

**FICTION WITHOUT PRETENSE**

by

Jillian Alexandra Isenberg

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR  
THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Philosophy)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2013

© Jillian Alexandra Isenberg, 2013

## Abstract

A No-Object theory of fiction denies that there is any sense of “object” in which the objects of fiction are objects at all. This is conjunction of two fundamental assumptions. The first is a metaphysical principle that asserts that there is nothing that does not exist. The second asserts that the individuals and events that figure in works of fiction do not exist. I call these assumptions “Parmenides’ Rule” and the “Non-Existence Postulate”. The No-Object theory also raises what I call the *subject-matter paradox*. If the objects of fiction are nothing, how can it be that we refer to them, ascribe properties to them, and draw inferences about them?

My dissertation dissolves the subject-matter paradox by providing an explanandum for philosophical theories of fiction. A theory of fiction must explain how we can know that there are no objects of fiction, while we respond as though there are. In order to better understand these responses to fiction, I consider recent empirical work in psychology. This work supports the claim that fictional narratives impact our beliefs and attitudes about both the fictional and the actual worlds and shows that we do in fact accept and act as though fictional statements are true, even when we are aware of their falsity. Empirical data concerning our responses to fiction supports a number of claims. First, fictions have objects. Second, we refer to, make true claims about, and draw correct inferences about the objects fiction. The Rule and the Postulate seem to cost us the truths of these two claims; given the Rule and the Postulate are true the claims must be false. If we accept the No-Object view, we shouldn’t feel philosophically obliged to honour our linguistic intuitions. What the data also show, however, is that the very people whose intuitions the No-Object view tramples have other commitments that actually support these intuitions. It is this seeming contradiction that a theory of fiction must accommodate. It must account for the fact that our responses to fiction are double-aspected. I provide a characterization of these double-aspected responses.

## **Preface**

This dissertation is original, unpublished, independent work by the author, Jillian Alexandra Isenberg.

# Table of Contents

Abstract.....	ii
Preface.....	iii
Table of Contents.....	iv
Acknowledgements .....	v
Dedication .....	vi
Introduction .....	1
1 Objects Not Included.....	8
1.1 The No-Object View .....	8
1.2 Ambiguation.....	13
1.3 Sainsbury.....	14
1.3.1 Negative Free Logic .....	16
1.3.2 Proper Names .....	18
1.3.3 Fiction .....	20
2 A Tale of Two Fictions .....	25
2.1 Representation .....	25
2.2 Imagination .....	29
2.4 Bootstrapping .....	37
2.5 Implicit Definition .....	38
2.6 Literary Fiction .....	40
2.7 Adaptation vs. Instantiation .....	44
3 We Can Believe It for You Wholesale (Or, Total Acceptance).....	49
3.1 So Close and Yet So Far?.....	52
3.2 Fiction by the Numbers .....	54
3.2.1 You Can't Not Believe Everything You Read .....	54
3.2.2 Fictional Narratives Change Beliefs.....	58
3.2.3 Replications with Mixed Corroboration .....	61
3.2.3.1 Initial Adaptation .....	61
3.2.3.2 Need for Cognition and Personal Relevance .....	63
3.2.3.3 Perceived Strength of Story Assertions and Levels of Scrutiny .....	65
3.3 Philosophical Implications.....	67
4 The Curious Case of Double-Aspected Responses.....	71
Afterword .....	87
Bibliography .....	89

## Acknowledgements

I would like to thank my committee for their support on this project. Individual thanks goes to my supervisor Dominic Lopes for all of the time, encouragement and helpful discussions that he so generously provided in order to help me accomplish my goals. Thanks to John Woods for your early guidance and encouragement as well as your ongoing help and support. Thanks to Roberta Ballarin for always providing a new perspective and helping me to see things in ways that I had not yet considered.

I would also like to thank my fellow graduate students in the Philosophy department for all of the productive and supportive conversations that we've had over the years. Special thanks goes to Nola Semczyszyn and Joshua Johnston who have been the two best travelling companions I could have asked for on this journey.

I would also like to give special thanks to Nissa Bell and Rhonda Janzen for always being there to listen to me as well as for immense amounts of help and support.

I would especially like to thank Elisabeth Ratz for her eleventh hour effort to help ensure that everything was formatted correctly and submitted on time. More importantly, thank you for your blind faith in my abilities and your amazing friendship.

Finally, I would like to thank my family for the moral, emotional, and financial support without which this project would not have been possible. You have been incredibly generous, patient and understanding. Words cannot express my gratitude.

This thesis is dedicated to my parents.

## Introduction

Philosophical attention to fiction has gathered momentum since the early 1970s, when it became a serious topic of interest for logicians and philosophers of language, and for philosophers of art. Philosophers of language take a semantic approach, concentrating on theories of reference, truth and inference. Philosophers of art concentrate on representational and normative questions. Recently, philosophers of mind have also done work that draws upon and connects with psychology and cognitive science in an attempt to explain our responses to fiction and aesthetics now draws upon this work in an attempt to give an account of our engagement with literary fiction. When, in the pages to follow, I speak of the modern literature, the reference I intend is to the body of work produced by these philosophers in that forty-year period.

