterprise then places constraints on a theory of grammar, that is, that theoretical elegance is not mathematic elegance but “brain elegance”, any theory of grammar has to explain how language is acquired and has to account for how language evolved after we separated from chimpanzees. Moreover, the philosophical tradition, from Frege onwards takes “language to be independent of its human users: it relates directly to the world.” (*ibid.*, 222) This approach however, despite attempts of philosophers like Katz to connect it with mentalistic generative grammar “disconnects generative linguistics from all sources of evidence based on processing, acquisition, genetics and brain damage.” (*ibid.*, 223) Jackendoff points out that it is absurd to think of Nicaraguan sign language as waiting for it to be “grasped” till the 1980s or English lingering since the Big Bang till someone grasps it. At best, it is an idealization and one can see this as our commonsense notion of language (this is explained further in the section on Chomskyan theories of reference).

It must be said however, that the theory or practice of generative linguistics is independent of Chomskyan mentalism (I will however assume mentalism). As Professor Nirmalangshu Mukherji, in personal correspondence, says:

The theory and practice of generative linguistics is independent of views on mentalism etc....Linguists don't care about the mind or the brain. They just study human languages. Just as physicists don't care whether they are physicalist or instrumentalist. Feature theory of lexical items is a standard way of composing computational atoms. Whether they belong to the mind or the brain or mind-brain is irrelevant for syntactic theory. Since syntactic objects such as pronouns don't float in the air, it's natural to think of them as internal to the speaker...But whether they are internal or external has no bearing on the theory.

### **3.1 Objections Against Bundles**

Miller (2019) considers a possible objection by van Cleve (1985) against bundle theories, which also partly applies here. The objection is that if a particular bundle changes, then it becomes a new bundle and thus a new individual. That is, if bundle theory is true, all properties that an object has are essential to it. But this would entail that if even a single property that makes up the token of a word changes then it becomes a different token. Of course, the account of fictional characters that I take up does not imply that a fictional character is identical to a bundle of mental representations which encode properties at time *t* but that it is constituted by such mental representations. The question then is what makes a

particular bundle of mental representations which encodes radically different properties the same fictional character (or idea) as another bundle? I appeal to interest relative identity analogously to how Miller does it for words.

In the case of words, Miller gives a counterexample that if he types the word “table” and merely makes the font bold to “**table**”, it still remains the same token despite a change in the bundle of properties. That is, the word token has not ceased to exist and a new one has not been created and yet a property not essential to it has changed. But if van Cleve’s objection was correct, then the changing of a non-essential property would have created a new object. Since this looks counterintuitive, Miller gives the example of false friends, where the same or similar form of a word in the same or different language is actually a different one. For instance:

(6) Gettare i **confetti**.

(7) Throw the **confetti**.

Now ‘confetti’ in both languages (Italian and English) are different words, although with similar meanings. But if one rubs “Gettare i” on the blackboard and replaces it with “Throw the” in (6) then they become two tokens of the same word. The explanation that he gives for a difference in intuitions between the table case and the confetti case is similar to a case of interest relative identity that I have mentioned before, the explanation being that the reason why we think that the words change in “confetti” is because semantic properties of words are more relevant to our purposes in language rather than orthographic properties. Thus, he says:

Tolerated differences for relevant communicative aims or purposes are, though, pragmatic concerns about successful and productive communication. The bundle theory can accept the influence of these pragmatic concerns whilst holding that, metaphysically, changing the font of a word creates a different particular word, just as a change in the semantic properties of a particular word will also create a new particular word. The counter-intuitiveness of the bundle theory in the first case is undermined by the intuitiveness of accepting the objects as being different particulars in the second. (Miller 2019: 10)

That “metaphysically, changing the font of a word creates a different particular word” is what I mentioned in the first chapter, that with respect to what are the constituents in the world that make up fictional characters, a fictional character is constituted by a bundle of mental representations at time *t* that encode a bundle of properties. This explains why the bundle can be different, metaphysically even when a property is changed. However, when



we ask the question, what makes a fictional character that particular character and not another one, the answer in this case is that what matters here is interest relative identity. It is not the intrinsic properties encoded by the bundle that are only relevant here (despite the changed bundle), even though it constrains what constitutes a fictional character with respect to our interests. That a particular bundle constitutes a fictional character y or not in that respect is dependent on our purposes, the way constitutive rules like (x is counted as y in context c) (Searle 2006) determines what makes a piece of paper a particular five dollar bill or not<sup>49</sup>.

In Chapter 1, we saw how even if all the properties of the fictional character changed, it would be the same one from Thomasson's case of Faust and Phaust. That is, if in another novel, Faust is named Phaust and where instead of making a pact with the devil, he was a quiet scholar. This becomes even more clear if we imagine a more radical case where the author assumes people's knowledge of context, so he creates a novel to imagine what would have happened to the first Faust if he led a different life and names the character Phaust. So instead of a scholar, he is a carpenter. In this case there are two radically different bundles of properties encoded by two token mental representations, but the same character. That explains why despite a change in properties, even though the new bundle is completely different, that does not entail that there is a new character.

Consider a similar case as the one on "confetti" above, which I mention in Chapter 5. Imagine a person who hears about Holmes, remembers one description ("wears a hat") and ceases to remember all others. Perhaps she has some form of dementia. Then she writes a

---

<sup>49</sup> It is knowledge of conventions that enable us to count two bundles of representations as the same character or the same bundle as a different character. Furthermore, the conventions exist in the brain. X knows that p is a mental state and the information is encoded in a brain state. With regard to conventions, Jackendoff says:

What are they, and where could they be, other than in people's heads? You might say they're in the practices of the community. But members of the community can only conform to these practices because of something in their heads. (Jackendoff 2012: 11)

My view is purely within naturalism. It does not even make sense to say that conventions are anywhere else but in at least somebody's head. Definitely not in group minds. The burden is on the person who opposes me (someone like David Lewis) to answer the following question: Apart from all the physical entities in the world, are conventions something over and above them? This is the same as asking the question social ontology starts with: Over and above all the physical things in the world, is there something over and above this such as money? Or the problem of collective intentionality: Over and above all the physical things and processes in the world, which include individual minds, are there group minds? In most of these areas, the answer is mostly a resounding no. So I really do not understand what someone means when someone says conventions are not in the head. Maybe not in my head at time t, but in someone's head. That's the null hypothesis, that they are in minds. I don't have to show they aren't. The opponent does.

novel with a character named “Holmes” who wears a hat, but with all different properties. This author goes to a Holmes nerd who is talking about Holmes, and tells him how she heard about Holmes and decided to write a novel on him, describing some features of Holmes in her novel. The nerd is bound to ask her, “Are we talking about the same character?”, the answer is that it is indeterminate whether it is at best and it is not the same at worst. Now, suppose she hasn’t written a novel, goes to the nerd who’s talking about Holmes and tells him enthusiastically about Holmes talking about the same mistaken properties as in the first case. This time the nerd definitely corrects her and says that she’s talking about Holmes but she’s got the properties wrong.

The only salient difference in the two cases is the nerd’s knowledge of the novel. It is this contextual knowledge in this case that determines whether the character the woman is thinking about is Holmes or not and whether they are thinking of the same character or not. The woman on the other hand, in her idiolect refers to the same bundle of mental representations encoding the same properties.

The point on interest relative identity also explains why I do not consider ideas as identical with mental representations but constituted by them. Let us go back to Chomsky’s tea example again which I quoted in Chapter 1.

Substances reveal the same kinds of special mental design. Take the term "water," in the sense proposed by Hilary Putnam: as coextensive with “H<sub>2</sub>O give or take certain impurities" (Putnam 1992, citing his now classic paper, Putnam 1975). Even in such a usage, with its questionable invocation of natural science, we find that whether something is water depends on special human interests and concerns, again in ways understood without relevant experience; the term "impurities" covers some difficult terrain. Suppose cup1 is filled from the tap. It is a cup of water, but if a tea bag is dipped into it, that is no longer the case. It is now a cup of tea, something different. Suppose cup2 is filled from a tap connected to a reservoir in which tea has been dumped (say, as a new kind of purifier). What is in cup2 is water, not tea, even if a chemist could not distinguish it from the present contents of cup. The cups contain the same thing from one point of view, different things from another; but in either case cup2 contains only water and cup1, only tea. In cup2, the tea is an "impurity" in Putnam's sense, in cup, it is not, and we do not have water at all (except in the sense that milk is mostly water, or a person for that matter). If cup3 contains pure H<sub>2</sub>O into which a tea bag has been dipped, it is tea, not water, though it could have a higher concentration of H<sub>2</sub>O molecules than what comes from the tap or is drawn from a river. (ibid., 127-128)

Both cups are constituted by water. We don’t count both as tea, however. Similarly, the same token mental representation (in the same way as the cups in two contexts which are constituted by the same ingredients) can be counted as Holmes in one context and another character in another (like in the nerd case above). There’s a problem with transitivity if we

say the character is identical. I can just say ideas are constituted by mental states, but like tea, they are counted as one fictional character as opposed to another. The same goes with money. A five dollar bill is constituted by paper, but is not identical to that paper. The paper gains that significance after we count it as such due to constitutive rules. Similarly, a mental representation gains the significance of being one character rather than another after we count it as such. *In effect, the idea is the result of all the ingredients involved. The “stuff” in the world that constitutes it. The conceptual metaphor. The constitutive rules.* Same with the tea and money. Our interests are involved in imbuing constituents in the world the significance they do (mental states being objective parts of the world too). It goes this way too: ideas are significant to us in this context. Mental states construed as mere brain processes in themselves aren't. So it's better to think of ideas as constituted by token mental states.

#### **4. Referring to Mental Representations**

Till now, I have attempted to show that fictional entities are constituted by mental representations. Furthermore, the content of mental representations pertinent to fictional characters are the properties encoded by mental representations. In what follows, I shall try to show how we can refer to a mental representation using research on reference processing as well as Lauri Karttunen's (1976) work on discourse referents.

In the literature on reference processing, Garrod (2011) says that

Situation models are assumed to be multi-dimensional representations containing information about space, time, causality, intentionality, and currently relevant individuals....They capture what people are “thinking about” while they understand a text, and therefore are in some sense within working memory (they can be contrasted with linguistic representations and with general knowledge).....The basic idea is that models of the “discourse world” mediate between references in the language and the entities they refer to.” (Garrod 2011: 274)

Thus, in the absence of real world referents, situation models serve as “surrogates” for the “world of discourse.” (*ibid.*, 274) First, such models are constructed as representing the world in the discourse and second, referents are mapped on to that world. As I said before, the situation model is possible to have explanation within my view, that the representation of the world happens because of conceptual metaphor wherein the representation is likened to the concrete world. This is why the representation appears as the world.

To understand how situation models and reference in them work, consider the following (*ibid.*, 274 adapted from Stenning 1975):

(8) In the morning Harry let out his dog Fido.

(9) In the evening he returned to find a starving beast.

Garrod points out that (8) introduces two referents into the situation model and in the situation model, readers then map a starving beast onto the referent of “Fido”, that is, they are interpreted to be “coreferential”. Garrod then invokes Stenning’s (1975) suggestion that the reason that there are two and not three referents (that is, there is no referent of “a starving beast” additionally) introduced into the discourse model is because of parsimony. This inference is made in the absence of semantic or syntactic cues, where there is no semantic association between *His dog Fido* and a *starving beast* and that there is no anaphoric cue, and a starving beast is presented as an indefinite rather than a definite Noun Phrase.

On the other hand, in the context of Figure 1. below, one is more likely to interpret a starving beast as referring to the tiger, rather than the dog Fido, since the tiger is salient. Now the reason why Garrod argues for the need for a situation model in a real world visual context like the figure is precisely because one might argue that a situation model in this case is redundant since “with a picture the viewer simply links the linguistic reference directly to what is accessible” (*ibid.*, 276) from the picture. However, if one introduces a sentence like (10) with anaphora that follows (9), then it seems more likely that “it” refers to the dog, and the situation model is required since one needs to keep in mind how the previous referents have been mapped.

(10) It had nothing to eat all day.



Figure 1. Reference Picture for sentences (8), (9), (10). (Reprinted with permission from Garrod, S., “Referential Processing in Monologue and Dialogue with and without Access to Real-World Referents” in Gibson, E., & Pearlmutter, N. J. (Eds.). (2011). *The processing and acquisition of reference*. MIT Press. )

Garrod points out that the difference between the visual and the reading task is that while in the former, one already has a readymade situation based on the picture, in the reading task, one uses information provided in the text and knowledge of what the text is talking about to construct a situation model in the case of reading.

Garrod’s examples show an interesting instance of the confrontation/representation distinction that Matravets (2014) brought out. For Matravets, if we recall from Chapter 5, content in fiction belonged to a broader category of representations\* (for convenience, I shall call them representation\* to prevent confusion with mental representations) as opposed to confrontations, instead of determined by propositional attitudes. Confrontations, we saw, were situations in which there is the possibility of action in our egocentric sphere and representations\* were situations in which such actions were not possible, since that which was represented was could not be acted upon in our egocentric sphere. This distinction also:

...aligns with situations in which our mental states are online and situations in which our mental states are offline. That is, in confrontation relations our mental states are caused by perceptual inputs from the objects of those states, and which cause actions towards objects in our egocentric space. In representation\* relations our mental states are not caused by perceptions of the objects of those states, and do not result in actions towards objects in our egocentric space (although, of course, they can still cause actions). (Matravets 2014: 50)

The referential processing is the same for fiction and nonfiction (especially since in the examples, Fido doesn't necessarily have to be a real dog and in the example, could just as well be fictional) in the absence of a visual cue (Matravers brings in a host of empirical evidence to show that the processing is the same too). This explains Zemach's objection to Thomasson above that even though books like the Bible and embellished biographies regularly switch between what is fictional and non-fictional, we still understand them because we know English, not because the sentences were determined to be fictional or non-fictional. However, the moment the picture of the tiger comes up, due to a confrontation relation, the picture receives the dominant interpretation, and becomes the salient candidate for the anaphoric referent. As Garrod says "With a visual context what you see dominates the referential interpretation — what Stephen Dedalus called "the ineluctable modality of the visible" in *Ulysses* (Joyce 1932). This is consistent with Matraver's view that I have adopted, since what explains the "ineluctable modality" is precisely the possibility of action with respect to the confrontation relation with the perceptual input.

Furthermore, in the light of what Friedman and Leslie<sup>50</sup> say for fiction, we need to make sense of what it is to say that the content is decoupled from one's representation of the world. My interpretation is that the mental representations involved in fiction shift the content from properties in the world that we encounter in confrontations to mind-dependent properties that are present in representation\*. Furthermore, along with mental representations pertaining to fictional entities, even discourse referents in representation\* that are unified mental representations of external concrete objects have content which is *sui generis*, which is what allows them to be discourse referents within situation models (of which more below). As we saw in Chapter 5, Matravers (2017) makes a distinction between R-properties which are properties that are possessed by the representation and W-Properties which are "properties that are represented as belonging to the world". So while the W-Properties allow the representations to be about objects and events in the world which are part of its content with respect to W-Properties, it is the R-Properties which also form part of their content which allow it to be an independent discourse referent and explain why the

---

<sup>50</sup> To quote them again:

...the content of Sally's pretense representation (e.g., THIS BANANA "IT IS A TELEPHONE") must be decoupled and quarantined from Sally's representations of the world. If this representation were not decoupled then Sally would represent THIS BANANA IS A TELEPHONE as just another fact about the world, and hence would really believe that the banana is a telephone. (Friedman & Leslie 2007: 108)

processing is neutral between fiction and non-fiction. Perhaps (and it looks like that from my account) misrepresentation in perceptual illusions and false memories can occur because of R-Properties and the basis for correcting a person and saying that they are wrong is because of W-Properties.

Second, that it seems plausible to say that attributes of discourse referents are updated through further information in the text, especially when we encounter predicates and as our constructions are updated with textual information, so does the nature of the referent as given in the construction. This bolsters my view that fictional characters are constituted by bundles of mental representations to which reference is made which keep changing as well as the properties encoded in them, as readers update information. Think of a reader or author encountering or thinking of a fictional name for the first time to which they add an indexical feature (Jackendoff 2002), that is a “minimal identifiable perceptual feature” of the object which we represent to ourselves. Jackendoff describes indexical features as follows.

Suppose someone tells you to look at a patterned rug, saying “Hey, look at that!” Initially, you see nothing, but as you observe, a bug “pops” out into your experience. Jackendoff says that even though there is no change in the external world or on one’s retina (since the bug was already present), there is a change in the way one’s percept (which is limited to one’s brain) is organized (percepts are constructed by the brain on information received from the retina, and are non-linguistic and tracked by us, in which individuals have been individuated). The bug emerges as figural and Jackendoff calls this figural feature an indexical feature of a percept, that is, that which makes the bug representation in the percept a something to which other features are “bound”. This is the “minimal identifiable perceptual feature” he talks about to which one identifies in “I don ’t know what that was, but here it comes again!”

Indexical features are *sui generis* and exist in our minds (but they “indicate” and are causally related to properties of the brain and the shape features in the external world) even when there is no perceptual input and are what account for our sense that objects exist even when we are not perceiving them. Crucially, these are what are very important to reference, and one cannot refer using a deictic or linguistic expression without features being bound to indexical features. Consistent with Leslie’s interpretation of decoupled representations we saw before, such features look like they can be decoupled from representations of the world to form part of representations\*, in Matraver’s sense<sup>51</sup>.

---

<sup>51</sup> Think of these discourse referents as doing the work of Frege’s senses, except that they are psychological associations and the wrong properties can be associated with names too

Nevertheless, this indexical feature gets subsequently enriched with more properties. At time  $t$ ,  $x$  refers to a bundle of mental representations that is the fictional referent at time  $t$ , and at time  $t_1$   $x$  refers to the referent with different properties even as the mental construction of the person who thinks about it enriches properties. Furthermore, what accounts for these representational properties to be about objects in both instances is precisely that there is a conceptual metaphor operating and that there is a causal link of the mentally represented properties with the external world.

#### 4.1.1 Discourse Referents and Denotation

Laurie Karttunen (1976) argues that discourse referents are introduced into the discourse when a sentence entails the existence of the referent, so that the same referent is spoken of again. However, Karttunen thinks that discourse referents are a linguistic phenomenon, and not ontological. The reason it is not ontological but linguistic is because in Karttunen's view, discourse referents are introduced only if the existence of the discourse referent is entailed (it doesn't have to be the existence of a concrete object and so it is ontologically neutral; the entailment seems purely linguistic) I argue that that's not the case. So if *prima facie*, existence is not entailed *prima facie*, you still get a discourse referent. I argue, in the light of some counterexamples, that even if they do not, *prima facie*, entail the existence of the referent, it does not imply that a further discourse referent is not introduced, merely that it is unresolved whether it is the same as the previous one. There is more that is happening

---

as we shall see in the objections section. The discourse referent of "Hesperus" is a unified mental representation that, if a person is correct about it, encodes the property *being a thought of being the morning star* (the representation is likened to a concrete object). The discourse referent of "Phosphorus" includes the property *being a thought of being the evening star*. To account for the triviality of "Hesperus is Hesperus" and new knowledge in the case of "Hesperus is Phosphorus", Jackendoff (2002) says this:

We can understand Frege's example in present terms as reporting the merger of indexicals associated with different descriptive features, exactly as in our example of the circle that is sometimes red and sometimes blue. So Frege's problem is not a uniquely linguistic problem; rather, it lies in a more general theory of how the  $f$ -mind keeps track of individuated entities. (Jackendoff 2002: 314-315)

The merging of indexical features of these two discourse referents accounts for Russell's intuition that Mark Crimmins (1998) points out and tries to solve:

Identity is a rather puzzling thing at first sight. When you say 'Scott is the author of Waverly', you are half-tempted to think there are two people, one of whom is Scott and the other the author of Waverly, and they happen to be the same. That is obviously absurd, but that is the sort of way one is always tempted to deal with identity. (Russell 1985: 115 as quoted by Crimmins 1998: 1-2)



than mere linguistic entailment. In the light of this, I argue that discourse referents are in fact ontological and not merely linguistic.

Cumming (2014) further argues that these discourse referents are not referents but there is only a relation of denotation as opposed to reference. In what follows, I shall specifically argue against Cumming (2014), why discourse reference is not merely denotation, but in fact reference. This would bolster my case that at time *t*, there is always a reference to a mental representation (Everett and Schroeder also talk about reference to mental representations in this context) and that it happens while we are interpreting fiction or while an author is writing a novel. As I have mentioned before, I try to account for why an author starts out with thinking of a fictional character and refers to it throughout despite changes in properties, and even allowing for the possibility of changing all properties.

#### 4.1.2 Karttunen's Discourse Referents

In the case of fictional entities, Lauri Karttunen (1976) brings up the following parallel between the introduction of a discourse referent in an ordinary and in a fictional case (*ibid.*, 366):

(11a) Bill has a car.

(11a) can be followed by either of these.

(b) It is black.

(c) The car is black.

(d) Bill's car is black.

On the contrary, Karttunen says (12a) cannot be followed by (12b), (12c) and (12d) (*ibid.*, 366):

(12a) Bill doesn't have a car.

(12b) \*It is black.

(12c) \*The car is black.

(12d) \*Bill's car is black.