The modern literature subdivides in telling ways under a pair of fundamental assumptions. The first is a metaphysical principle that asserts that, owing to the existentially loaded nature of quantification, there is nothing that does not exist. While this is a general principle about everything, the second assumption is a particular claim about fiction. It asserts that the individuals and events that figure in works of fiction do not exist. For ease of reference, I will call these assumptions “Parmenides’ Rule” and the “Non-Existence Postulate” respectively.<sup>1</sup> Together they induce a partial partition of the modern literature.

Subscription to, and rejection of, Parmenides’ Rule and the Non-Existence Postulate carve out some of the more important rivalries in contemporary theories of fiction. For example, in holding that there are some objects that do not exist and that the objects of fiction are among them, Meinong rejects Parmenides Rule while accepting the Non-Existence Postulate.<sup>2</sup> That is, he thinks that since there are things that don’t exist, these non-existents are in the range of the quantifier “there are.” This view admits of some interesting refinements and extensions, producing theories that fall with varying strictness into the rubric of a generic Meinongeanism, and it affords an important insight about our two assumptions. It is that they are independent of each other. The Rule does not imply the Postulate nor does the Postulate imply the rule.

---

<sup>1</sup> John Woods and Jillian Isenberg, “Psychologizing the Semantics of Fiction,” *Methodos* 10 (2010), accessed April 7, 2010, doi: 10.4000/methodos.2387.

<sup>2</sup> Alexius Meinong, “The Theory of Objects,” in *Realism and the Background of Phenomenology* ed. R. Chisholm (Atascadero: Ridgeview Publishing Co., 1960), 76-117.

Not every way of rejecting the Postulate is compatible with every way of fulfilling the Rule. For example, when he asserts the existence of fictional beings, David Lewis rejects the Postulate.<sup>3</sup> But when he asserts that the beings of fiction lack residency in the actual world, he rejects any version of the Rule that restricts the domain of quantification to actualia. By the same token, Lewis' existent non-actualia satisfy versions of the Rule permitting quantification over possible, non-actual existents. However, one can only wonder what now remains of the original intent of the Postulate. In particular, did it envisage non-actual possibility as a way of qualifying for existence? Or is it rather more sensible to see Lewis as affirming the Postulate (“nothing actual is fictional”) and rejecting the Rule (“there are things that are not actual”)? On more fronts than one, Lewis is an interesting case.

Peter Van Inwagen also advances the idea that the creatures of fiction actually exist and are among the bound variables of quantification but here too with a wrinkle.<sup>4</sup> The wrinkle with Lewis was the existence of non-actual possibilia. The wrinkle introduced here is the theoretical entity. Seen in Van Inwagen's way, Anna Karenina exists and electrons exist. They are both objects of theory – of physics in the case of electrons and of literary theory in the case of Anna – and they are both open to quantification. So the Rule and the Postulate are both satisfied. What makes this a wrinkle is that there is no intuitive support for the idea that Anna is a *creature* of Amy Mandelker.<sup>5</sup>

That said, the Rule and the Postulate are the consensus in the philosophies of language, mind and art. Indeed the motivation for much of the work done on fiction in philosophy is the result of a general acceptance of them within these fields.<sup>6</sup> Moreover, among those who are concerned with providing a philosophical account of fiction, there is a further divide within the literature between those who are interested in fiction because of this acceptance. Some are interested in metaphysical fictionalism, using theories of fiction to further explain how we individuate, talk about and have knowledge of theoretical entities that do not actually exist. These include modal, mathematical and moral fictionalists.<sup>7</sup> Others are more concerned with art

---

<sup>3</sup> David Lewis, “Truth in Fiction,” *American Philosophical Quarterly*, 15 (1) (1978): 37-46.

<sup>4</sup> Peter Van Inwagen, “Creatures of Fiction,” *American Philosophical Quarterly*, 14 (1977): 299-308; Van Inwagen, “Quantification and Fictional Discourse,” in *Empty Names, Fiction, and the Puzzles of Non-Existence* edited by Thomas Hofweber and Anthony Everett (Stanford: CSLI Publications, 2000).

<sup>5</sup> Amy Mandelker, *Framing Anna Karenina: Tolstoy, the Woman Question, and Novel* (Columbus: Ohio State University Press 1993).

<sup>6</sup> W.V.O. Quine, “On What There Is,” in *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1980), 1-19.

<sup>7</sup> Among them, Mark Balaguer, *Platonism and Anti-Platonism in Mathematics* (Oxford: Oxford University Press, 1998);



or literary fictions. In some cases they are interested in this topic because of the insights that this special kind of discourse can provide into concerns with vacuous names and non-referring expressions in the philosophy of language. In others, they are interested in our responses to fiction, as in both philosophy of mind and aesthetics. While it may be the case that philosophical results from inquiry into our discourse about and psychological engagement with literary fictions will prove interesting and helpful to metaphysical fictionalists, my concern here is with providing an account of literary fictions that is not concerned with its implications for such theories.