For Karttunen, that the sentences given above that follow (11a) and do not follow (12a) shows that in order to interpret (11), the indefinite NP has to be seen as introducing a discourse referent and entailing its existence in order for it to be spoken about again by means of the pronoun or a definite description. However, in the second case, since (12a) does not entail the existence of a car, (12b), (12c) and (12d) do not follow and no discourse referent is introduced.

Now, Karttunen insists that the discourse referent is linguistic and not ontological due to the following parallel case where (13) parallels (11) and (14) parallels (12) (*ibid.*, 366):

(13) Bill saw a unicorn. The unicorn had a gold mane.

(14) Bill didn't see a unicorn. \*The unicorn had a gold mane.

Thus, she says:

Let us say that the appearance of an indefinite noun phrase establishes a discourse referent just in case it justifies the occurrence of a coreferential pronoun or a definite noun phrase later in the text. (Karttunen 1976: 366)

However, the definition does seem correct, but it does not follow from this that there is no discourse referent in (14). For instance, (14) can also be interpreted that two discourse referents were introduced into the discourse, one for “a unicorn” and the second for “The unicorn” in the second sentence. However, it just wasn’t inferred that they are about the same referent because the first one is where the unicorn is supposed to be a concrete object, but there wasn’t any seen and the second one refers to a mental entity with no clear indication whether the unicorn is supposed to be in the external world or is internal to the mind (this shall be made clear below).

Consider the following:

(15) Bill didn’t see a unicorn. It was in his imagination.

(16) Bill didn’t see a unicorn. He thought he’d seen it.

(17) Bill didn’t see a unicorn. The unicorn had a gold mane. It is there alright, but in his head.

(18) Bill didn’t see a unicorn. Such things cannot be seen but thought of.

Thus, the above will come out to be correct, where they are inferred to be the same referent and involve coreferential pronouns even when on Karttunen's condition, the first sentence does not seem to entail the existence of the discourse referent. Moreover, (17) is even more troublesome for Karttunen, since the appearance of the NP "a unicorn", does not immediately allow for coreference by "the unicorn" in the second sentence, but we find coreference when the third sentence appears. So (17) shows that there were in fact two discourse referents introduced, but it was resolved that they are the same in the light of further evidence. In the light of Garrod's argument above for situation models, the resolution can be explained precisely because by the time the reader comes to the third sentence, the reader had to keep in mind how the referents in the previous two sentences were mapped in a situation model, in order to make the resolution. As we shall see, there's more than something merely linguistic happening for this resolution to take place.

#### **4.1.3 Negative Existentials and Valuation Features**

In what follows, I shall use mine and Solodkoff's analysis of negative existentials as a prelude to further clarifying the issue. I will also explain valuation features like +EXTERNAL and +INTERNAL to explain the difference between two bundles of mentally dependent properties that serve as discourse referents.

Consider my general analysis of statements like (19) given as (EXT) for statements like the following (from Chapter 4):

(19) There is a fictional detective, named 'Sherlock Holmes', but he does not exist-he is merely fictional after all.

(EXT) What you thought of as an object in the external world is not really an object in the external world

Furthermore, I had argued that in certain contexts, "x does not exist" functions to deny that something has a particular nature, rather than categorically deny the existence of a phenomenon. For instance, suppose a linguist like Deborah Tannen argues that mansplaining, where sexist men patronizingly explain and oversimplify things to women, does not exist. So, in that view, men do not do it to assert patriarchal dominance but do it to gain status through dominating a conversation, the way women achieve status by making connections and networks. Moreover, they do the same thing with both women and other

men, so it is not really sexist. Now suppose I go to someone, and without giving him and Tannen's explanation, claim "Mansplaining does not exist". The other person says, of course it does, look at this phenomenon where men do such things. In reply, I tell him, "The particular phenomenon that you think of as mansplaining is not really mansplaining but something else. Thus, there is no such thing as mansplaining". In effect, "x does not exist" functioned to deny of something that it had a particular nature.

Furthermore, (EXT) above links up with the purposes I mention in (MET<sub>FIC</sub>)\*, where asserting "x does not exist" is relative to our purposes, in this case our purpose being to deny that something has a particular nature.

(MET<sub>FIC</sub>)\* There is an x such that x is fictional and x is likened to a real person and x is and ought to be treated/counted as a real person for all relevant intents and purposes.

More specifically, Solodkoff's (2014) analysis of negative existentials is as follows, which I think is on the right track, where  $K_c$  is a contextually specified kind., although my claim was that it presupposes (MET<sub>FIC</sub>)\* to account for why n can be about a flesh and blood person.

"n doesn't exist" is true in a context c iff n isn't a  $K_c$  (*ibid.*, 338)

This gets the correct prediction in the context of Kripke's (2013) Moloch case. Consider Kripke's argument from Chapter 1 on the existence of Moloch. Moloch is construed to be a pagan god, but Moloch might mean Yahweh or a kind of sacrifice. In that case, it can be concluded that that Moloch was construed to be a pagan god was a mistaken assumption and there was no pagan god as Moloch. This point is highlighted from a joke with Harry Frankfurt. Kripke goes to Frankfurt and says that there is no such pagan god but that it was a sacrifice that was referred to by "Moloch". To this, Frankfurt responds, "Of course there was not such a god. You don't believe in pagan deities, do you?" (*ibid.*, 71) However, Frankfurt's response is indicative of an ambiguity, namely that there are two questions: the first is that "Was there a pagan god as Moloch?" and the second is "Is there such a legendary character?" (*ibid.*, 71) In this case, the answer to the first is no, but to the second, the answer can be yes or no, depending on what the empirical historical investigation reveals. But such a character exists if the answer is yes. Consider what I say about the contexts in which "x does not exist" functions to deny that something has a particular nature, rather than categorically deny its existence.

Frankfurt's comment categorically denies the existence of pagan gods, since he is not attentive to the purpose to which Kripke is putting the negative existential. On the other hand, Kripke says that Moloch isn't one contextually specified kind (pagan god) but another contextually specified kind, a sacrifice. He's saying Moloch does not have the nature of what it was construed to be before, where x was counted as y in context c (x being relevant mentally dependent properties associated construed as a pagan god in context c). So the assertion, "there is no such pagan god" is that look, Moloch is not this thing which has a particular nature, but it is of a different nature, because it is something else. So by denying its nature, what it was construed to be, Kripke is saying it's something else which has a different nature. It's not a pagan god. It's a sacrifice.

Consider an objection to my general account of fictional objects as mental entities. Crane (2001) recounts Gilbert Harman's objection that intentional objects like the fountain of youth cannot be mental objects because Ponce de Leon is not looking for something in his mind. However, this sort of objection is answered by Jackendoff (2002), who talks not only about perceptual features, but also about the "feels" attached to entities in our percepts, which he calls "affects" or "valuation" (Jackendoff's terminology on features is imprecise compared to mine where I would specify the property as the property of *being perceived*, *being thought of* or *being felt of being external*: I shall write [+EXTERNAL] for these for convenience and [+INTERNAL] for the internal counterpart). Jackendoff also draws attention to valuations like *familiar vs novel* (where "deja vu is a sense of familiarity associated with a situation known to be novel" (*ibid.*, 313)). The valuations *meaningful vs non-meaningful* come into play, for instance, when a schizophrenic finds his experiences more meaningful than other people and autistic people find less meaning than they're supposed to. On the internal vs external valuations, Jackendoff says:

Dreams feel external and non-self-produced but are of course internal and self-produced. People with phobias attach illusory negative value to all manner of inappropriate entities in the environment. Hence valuations too are constructions of the f-mind.

Returning to our bug, it is not simply experienced "out there" automatically: it is "out there" because the percept that gives rise to the experience contains the valuations external and non-self-produced. Visual processing normally settles on these valuations in response to retinal input: "seeing is believing." (*ibid.*, 313)

Thus Saul Kripke would have the valuation [+EXTERNAL] when someone is seeing him. On unicorns, Jackendoff says:

...consider the unicorn in my dream. There is a remembered entity which has an indexical feature and descriptive features in the visual modality (in particular a one-horned shape). In memory, it is assigned the valuation internal, so it is not experienced any more as having been “in the world.” Still, because it has an indexical feature, it can be connected to a referential expression. (ibid., 314)

Again, Jackendoff is imprecise in his terminology of features. He talks about a “one-horned shape” as a feature in the mind. This is simply false, because there is literally being one-horned or being red can only belong to one horned things or red things. I shall revise all of them to *being an appearance of being red*, *being an appearance* or *being a thought of being one horned*.

Nevertheless, the answer to Harman’s point is simply that if Ponce de Leon is not looking for something in his mind, it is because his thought has the valuation of the property [+EXTERNAL] because he thinks it is out there. If he finds out there is no such concrete place but is actually internal and is fictional, the valuation most likely would be revised to [+INTERNAL] since it in fact is. Moreover, to recall what I said above, just because one has a thought with the property being a detective doesn’t entail that there is no such person with the property *being a detective*. Analogously, just because a thought has the valuation *being a thought of being internal* doesn’t entail that the thought isn’t in fact internal. Of course, we might also mistake an external object to be internal, in which case, the valuation changes when we find out the object of our thought is an external object (on the presupposition that we are fully rational).

#### 4.1.4 Discourse Referents Clarified

In the light of the above discussion, consider both (12) and (12a) again:

(12)’ Bill doesn’t have a car [+EXTERNAL]. It [+INTERNAL] is black.

The better explanation for why both do not match up is because there are in fact two different discourse referents which are two bundles of mental representations encoding two bundles of properties as their content. So, the process can be described as follows: the first use of “a car” introduces a discourse referent (a bundle) and then the second “the case” introduces another bundle, with the same properties except that the valuation feature [+INTERNAL] is added. They’re two different bundles but, there’s no criteria available for counting them as the same. On the other hand, this explains why the reference gets resolved in (17)’ due to revising the valuation of a “unicorn” in the first sentence using (EXT) to

[+INTERNAL] because one knows that the properties are internal, as evidenced in the third sentence:

(17)' Bill didn't see a unicorn [+EXTERNAL]. The unicorn [+INTERNAL] had a gold mane. It's [+INTERNAL] there alright, but in his head.

This also matches with the intuition that both are about the same thing (recall what I said in previous chapters between facts about reference and facts about aboutness diverging; I shall also expand on this below), as in one looks at "Bill didn't see a unicorn. The unicorn had a gold mane", there feels like there's a good "connection" between the two NPs, but they just cannot "connect". I would add, they are about the same thing in a different manner or way, since they cannot match. It also looks like there is more of a "connect" between "a unicorn" and "the unicorn" than "a "unicorn" and "the centaur". The further intuition is that it's not just what is in common with the words but what we understand by them.

(14) Bill didn't see a unicorn. \*The unicorn had a gold mane.

Compare:

(14)' Bill didn't see a unicorn. \*The centaur had a gold mane.

(12)' Bill doesn't have a car. It is black. I mean his cat. Not the car. He doesn't really have one.

In (12)', in the first two sentences, two discourse referents are introduced, but they do not match because of different valuations. Then it is resolved in the third sentence when "his cat" is introduced. So there is a new bundle, but counted as about the same thing. Then we can see that "one" in the last sentence succeeds in anaphoric reference to two previous instances of "a car" and "the car".

In the case of proper names, it is more clear than indefinite NPs for Karttunen, where in the case of an unknown name, one records to memory what is said about him as well as the proper name and in the case of a known name, by means of the name, one recalls from memory the file associated with the name. If the situation model in fact involves mental construction of the model as well as the referent and the memory is also involved, then it looks like what seems to be merely a linguistic phenomenon can be given an ontological

interpretation by invoking what constitutes the mental processes involved in the construction.

#### 4.1.5 Reference, Not Denotation

Cumming (2014) points out that the phenomenon Karttunen points out is not really reference, but denotation. The following example is invoked (*ibid.*, 215):

(20) John came; so did Mary. The one who arrived first brought a cake.

In this case, he says that the definite NP, “the one who arrived first” in case it is a referring expression, might refer to John or Mary but it is not dependent on John or Mary anaphorically. Since this is so, the anaphoric denotation of the phrase is different from “John” as well as “Mary” and therefore anaphoric denotation and reference are not the same.

But Karttunen specifically mentions that definite descriptions in general pose a problem with anaphoric reference. So, it is a specific issue with the way definite NPs are associated with anaphoric reference and not whether there is anaphoric reference per se.

This problem of identification will be more difficult where a definite description—a definite noun phrase such as the man Bill saw yesterday, is used, since there will, in general, not be any simple look-up procedure for associating the description with the right individual. With definite noun phrases there is also the problem that it is not possible to tell just from the noun phrase itself whether or not it is supposed to refer to an individual at all. For example, it is clear that the phrase the best student is not used referentially in a sentence such as Bill is the best student. (*ibid.*, 364-365)

Second, Karttunen specifically writes about the need for discourse referents in the context of them needing to be invoked in the process of interpretation.

Suppose one says:

(20)' John came; so did Mary. The one who arrived first brought a cake. Mary must be given credit, since she arrived first.

It is clear in (20)' that the anaphoric denotation of “the one who arrived first” and “Mary” coincide. So the explanation for why in (20) the anaphoric denotation looks different from “John” or “Mary” has more to do with the interpreter’s epistemic state and uncertainty in the



absence of information, than it not being anaphorically dependent on John or Mary. It is in fact anaphorically dependent on John or Mary in the light of further information, since reference is secured in the mind of the interpreter the moment information comes in. This does not look very different from the process of anaphoric resolution that we saw in Garrod above, where the referent of “a starving beast” is resolved in favour of the dog rather than the tiger using further information in (10) and the overall situation model, despite the possibility of either of them being anaphoric referents. In the light of this, (20) can be interpreted as involving a discourse referent introduced by the occurrence of “the one who arrived first”. And the indexical feature in the abstract has been introduced as well as the descriptive feature *being a thought of being the one who arrived first* (both being in the same bundle. Mind-dependent properties of Mary are filled in when it is resolved in the third sentence to form a new bundle of representations.

## 5. Idiolects and I-Language

In this section, I shall justify how reference to fictional entities can take place in our minds, due to idiolects in our mind, since what the nature of language is, primitively. I shall bring up arguments in the literature on Chomskyan ideas of reference. I do not however adopt one radical conclusion from these arguments that languages like English, Chinese etc. do not exist, since I think that these languages form an average of mental systems in the head, and this can be made consistent with direct referential views at this level of explanation. I shall now out these arguments to show how at least such idiolects are possible.

Stainton (2006) argues against the view that natural languages such as English and Swahili involve shared public signs, words, phrases and sentences or have referents. One of the objections outlined by Stainton is that there is a problem in individuating words, especially as belonging to a language or a dialect. Commonsensical or philosophical understanding of “language” do not track any “robust divide” and do not track “nature’s joints”. (*ibid.*, 918) This is so because there is a problem with the notion of a language or dialect itself as commonly conceived. For instance, while Chinese is considered a single language as when a person says “y word in the language Chinese refers to x, there is a problem saying this since dialects of Chinese like Mandarin and Cantonese are mutually unintelligible. On the other hand, Italian and Spanish are considered different languages even though they are dialects of Romance even as there is much more mutual intelligibility between them than between Cantonese and Mandarin. One reply may be that one can replace “y word in dialect Mandarin refers to x” and individuate words as parts of dialects. But the problem again is

that Canadian English is a dialect, but there are rural, urban, Eastern, Central and Western Canadian varieties.

Stainton says that one that rather than public languages and dialects,

the real objects that one finds are (i) individual idiolects, (ii) sets of idiolects that share some non-obvious underlying parametric feature (e.g. having complements falling after heads), and (iii) the universally shared language faculty. None of these, however, corresponds even remotely to “public languages” like English and Urdu. (*ibid.*, 918)

If this is so, Stainton asks, what makes the different pronunciation of “photographer” in Bombay (‘fotoGRAFer’) and Toronto (‘foTAHgrafer’) the same word, different from one ‘fotOgrafo’ in Buenos Aires? One cannot without circularity say that it is because the former are English and latter is Spanish or even individuate it by saying it is part of a dialect. Moreover, one cannot appeal to local dialects, since there would be many differing pronunciations of ‘Harvard’ in the Eastern United States and even children and adults, men and women have different pronunciations. The problem is as bad if one says we individuate by appealing to the community that shares the symbol system, but the problem is that this too is circular, since one cannot individuate the correct linguistic community without appealing to shared language. Now the radical conclusion that Stainton draws is that if there is no entity without identity, and one cannot individuate these words as “words in English” “words in Swahili” etc., then such languages do not exist. If this is so, then there can be no “science of “word-world” relations that pairs “public words and sentences” with worldly objects, sets, and proto-thoughts.” (*ibid.*, 918) There are, however, non-linguistic explanations for these classifications. Thus, as Chomsky says “This idea [of a common public language] is completely foreign to the empirical study of language... What are called “languages” or “dialects” in ordinary usage are complex amalgams determined by colors on maps, oceans, political institutions and so on, with obscure normative–teleological aspects” (Chomsky 1993: 18–19 as quoted by Stainton *ibid.*, 918)

Of course, there are genuine worries about Stainton’s radical conclusion that languages like Swahili etc. do not exist and I think they ought to be conceded to. Stainton himself recounts the worry that this is too strong a standard for existence, since on the same grounds, human artefacts like corporations, songs, TV shows etc. will not exist. I think the best reply to the problem of individuation is Jackendoff’s (2012) view that English is an idealization or an average of the mental systems in our heads, an idealization that is done away with when child language acquisition is studied scientifically. For him, ordinarily, public language is understood as “out there” and not in the mind, the way we mistakenly but for useful

purposes think of sunsets, phenomenal colours and rainbows as existing externally, independent of us.

In all these cases, we project “out there” something that is really an amalgam of things going on inside the mind, and (non-obvious) things that are going on in the external world: “the structure of language is not “out in the world” but [is] rather a consequence of the mental organization of language users.” (Jackendoff 1987: 133 as quoted by Stainton). (Stainton 2006: 922)

Thus, there should be no problem of individuation anymore if we think of words as an average of mental systems in our heads.

Chomsky calls this internal, individual and intensional idiolect I-language, which is a property of the brain. (Chomsky 2000: 26-27) In my account, I consider I-Language as primitive, but I do acknowledge that considered in the light of how language is idealized, especially with regard to conforming to social norms, referential semantics has its place as well as a causal theory of reference. For my account, it merely suffices there are these idiolects, which we use to refer to mental representations and communicate thoughts, constrained by our knowledge of the world, our causal relation to the world, our human cognitive systems and human norms and interests. In fact, this is not a private language that is incommunicable, since our token mental representations and private associations have properties that are causally related to the world and to our common cognitive systems. Thus what we have in common with other people still forms the basis for communication.

### **5.1 Refers (to) as an Intentional (*sic.*) Transitive Verb**

In this section, I utilize Ermanno Bencivenga’s (1983) Neo-Kantian “Epistemic Theory of Reference” to show how “to refer” construed as an intentional transitive verb allows for reference to token mental representations. Bencivenga argues that “refers (to)” is an intentional transitive verb like “thinks (of)” in “x thinks of y” rather than a non-intentional verb like “hits” in “x hits y” (D’Ambrosio (2019) makes this case too although he calls them “intensional”, but Bencivenga is way earlier in the literature). For instance, suppose someone is attacked by animals and succeeds in hitting one of them with a stick. He goes back home and is asked “Which animal did you hit?”, to which he replies “I am pretty sure I hit a winged horse”. The reply “You can’t possibly have hit a winged horse as there is no such thing” would suffice to show to him the error of his belief. On the other hand, if someone says “I am thinking of a winged horse” to which another replies “You can’t possibly be thinking of a winged horse, for there is no such thing”, the reply would not show

that the belief is in error. One can always appeal to Brentano's thesis that "real existence" is not required for thought (I ought to say that there is reference to concrete objects or mental representations, since it is just that the nature of the referent is different. However, in my view being is univocal). On the other hand, verbs like "thinking (of)" for which existence is not necessary (in my view, it's just the existence of concrete objects in the external world which is not necessary) are intentional transitive verbs.

Bencivenga says that even if "refers (to)" is ambiguous between both, then it seems likely that the intentional reading is more primary for reasons of economy, one does not need the existence of the external world in order to refer. Furthermore, for Bencivenga winged horses and Sherlock Holmes do exist, as they exist in the conceptual space to which they belong. (*ibid.*, 791)

For clarity, let's take consider Bencivenga's example (*ibid.*, 793):

(25) The brown table over there is rectangular.

Now suppose there is no really existing brown table over there or it's a blue one. Bencivenga then poses a dilemma "either you say that I am not really referring to what I think I am referring to, or I find my reference in a different world from the actual one." (*ibid.*, 793)

If it is the former, then it is not possible because "referring" has an "intentional character", that is, "if I *intend* a given object, it sounds illegitimate for somebody else to tell me that this is not what I "really" intend." (*ibid.*, 793) If the latter, then one refers to it in the non-actual world (non-external "world", that is in the mind in my view) of the speaker's cognitive space, that is, a world "determined and shaped" by speaker's belief system. Moreover, he is not referring to senses since senses are not winged horses or brown tables (and therefore fail to capture the phenomenology). It is also not a Meinongian view in which the brown table is a nonexistent object, since the brown table might exist and I might be referring to an existent brown table in that circumstance, if he said (*ibid.*, 793):

(26) The existent brown table over there is rectangular.

It is instructive to see how Bencivenga's account deals with Kripke's Gödel-Schmidt case. Kripke's argument is as follows. Descriptivism about names holds that "Gödel" means "the

man who proved the incompleteness theorem". But Gödel might have stolen it from a man named Schmidt. If descriptivism is true, then when someone says "Gödel lived in Vienna", he is referring to Schmidt. But he is not referring to Schmidt, but Gödel. Thus, descriptivism is false since it leads to the wrong result. However, Bencivenga also says that a person might utter the following in the same circumstance as the utterance "Gödel lived in Vienna" (*ibid.*, 795)

(28) The man who proved the incompleteness of arithmetic was born in Vienna.

With (28) it is plausible that the person is referring to Gödel. Thus, with the description too, one does not need to possess the "real" properties of Gödel to refer to him. Of course, Bencivenga says that one might invoke Donnellan's referential/attributional distinction to say that this is the referential use, but the description does refer to Schmidt in its attributional use. But he points out a problem with this reply:

And here is a pretty serious problem: how is it that descriptions, apparently the referring expressions whose use depends most strictly on the correct identification of some property or properties of the object of reference, may succeed in referring to objects that blatantly contradict the description used?

I have an answer for this question. Reference is primarily an intentional operation, which is performed in an intentional, cognitive domain. The properties "real" objects "really" have do not matter for reference because reference is not to such objects. But of course, the properties objects in cognitive spaces have do matter. It is only because the ordinary man's cognitive space contains a Gödel who proved the incompleteness of arithmetic that he can succeed in referring to Gödel (whatever the case "really" is) by the description 'the man who proved the incompleteness of arithmetic'. He could not, and would not, refer to Gödel (referentially or otherwise) as 'the man who discovered America' (unless his cognitive space were pretty peculiar). (*ibid.*, 795-796)

However, the flaw in Bencivenga's account is that from reference being primarily intentional, it does not follow that "real" objects do not matter for reference and that reference cannot be to them. As I said above, I can think of the property *being a detective* with the property of my thought *being a thought of being a detective*, but it does not entail that there is no concrete detective of which it is a thought. Similarly, with a chair. I can think of a chair, but it does not entail there is no concrete chair I am thinking of. His argument can still follow through, since in (25), "the brown table" can directly refer to mental representations in cognitive space as well as the concrete brown table when there is one in (26). So, Bencivenga overstates his case. In my view, this is acknowledged, since there is direct reference due to the confrontation relation, and mistaken properties of Gödel in

cognitive space are part of representation\* and can be corrected by a person with knowledge of the referent. So Kripke is still Kripke in all possible worlds but with a difference between how Kripke is and how he is thought of as, and with respect to the latter, is that we refer to mental representations.

Nevertheless, from this, Bencivenga gives a diagnosis of the problems with the “Frege-Russell” view of names. In it, there is a clear demarcation between a description and a name: a name picks out an object in the real world by means of its associated description and a description picks out the unique object that is described by it.

On the other hand, when the intentional character of referring is rightly emphasized, things fall back into place. It then becomes possible to say again that descriptions refer to the entities described, and names like ‘Gödel’ get their reference (in a cognitive space) through association with some description. And this is reasonable, since for most of us Gödel is nothing more nor less than the man who proved the incompleteness of arithmetic. (*ibid.*, 796)

## 6. Objections

### 6.1 Objections on Aboutness and Reference

One of the key objections to Everett and Schroeder’s account we saw above was that fictional characters are not ideas, just as the idea of the law is not the law itself. While they responded to the objection by pointing out that unlike laws, fictional characters are not public institutional entities like the law, and could be thought of even without having created any work of fiction. However, the worry about this still persists, since we do not think that we are thinking about an idea when we think of a fictional character, but a flesh and blood person. In what follows, I shall attempt to solve this issue by arguing that there is a distinction between facts about reference and facts about aboutness, which accounts for how we can refer to mental representations and yet our thoughts are about flesh and blood persons.

In Chapter 4, I looked at Hayaki’s objection to Salmon, due to which I made a distinction between facts about reference and facts about aboutness. Hayaki’s objection to Salmon’s single tiered account, where names within fiction as well as metafictional statements refer to abstract objects, was as follows:

Consider the following statements (*ibid.*, 141):

(29) The Canon is about Sherlock Holmes.

(30) The Canon is about a brilliant private detective who solves many crimes that baffle Scotland Yard.

(31) The Canon is also about Queen Victoria (among other things).

While (29) is about Queen Victoria and (31) in Salmon's account is about an abstract object, since the fictional character Holmes is an abstract object. Thus not only are (29) and (31) disanalogous, (29) is also supposed to be about a flesh and blood person, not an abstract object. Furthermore, Hayaki argues that aboutness is not a genuine relation, in the context of "the Canon is about x", since in (30), the Canon is about a brilliant detective but since there is no such detective, there is no relation between the canon and a detective. This makes such a context an intensional one.

One response that I say is available to Salmon is that in fact (29) and (31) are not disanalogous is because since it is a fictionalized account of Queen Victoria, it is pretended of her that she is a fictional character. This accounts for cases like the Schlomes case I bring up, where based on the metafictional principle (F) below, that may be introduced by an author, an author can fuse (not a mereological fusion) Queen Victoria and Holmes to be a new character Schlomes.

(F) For any actual or fictional characters x and y appearing in a work of fiction Z, if x and y are fused into conjoined twins, a new object is thereby created in the world of the fiction.

So in the Schlomes case, one pretends that Queen Victoria and Holmes are fused like twins and becomes one character. From this, I had made two observations. The first is that facts about reference and aboutness can diverge. The names refer to two entities, but pretense allows Schlomes to be about one character, thought of as a fused twin. The second was that facts on aboutness can be accounted for by a psychological mode of presentation where there is a way our minds present something to us. The reason why I mention psychological modes of presentation is because mental representations which are computational states themselves, in the absence of environmental input are simulations of the real world and have intrinsic properties which recreate the experience of the world to us. (Bergen 2012)

The place where I diverged from Salmon's possible response above however, was in the light of a further objection from me. If we in fact pretend that Queen Victoria and Holmes are a new character called "Schlomes", in Salmon's account, since the pretend use of a name gives rise to a new character, there is in fact a new character in the fiction. To deal with this

problem, I proposed that the fictionalized version of Queen Victoria is constituted by mental representations which encode properties abstracted from the real Queen Victoria and Holmes involves mental representations which encodes properties abstracted from real people in general. In creating a new fused character, the two bundles of mental representations are combined to form a new bundle which is counted as the same character<sup>52</sup>. My account had the distinct object that whereas it is difficult to account for such a fusion for an abstract and a concrete object, in this case, it is mental representations that combine.

My account is also consistent with facts about aboutness and reference diverging. In my account, unlike abstract objects, fictional names refer to unified mental representations which are formed by bundles of mental representations but it is about a flesh and blood person due to the “likened to” relation as well as a combined effect of the processes involved in mental representation. Thus, “x is about y” is explained in terms of a more primitive relation, “x is likened to a flesh and blood person” as well as processes involved. This would also account for a similar objection that was raised against Salmon by Hayaki that can also be levelled against Everett and Schroeder, namely that we do not think of ideas

---

<sup>52</sup> In the context of my discussion of discourse referents as well as the Schlomes case, we are in fact referring to mental representations, especially since discourse referents occur not when we are in a confrontation relation with an object but within Matravers’ representation\*. Firstly, reference within the situation model allows me to claim that I am referring to a surrogate representation of the world, which is better understood as a simulation (see p. 279 above). Secondly, the Schlomes case allows me to claim that we are in fact referring to abstracted properties from Queen Elizabeth which are mind dependent and which allow us to fuse mental representations that encode these properties with mental representations that encode Holmes’ properties to form a new character Holmes. However, Friend (2011) argues against fictional surrogates and argues that “London” in Orwell’s 1984 actually refers to London, since the reference to London and descriptions of the actual London are interspersed in the novel even as the author describes how Airstrip One used to be called England or Britain but London in the novel was still called “London” before the war. (*ibid.*, 192) More pertinently, she argues that one is prescribed imagining about London since “Insofar as the city is fictionalized, the distortion is crucial to the power of Orwell’s vision: readers are supposed to imagine what London would be like if England had pursued a certain path.” (*ibid.*, 192) My argument against her (suppose she’s an abstractionist about fictional objects) would be analogous to the Schlomes case, where I can always think of Schmlondon, which is a new fictional city, a fusion of London and Atlantis. Since an abstract entity (the Holmes character) cannot literally fuse with a concrete entity, the fusion is only possible if I am merging mental representations and I am referring to the London of my conceptual space, which is my idea of London (see Bencivenga above). Nevertheless, the reason why Friend’s intuitions persist is because the mental representation has R properties, but also has W properties. While we refer to the R properties in the representation (and representation\*) which forms part of descriptive meaning and involves conceptual metaphor due to which we think of the fictional London as a bounded entity, the normative aspect of meaning which comes from knowledge of W Properties (because the normative meaning of “London” is London) is what enables us to think of the London of 1984 in line with the city of London in the external world. This is consistent with us being able to refer to London as well as our idea of London, in which may get properties of London wrong. Thus, since we do need a conceptual metaphor to think of the London in the fiction and we refer to the London in our conceptual space, we need to presuppose (MET<sub>FIC</sub>\*).



or mental representations when we think of a fictional character, but of a flesh and blood person.

Furthermore, the representations are stored away in our memories in mental files. Fictional characters, since they “flit in and out of existence”, exist only when we think of them. I can think of some of Holmes’ attributes and yet think of Holmes, because the aboutness of my thought is determined by my application of constitutive rules. That I can think of Holmes without thinking of all of Holmes’ attributes is precisely why I think of a fictional characters as constituted by bundle of mental representations at time  $t$  a subject  $x$  thinks about.

There is another phenomenological objection concerning the aboutness of the properties involved. The objection is that when a person thinks of oneself as a human being and Holmes as a human being, it looks like the same property is involved. The property does not all of a sudden become mind-dependent when it is involved in a thought about fictional characters. My response is as follows.

Almost everyone (if not all) in the literature acknowledges that the same property is involved in two different ways. Van Inwagen would say fictional characters *hold* them versus Kripke has the property *being a human being*. There are issues with that which we saw in Chapter 2. Thomasson would say it is pretended that Holmes has the property but he does not really have it. In the scope of pretense, is also unclear whether it is pretended of the abstract object that it has the property or like in Walton it is a pretend property or it is pretended of Holmes that he has that property. If we say it is a pretend property, I have already outlined the problem that it does not account for content. If we say we pretend that an abstract object has the property we end up with my Schlomes problem to which I offered a plausible solution above.

There is also a problem with Everett and Schroeder’s account as we saw above. Plus my account is a uniform account with no strict distinction between fictional and metafictional contexts. The “according to the fiction” operator has no semantic role and is redundant. Pretense and fictionality is also irrelevant to semantics (also from Matravers’ account). So I cannot allow for mind independent properties (according to the fiction) as being the content of mental representations for fictional characters.

As for the Neo-Meinongians, is it any less plausible that properties like *being a human* exist apart from flesh and blood people, in Platonic heaven, only to be encoded in a fictional character and exemplified in a flesh and blood person or with two types of them, nuclear and extra-nuclear? We also saw Thomasson’s objection to that above.

Moreover, Crane (2013) argues that since properties constitute the nature of things, properties in the external world are existence entailing (my version would be that the entail the existence of a concrete object). So the property being a penny would entail the existence of a penny. It just seems very strange that “existence entailing properties” would be the content for fictional characters.

More pertinently, even I say the same property is involved in two different ways. Kripke has the property *being a human*. The property in the external world causes me to have thoughts of the property being a human. The thought is partly individuated by it having been either caused by the property *being a human* in my egocentric space or there is a causal link in the past. Surely the property is involved, but in an indirect manner. This is more so, since my mental representation carries information on that property. I’m just saying the same information is passed on when it comes to a fictional character. Also, when the objector says “When I think of Holmes as a human being and when I think of myself as a human being, the same property is involved”, it’s not clear. When one thinks of Holmes literally as a human being? No one thinks that. Why should Holmes literally have the same property? I explain the intuition of my objector by saying that he or she thinks of Holmes as a human being because of the information encoded in their mental representation which likened to a flesh and blood person substance. There is nothing more required.

## 6.2 A Fregean Objection

The most direct objection to the views I have laid out might be that if we take fictional names to refer to token mental representations whose content is given by mind-dependent properties, then they become personal associations. However, personal associations ought to be set apart from meaning. Frege says:

If the reference of a sign is an object perceivable by the senses, my idea of it is an internal image, arising from memories of sense impressions which I have had and acts, both internal and external, which I have performed. Such an idea is often saturated with feeling; the clarity of its separate parts varies and oscillates. The same sense is not always connected, even in the same man, with the same idea. The idea is subjective: one man's idea is not that of another. There result, as a matter of course, a variety of differences in the ideas associated with the same sense. A painter, a horseman, and a zoologist will probably connect different ideas with the name 'Bucephalus.' This constitutes an essential distinction between the idea and the sign's sense, which may be the common property of many and therefore is not a part or a mode of the individual mind. For one can hardly deny that mankind has a common store of thoughts which is transmitted from one generation to another. (Frege 1948: 212)

Jackendoff's argument against this is to take a case of two friends, who might have grown up together and have personal associations with some words. In fact, not only can they communicate them with each other, but also the very fact of them communicating them with each other makes the meanings "public". It is merely that the meaning is restricted from the rest of the world and it is tough, if not impossible for outsiders to "get" them. Similarly, in a larger community, linguists have associations with the word "language" by means of shared experience, and these associations are not what a linguist communicates to a dentist while using the same word language, even though the word he is using is the same.

So given Jackendoff's point, the problem that I see with Frege's picture is that even though the painter, the horsemen and the zoologist will use similar shared associations of "Bucephalus" when they are communicating with each other, they do in fact succeed in communicating shared associations of "Bucephalus" within their own professional communities. In that respect, their meanings are public. Moreover, Jackendoff says, that for a linguist like him, he needs to estimate which associations of his he can share, based on what he finds common with a continuum of communities ranging from neuroscientists, philosophers, biologists and psycholinguists. So, it is the same word "language" that he uses but he chooses to select which facets of the meaning to communicate so he can be understood. With a popular audience, he uses the least specialized meaning which he finds in common with them.

Thus, the explanation for why people succeed in communicating public meanings is not that senses and associations are two different kinds, but because of a type of Gricean principle: "one produces utterances in such a way that one can expect the hearer to reconstruct the intended message." (*ibid.*, 2002: 284) So, when one is trying to prevent ambiguity or make himself understood, he uses this principle. Moreover, in my view, they also succeed in communicating, since they are causally related to the same properties as well as objects in the world their thoughts represent.

The picture becomes clearer when one sees that if one is trying to communicate about a fictional character to another person, the other person recognizes predicates and descriptions that the other person talks about and relates them to his associations in his own experience. Thus, the predicates are applied to properties in his own head. Nevertheless, it is not entirely subjective, since the properties, as I wrote before are causally related to properties of objects in the external world and carry traces of them, as Jackson says.

If a person has no associations with the meaning of a predicate, he asks the speaker to explain the meaning of the predicate to him. Thus, the speaker thinks of other predicates and

descriptions by means of which to communicate the same thing in other words that the other person can relate to his own experience and thoughts. If he cannot relate to them, he asks again. None of this shows that Direct Reference theories are false. In fact, it shows that someone with a knowledge of the referent appropriately related to a causal chain can correct a person with false associations with the referent like in the Gödel-Schmidt case as I said above, that if someone mistakenly thought of Gödel as the man who solved the incompleteness theorem even though Schmidt did, a person who has knowledge of the referent can correct him. This is consistent with saying that in I-Language idiolect, the mistaken person thinks of a different meaning of Gödel. This meaning is not *normative* and yet, it can be *described* as meaning. The normative meaning comes from knowledge of the referent. Thus, there is a distinction between descriptive (descriptive not in the sense of using “descriptions” but in the sense of normative vs descriptive grammar) and normative meaning.

This explains the reason why in the case of the nerd and the other person getting “Holmes” all wrong, the nerd can correct the latter, since he has appropriate knowledge of conventions and other norms associated with Holmes. This also explains why two people can stipulate or even one person can stipulate that Kripke is a clown and associates completely different properties with him and carry on a conversation with each other associating newer properties as the conversation proceeds. Then then a third person overhears them and corrects them based on a knowledge of the referent and they tell him it was a stipulation. This is almost Humpty Dumpty’s theory of meaning, in which one can make a word mean whatever they want to mean, descriptively speaking, except that the common world and common cognitive mechanisms come into that context, so communication can happen even then and the normative context comes in as well. So they think of Kripke by referring to properties encoded by mental representations, like *being a thought of Kripke being a clown*, *being a thought of Kripke being y* etc. in their representation\*. Their thoughts are causally related to Kripke, because of which they can be corrected. Since their thoughts are partly causally individuated, they might not even know that their thoughts are about Kripke<sup>53</sup>.

---

<sup>53</sup> It is in place to add here how mental representations with mind-dependent content and mental representations with mind independent content are to be distinguished. Simply put, the distinction is that unified (complex) representations that correspond to objects in the external world have mind-independent content while those that do not correspond only have mind-dependent content. However, it will be in place for me to elaborate here.

As we saw, mental representations in representation\* when our mental states are offline and not caused by a perceptual input, can be discourse referents. Now a representation of Kripke has R-Properties as well as W-Properties, where R-properties indicate W-

Moreover, as I mentioned in Chapter 5, Chomsky replies to Lepore and Fodor's (1992) worry that if Chomsky holds his view on I-Language as a property of the brain as opposed to it being an external social object (E-Language) then there is problem with possibility of "communication between time slices of an idiolect" in the absence of "shared meanings", "shared reference" and "shared language" :

It may be that when he listens to Mary speak, Peter proceeds by assuming that she is identical to him, modulo M, some array of modifications that he must work

---

Properties of objects in the world. When it comes to the confrontation relation too, R-properties indicate W-properties. But both ways, mental representations do in fact have content that is mind-dependent, which allows us to distinguish between our idea of x, which we can get wrong with the wrong properties and x itself. For instance, this allows us to say things like "this matches my idea of it" or "this does not match my idea of it", where the R-properties match or do not match the W-properties.

Nevertheless, as I wrote in the first chapter, that it is better to think of mental representations as simulations rather than representations, where while some of the properties of simpler representations resemble or correspond to properties of objects in the external world that they are simulations of, the simulations have sui generis properties too. Furthermore, while simple representations can correspond to properties in the external world, and as I said the mind dependent properties of these representations are individuated partly in terms of being causally linked to a mind independent property in the egocentric space or in the past. However, when it comes to unified representations that emerge out of the bundling of simpler representations, since representations also represent themselves, there can in fact be a unified representation without there being an object it corresponds to. The unified mental representations (simulations) with mind independent content are merely those that have resemblances/correspondences to objects in the external world. The reason why the representations in fiction have mind dependent properties as their content is because they are "decoupled" from representations of the world.

However, the picture gets slightly more complicated, since there are two levels of content for misrepresentations too (R properties and W properties), as well as the content of names in idiolects, which I mentioned before as referring to discourse referents too, in representation\*. And what determines whether we take up the W properties or the R properties as the content of the same representation is based on our interests. So for instance, I can talk about my idea of evolution where I am focusing on the R properties but I can talk about evolution as such, where I'm looking at W properties of my idea. They get their descriptive content from mind dependent properties, and yet since they are surrogate representations which also are directed at the world, the normative content based on which I can correct what I think of, is a mind independent property. For instance, when a representation is causally linked to a cow but it represents it as a horse, it counts as a case of misrepresentation. Causal theories of mental representation have a problem with such cases, since the representation turns about to be about a horse, but if what it is causally linked to determines its content, it will be about a cow. Moreover, one cannot explain why it is a misrepresentation, since the representation is about a horse. If however, we take one level of the content of representation which is about a horse as mind dependent, we can identify the misrepresentation by means of its content, and say it is about a horse and then say that it is a misrepresentation because of the fact that it is causally linked to a cow and ought to show the properties of the cow, but it is about a horse. I also recognize that it is about a horse, since I have been related causally to properties of a horse in the past.

out. Sometimes the task is easy, sometimes hard, sometimes hopeless. To work out M, Peter will use any artifice available to him, though much of the process is doubtless automatic and unreflective.' Having settled on M, Peter will, similarly, use any artifice to construct a "passing theory" - even if M is null. Insofar as Peter succeeds in these tasks, he understands what Mary says as being what he means by his comparable expression. The only (virtually) "shared structure" among humans generally is the initial state of the language faculty. Beyond that we expect to find no more than approximations, as in the case of other natural objects that grow and develop. (Chomsky 2000: 30)

## **7. Everett's Problem of Indeterminacy**

The account I have provided of mental representations and properties that are the content of these representations answers a key problem facing realist theories from Everett. Consider again the following case that I speak of in Chapter 5:

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. (Everett 2005: 629)

If it is assumed that people are unsure because of metaphysical indeterminacy, the realist would be bound to take up a contradiction that Frick is distinct from Frack and Frick is the same as Frack. However, my theory solves this problem. The interest relative identity of fictional characters is determined by our information states, and we can count the same bundle of mind-dependent properties as two different characters and two different bundles as the same character. So in cases where Frick and Frack are seen in different places at the same time, one merely suspends judgement about whether the two unified mental representations that "Frick" and "Frack" refer to are the same character or not. So Frick and Frack would be set apart in the mind as two bundles of properties since (two unified) representations are referred to by two names, and in parts of the story when it is said that they are simultaneously in two places, one thinks of the two bundles. But in the latter cases of Frick changing to turn into Frack, one thinks of one bundle by merging the two bundles. But one suspends judgement on whether the two bundles and the single bundle are to be counted as the same character.

## **8. Conclusion**

In conclusion, I have argued for why fictional entities, which are ideas as Everett and Schroeder content, are constituted by mental representations whose content are mind-dependent properties they encode and yet at another level of explanation, they are constituted by our social practices and their identity is interest relative. I started by laying out Everett and Schroeder's view that fictional characters are ideas, as well as problems they face explaining what the content of ideas of fictional character are. After giving a revision of van Inwagen's argument for fictional characters, which forms my master argument, an account of mental representations from Thagard was given, showing how simpler mental representations are bound to form unified representations. I then gave arguments about what these properties encoded by mental representations are and accounted for representational properties of mental representations. I then argued from Kripke, that since fictional characters have no internal nature, they cannot be substances in the external world. Moreover, they cannot be eternal and necessary entities either, since the consequence of a view is that we end up with a bare particular, which is unpalatable. I explained that the reasons for van Inwagen's mistake and Kripke's argument was that both were actually thinking of mental representations and the properties they encoded. Furthermore, if we take fictional entities to be non-spatiotemporal abstracta that cannot "flit in and out of existence", dependent on a spatiotemporal entity like a literary work, then we end up with a counterexample which shows that they can in fact "flit in and out of existence". From this, I argued that, since mental representations are the sorts of things that do flit in and out of existence, and we think about fictional characters, the most likely explanation for why that is so is that they are mental entities.

Following this, I argued, inspired by James Miller's (2019) account of a bundle property of words, interest relative identity of fictional characters is possible via content.

Then I argued, using empirical work on reference processing and Karttunen's work on discourse referents, that reference to fictional characters is possible in representation\*. I also utilized Chomsky's I-Language idiolects and Bencivenga's view that "refers (to)" is an intentional transitive verb, that reference to such mental representations is possible in the way I have described it. I also attempted to answer possible objections to the views I have laid out, especially showing how facts about reference and aboutness diverge.

In the light of the cumulative picture, I showed a major advantage of my view, which is a modification of van Inwagen's and Everett and Schroeder's, in which a solution to the

thorny problem of metaphysical indeterminacy of fictional characters that abstractionist views face was proposed.



## BIBLIOGRAPHY

- Armstrong, D. M. (1989). *Universals: An Opinionated Introduction*, Westview Press.
- Barsalou, L. W. (1999). Perceptual Symbol Systems. *Behavioral and Brain Sciences*, 22(4): 577–660.
- Bencivenga, E. (1983). An Epistemic Theory of Reference. *The Journal of Philosophy*, 80(12): 785-805.
- Bergen, B. K. (2012). *Louder than words: The new science of how the mind makes meaning*. Basic Books.
- Braun, D. (1993). Empty Names. *Noûs*, 27(4): 449-469.
- Brock, S. (2002). Fictionalism about Fictional Characters. *Noûs*, 36(1): 1-21.
- (2010). The Creationist Fiction: The Case Against Creationism About Fictional Characters. *Philosophical Review*, 119(3): 337-364.
- Caplan, B. (2004). Creatures of Fiction, Myth, and Imagination. *American Philosophical Quarterly* 41 (4): 331-337
- Carter, W. R. (1980). Do Creatures of Fiction Exist?. *Philosophical Studies*, 38 (2): 205 - 215.
- Chandler, R. (1955). *The Little Sister*. Harmondsworth, Penguin.
- Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge University Press.
- Contessa, G. (2009). Who is Afraid of Imaginary Objects? In N. Griffin & D. Jacquette (Eds.), *Russell Vs. Meinong: The Legacy of "On Denoting"*. Routledge.
- Corballis M. C. (2013). Wandering Tales: Evolutionary Origins of Mental Time Travel and Language. *Frontiers in Psychology*, 4, 485: 1-8.
- Crane, T. (2013). *The Objects of Thought*. Oxford University Press.
- Crimmins, M. (1998). Hesperus and Phosphorus: Sense, Pretense, and Reference. *Philosophical Review* 107 (1):1-47.
- Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics*. (Vol. 30). John Wiley & Sons
- Cumming, S. (2014). Discourse Content. *Metasemantics: New Essays on the Foundations of Meaning*. Oxford University Press
- Currie, G. (1990). *The Nature of Fiction*. Cambridge University Press.
- D'Ambrosio, J. (2019). Semantic Verbs Are Intensional Transitives. *Mind*, 128(509): 213-248.
- Dennett, D. C. (1981). *Brainstorms: Philosophical Essays on Mind and Psychology*. MIT Press.
- (1987). *The Intentional Stance*. MIT Press.
- (1988). Precis of the Intentional Stance. *Behavioral and Brain Sciences* 11 (3): 495-505.
- (2006). *Breaking the Spell: Religion as a Natural Phenomenon* (Vol. 14). Penguin.

- (2013) *Intuition Pumps and Other Tools for Thinking*. W. W. Norton & Company
- Edwards, D. (2014). *Properties*. John Wiley & Sons. Kindle Edition. Retrieved from Amazon.com
- Evans, G. (1982) *The Varieties of Reference*, Clarendon Press.
- Everett, A. (2005). Against Fictional Realism. *Journal of Philosophy*, 102(12): 624–649.
- (2007). Pretense, Existence, and Fictional Objects. *Philosophy and Phenomenological Research*, 74: 56-80.
- (2013). *The Nonexistent*. Oxford University Press.
- Everett, A. and Schroeder, T. (2015) Ideas for Stories. In S. Brock, and A. Everett (Eds). *Fictional Objects* Oxford University Press.
- Fodor, J. A., & Pylyshyn, Z. W. (2015). *Minds without Meanings: An Essay on the Content of Concepts*. MIT Press.
- Frege, G. (1948). Sense and Reference. *The Philosophical Review*, 57(3), 209-230.
- Friedman, O., & Leslie, A. M. (2007). The Conceptual Underpinnings of Pretense: Pretending Is Not 'Behaving-as-If. *Cognition*, 105(1): 103-124.
- Friend, S. (2000) Reviewed Work: *Fiction and Metaphysics* by Amie L. Thomasson. *Mind*, 109 (436): 997–1000.
- (2007). Fictional Characters. *Philosophy Compass* 2 (2): 141–156.
- (2011). The Great Beetle Debate: A Study in Imagining with Proper Names. *Philosophical Studies* 153: 183– 211.
- (2014). Notions of Nothing. In Manuel García-Carpintero & Genoveva Martí (Eds.), *Empty Representations: Reference and Non-Existence*. Oxford University Press.
- García-Carpintero, Manuel (2020). Co-Identification and Fictional Names. *Philosophy and Phenomenological Research* 101 (1): 3-34.
- Garrod, S. (2011) Referential Processing in Monologue and Dialogue with and without Access to Real-World Referents 273 in Gibson, E., & Pearlmuter, N. J. (Eds.). (2011). *The Processing and Acquisition of Reference*. MIT Press
- Visual context for interpreting sentences (1) and (2). Reprinted (Permission Awaited) from Garrod, S. (2011) Referential Processing in Monologue and Dialogue with and without Access to Real-World Referents (p. 275) in Gibson, E., & Pearlmuter, N. J. (Eds.). (2011). *The processing and acquisition of reference*. MIT Press.
- Goodman, J. (2004). A Defense of Creationism in Fiction. *Grazer Philosophische Studien* 67 (1):131-155.
- Gopnik, A. (2010). *The Philosophical Baby*. Random House..
- Grady, J. E. (2007). Metaphor. In Geeraerts, D., & Cuyckens, H. (Eds.). (2007). *The Oxford Handbook of Cognitive Linguistics*. Oxford University Press.
- Greene, R.L. (2014, July). Johnson: Why Men Interrupt. *The Economist*. Retrieved from <https://www.economist.com/prospero/2014/07/10/johnson-why-men-interrupt>

- Guthrie, S. (1993). *Faces in the Clouds: A New Theory of Religion*. Oxford University Press.
- (2008). Spiritual beings: A Darwinian, cognitive account. In J. Bulbulia, R. Sosis, E. Harris, R. Genet, C. Genet, & K. Wyman (Eds.) *The Evolution of Religion: Studies, Theories, and Critiques*. Collins Foundation Press
- Haack, S. (2020). The World and How We Know It: Stumbling Towards an Understanding. *Journal of Critical Realism*, 19(1): 78-88
- Hanley, Richard (2009). Fictional objects. In Robin Le Poidevin (Ed.), *The Routledge Companion to Metaphysics*. Routledge.
- Hauser, L. (2002). Behaviorism. *Internet Encyclopedia of Philosophy*. Retrieved from <https://www.iep.utm.edu/behavior/>
- Hayaki, R. (2009). Fictional Characters as Abstract Objects: Some Questions. *American Philosophical Quarterly* 46 (2):141 - 149.
- Hicks, Michael. (2010). A note on pretense and co-reference. *Philosophical Studies* 149 (3):395–400.
- (2015). Pretense and fiction-directed thought. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 172 (6), 1549-1573
- Howell, R. (1996). Review of *Mimesis as Make-Believe*. *Synthese* 109: 413–434.
- (2002). Reviewed Work: *Fiction and Metaphysics* by Amie L. Thomasson. *The Philosophical Quarterly*, 52 (207):282-284.
- Iacona, A, and Voltolini, A. Reviewed Work: *Fiction and Metaphysics* by Amie L. Thomasson. *Dialectica*, 56(3), 284-287.
- Jackendoff, R. S. (1988) . Conceptual Semantics. In Eco, U., Santambrogio, M., & Violi, P. (Eds.). (1988). *Meaning and mental representations* (Vol. 496). Indiana University Press
- (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press.
- (2006). Locating meaning in the mind (where it belongs). In R. J. Stainton (ed.), *Contemporary Debates in Cognitive Science*. Blackwell.
- (2012). *A user's guide to thought and meaning*. Oxford University Press.
- Jackson, F. (2015). Reviewed Work: *Minds without Meanings: An Essay on the Content of Concepts*. Jerry A. Fodor and Zenon W. Pylyshyn. MIT Press,
- Jansen, L. & Näger, P. (eds.) (2018). *Peter van Inwagen. Materialism, Free Will and God*. Springer.
- Jung EM., Pellet F. (2018) Theoretical Entities of Literary Criticism and Science: What Mrs. Gamp and Electrons Do Not Have in Common. In L Jansen. & P. Näger (Eds.) *Peter van Inwagen. Münster Lectures in Philosophy*, vol 4. Springer
- Karttunen, L. (1976). Discourse referents. In J. D. McCawley, ed., *Syntax and Semantics*, vol. 7. Academic Press.
- Kim, J. (1996). *Philosophy of Mind*. Westview Press.
- King, J. C. (2007). *The nature and structure of content*. Oxford University Press.
- Kripke, S. A. (2013). *Reference and Existence. The John Locke Lectures*. Oxford University Press.

- (1972/1980). *Naming and Necessity*. Blackwell.
- Kroon, F. (2013). The Social Character of Fictional Entities. In C. Barbero, M. Ferraris, A. Voltolini (Eds.) *From Fictionalism to Realism*. Cambridge Scholars Publishing.
- Kroon, F. and Voltolini, A. (2018). Fictional Entities. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2018 Ed). <https://plato.stanford.edu/archives/win2018/entries/fictional-entities>
- Lakoff, G., & Johnson, M. (1980). Conceptual metaphor in everyday language. *The Journal of Philosophy*, 77(8): 453-486.
- (1980/2003). *Metaphors we live by*. University of Chicago Press.
- (1993) The Contemporary Theory of Metaphor. In A. Ortony (ed.), *Metaphor and Thought*. 2d ed. Cambridge University Press.
- (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought* (Vol. 28). Basic books.
- Lamarque, P. (1993). Reviewed Work: *The Nature of Fiction*. by Gregory Currie. *The Philosophical Quarterly*, 43(171): 253-256
- (1996). *Fictional Points of View*. Cornell University Press.
- (2010). How to Create a Fictional Character. In P. Lamarque *Work and Object: Explorations in the Metaphysics of Art*. Oxford University Press.
- Landman, F. (1986). Pegs and alecs. In *Towards a Theory of Information: The Status of Partial Objects in Semantics*. Dordrecht: Foris
- Lewis, D. (1978). Truth in fiction. *American Philosophical Quarterly*, 15(1), 37-46.
- Lodge, D. (1993). *The Art of Fiction: Illustrated from Classic and Modern Texts*. Viking
- Ludlow, P. (2011). *The Philosophy of Generative Linguistics*. Oxford University Press
- Matravers, D. (2014). *Fiction and Narrative*. Oxford University Press.
- (2013, March 2017) *Derek Matravers: Those pesky categories of fact and fiction* [Video file]. Retrieved from [https://www.youtube.com/watch?v=dfly\\_weEjR0](https://www.youtube.com/watch?v=dfly_weEjR0)
- Maurin, A.S. (2018). Tropes. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018 ed.). <https://plato.stanford.edu/archives/sum2018/entries/tropes/>
- Meinong, A. (1971). Über Gegenstandstheorie [originally in A. Meinong (ed.), *Untersuchungen zur Gegenstandstheorie und Psychologie*, Barth, Leipzig 1904], in *Gesammelte Abhandlungen, Gesamtausgabe bd. II*, Akademische Druck- und Verlagsanstalt, Graz, 481–535; transl. by I. Levi, D.B. Terrell and R.M. Chisholm, in R. Chisholm (ed.), *Realism and the Background of Phenomenology*, Free Press.
- Miller, G. A. (1999). On knowing a word. *Annual Review of psychology*, 50(1): 1-19
- Miller, J T. M. (2019). A bundle theory of words. *Synthese*, 1-18.
- Muller, C. (2012). *Harry Potter and the Rescue from Realism: A Novel Defense of Anti-Realism about Fictional Objects*. (Electronic Thesis or Dissertation). Retrieved from <https://etd.ohiolink.edu/>

- Onion, R. (2015). What if Historians Started Taking the What if Seriously. *Aeon Magazine*. Retrieved from [https://aeon.co/donate?utm\\_medium=web&utm\\_source=header-support-banner-default](https://aeon.co/donate?utm_medium=web&utm_source=header-support-banner-default)
- Orilia, F., & Swoyer, C. (2016). Properties. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 ed.). <https://plato.stanford.edu/archives/win2016/entries/properties/>
- Parsons, T. (1980). *Nonexistent Objects*, New Haven: Yale University Press.
- Pautz, A. B. (2008). Fictional coreference as a problem for the pretense theory. *Philosophical Studies* 141: 147– 156.
- Pelletier, J. (2003). Vergil and Dido. *Dialectica*, 57(2): 191-203
- Predelli, S. (2002). ‘Holmes’ and Holmes—A Millian Analysis of Names from Fiction. *Dialectica*, 56(3): 261-279.
- Quine, W.V.O. (1948). On What There Is. *Review of Metaphysics*, 2: 21–38. Reprinted in W.V. O. Quine, *From a Logical Point of View*. Harvard University Press.
- Reicher, M. E. (1995) Zur Identität fiktiver Gegenstände: Ein Kommentar zu Amie Thomasson. *Conceptus* XXVIII (1995), 72: 93-116.
- (1999). Amie L. Thomasson, Fiction and Metaphysics. *Grazer Philosophische Studien*, 57 (1): 325–344
- (2006). Two interpretations of “According to a story”. In Andrea Bottani & Richard Davies (eds.), *Modes of Existence: Papers in Ontology and Philosophical Logic*. Ontos Verlag
- Reimer, M. (2001). The problem of empty names. *Australasian Journal of Philosophy* 79 (4):491 – 506.
- Richard, M. (2000). Semantic Pretense. In A. Everett & T. Hofweber, *Empty Names, Fiction and the Puzzles of Non-existence*. CSLI.
- (2013). *Context and the Attitudes: Meaning in Context, Volume 1* (Vol. 1). Oxford University Press.
- Rosen, G. . (1990). Modal fictionalism. *Mind*, 99(395): 327-354.
- Roskies, A. L. (1999). The binding problem. *Neuron*, 24(1): 7-9.
- Russell, B. (1905). On Denoting. *Mind*, 14: 473–493.
- (1985). *The Philosophy of Logical Atomism*, ed. David Pears. LaSalle, IL: Open Court.
- Ryle, G. (1949). *The Concept of Mind*. Hutchinson & Co.
- Sainsbury, R.M. (2009). *Fiction and Fictionalism*. Routledge.
- Salmon, N. (1998). Nonexistence. *Noûs*, 32: 277–319.
- Schiffer, S. (1996). Language-created Language-independent Entities. *Philosophical Topics*, 24(1): 149-167.
- Sawyer, S. (2002). Abstract Artifacts in Pretense. *Philosophical Papers*, 31 (2): 183-198.
- Searle, J. (1979). The Logical Status of Fictional Discourse. In P.A. French *et al.* (eds.), *Contemporary Perspectives in the Philosophy of Language*,. University of Minneapolis Press.

- (2006). Social Ontology: Some Basic Principles. *Anthropological theory*, 6(1): 12-29.
- Singh, S.K. (2015). *Accountancy: Class XII*. SPBD Publications.
- Speaks, J. (2011). Facts, properties, and the nature of the proposition. *Unpublished ms.* <http://www3.nd.edu/~jspeaks/papers/facts-properties-propositions.pdf>.
- Stainton, R. (2006). Meaning and Reference: Some Chomskian Themes. In E. Lepore & B. Smith (Eds.) *The Oxford Handbook of Philosophy of Language*. Oxford University Press
- Terrone, E. (2017). On Fictional Characters as Types. *British Journal of Aesthetics*, 57(2): 161-176.
- Thagard, P. (2019). *Brain-Mind: From Neurons to Consciousness and Creativity (Treatise on Mind and Society)*. Oxford University Press
- Thomasson, A.L. (1999). *Fiction and Metaphysics*, Cambridge University Press.
- (2003a). Fictional Characters and Literary Practices. *British Journal of Aesthetics*, 43: 138–157.
- (2003b). Speaking of Fictional Characters. *Dialectica*, 57: 205–223.
- (2009). Fictional Entities. In J Kim, E Sosa & G Rosenkrantz (Eds.). *A Companion to Metaphysics*. Blackwell.
- (2015). Fictional Discourse and Fictionalisms. In S. Brock and A. Everett (Eds.), *Fictional Objects*. Oxford University Press
- Van Cleve, James, (1985). Three Versions of The Bundle Theory. *Philosophical Studies*, 47: 95–107.
- van Inwagen, P. (1977). Creatures of Fiction. *American Philosophical Quarterly* 14 (4): 299 - 308.
- (2003). Existence, Ontological Commitment, and Fictional Entities. In M. J. Loux & D. W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics*. Oxford University Press.
- (1983) Fiction and Metaphysics. *Philosophy and Literature* 7 (1): 67-77.
- (2018) Replies. In L. Jansen & P. Näger (Eds.) Peter van Inwagen. Münster Lectures in Philosophy, vol 4. Springer
- von Solodkoff, T. (2014). Fictional Realism and Negative Existentials. In Manuel García-Carpintero & Genoveva Martí (eds.), *Empty Representations: Reference and Non-Existence*. Oxford University Press.
- von Solodkoff, T. & Woodward, R. (2017). To Have and to Hold. *Philosophical Issues* 27 (1): 407-427.
- Voltolini, A. (2006). *How Ficta Follow Fiction. A Syncretistic Account of Fictional Entities*, Dordrecht: Springer.
- Walker, C. M., & Gopnik, A. (2013). Causality and imagination. In M. Taylor (Ed.), *Oxford library of psychology. The Oxford handbook of the development of imagination*. Oxford University Press
- Walters, L. (2015). The Problem of Nonexistence: Truthmaking or Semantics? Critical Notice of The Objects of Thought, by Tim Crane. *Disputatio* 7 (41): 231-245.
- Walton, K. (1990). *Mimesis and Make-Believe*. Harvard University Press.

—— (2003). Restricted Quantification, Negative Existentials, and Fiction. *Dialectica*, 57(2): 239-242.

Wolterstorff, Nicholas. Wolterstorff, Nicholas (1976). Worlds of works of art. *Journal of Aesthetics and Art Criticism* 35 (2): 121-132.

—— (1979). Characters and their Names. *Poetics* 8: 101-127

Yablo, Stephen (1998). Does Ontology Rest on a Mistake? *Aristotelian Society Supplementary Volume* 72 (1): 229 - 283.

Yagisawa, T. (2001). Against Creationism in Fiction. *Noûs* 35 (15): 153-172.

Zalta, E.N. (1983). *Abstract Objects: An Introduction to Axiomatic Metaphysics*. D. Reidel.

Zemach, Eddy M. (1998). Tom Sawyer and the Beige Unicorn, *The British Journal of Aesthetics*, 38 (2): 167–179

—— (2003) Amie L. Thomasson, Fiction and Metaphysics. *The Philosophical Review*, 112 (3): 427–431.