As important as they are, these views are by no means exhaustive. But they are representative of some basic approaches that philosophers take to fiction.

As regards the Rule and the Postulate, I am a No-Object theorist; I accept them both. That is, I accept the claim that there are no non-existent objects and the claim that fictional entities do not exist. Thus, I accept that fictional entities are not objects. I will be the first to concede the radicalness of the No-Object theory. It is greatly at odds with the ways – the considered ways – in which we all speak and think. If Anna Karenina is an object of fiction and yet there is no object that is *she*, how in the very saying of it can we avoid quantification over her? This is, in part at least, the problem of negative existentials.<sup>8</sup> But it also raises what I will call the *subject-matter paradox*. How, if nothing is Anna, can Tolstoy's novel and our theories of it have a subject-matter? If the objects of fiction are nothing, how can it be that we refer to them, ascribe properties to them, and draw inferences about them?

This apparent inconsistency, between the fact that we know that the objects of fiction do not exist and the fact that we respond to fictions as though they do, is one that must be accounted for by any adequate comprehensive philosophical theory of fiction. In what follows, I attempt to account for it by developing a general characterization of our responses to fiction which accommodates both horns of the apparent dilemma.

One instance of the *subject-matter paradox* is what is generally referred to as the *paradox of fiction*.<sup>9</sup> Simply put the paradox asks how it can be the case that we are genuinely moved by

---

Hartry Field, *Science Without Numbers* (Princeton: Princeton University Press, 1980); Gideon Rosen, "Modal Fictionalism," *Mind* 99 (1990): 327–54; and Mark Kalderon, *Moral Fictionalism* (Oxford: Clarendon Press, 2005).

<sup>8</sup> R. Cartwright, "Negative Existentials," *Journal of Philosophy* 57 (1960): 629–639; F. Kroon, "Disavowal Through Commitment? Theories of Negative Existentials," In *Empty Names, Fiction, and the Puzzles of Non-Existence* edited by Everett & Hofweber, (Stanford: CSLI, 2000).

<sup>9</sup> C. Radford, "How Can We Be Moved by the Fate of Anna Karenina?" *Proceedings of the Aristotelian Society* 49 (1975): 67–80; and R.J. Yanal, *Paradoxes of Emotion and Fiction*, (Philadelphia: Pennsylvania State University Press, 1999).

fictions if we know that what is portrayed in them is not actual. This is problematic because it seems that we can only be moved by what we believe is actual. This results in an inconsistent triad and so, while each claim is itself plausible, for any two of the claims to be true the third must be false. If we know that the objects of fiction do not exist, how can we have genuine emotional responses to them?

Does the No-Object view commit us to a blanket nihilism not only about Anna and her like but also about the responses that we have to stories in which they figure? It might reasonably be supposed that a theory of such nihilistic clout can only be disqualified by it. If there is no sense of “object” in which the objects of fiction are objects, no sense of “true” in which sentences about her are true, no sense of “instantiate” in which there are properties which Anna instantiates, and no sense of “infer” in which true consequences about Anna are inferable from true premises about her, then what could possibly justify the employment of Parmenides’ Rule and the Non-Existence Postulate in a theory of fiction? It will be a burden of the chapters to follow to answer this question. In discharging it, I shall propose that it is a peculiarity of fiction that the best account of our responses to it is one that allows these assumptions to stand.

There is a further assumption that also divides the modern literature in an important way. I call it Frege’s Dismissal.<sup>10</sup> It states that since fiction is of no scientific importance, it deserves no semantic notice beyond what is expressly fashioned for matters of science. If we accept this assumption, a semantic account of fiction can only be an extension of a more general theory, and therefore something of an afterthought. Russell’s “On Denoting” is a classic illustration of this principle.<sup>11</sup> It requires that a theory of reference for “Anna Karenina” be the same as the one for “Vulcan.” For philosophers who favour the course of reductionism and unification, Frege’s Dismissal is an attractive principle. And indeed it may be the case that a semantic account of fiction must not only respect but embrace this dismissal, as a theory of fiction must be just as concerned with our psychological responses to fiction as it is with either determining or developing the theory of reference, truth and inference that can best accommodate it. However, I intend to reject the spirit of Frege’s Dismissal, the view that fiction, by itself, deserves no serious philosophical attention, and to press for a purpose-built theory of fiction. I am not an anti-reductionist as a matter of philosophical principle. But in the present case, the peculiarity of fiction calls for specialized treatment.

---

<sup>10</sup> John Woods and Jillian Isenberg, “Psychologizing the Semantics of Fiction.”

<sup>11</sup> Bertrand Russell, “On Denoting,” *Mind* 14 no. 56 (1905): 479-493.

The first chapter of this thesis discusses the work of Mark Sainsbury.<sup>12</sup> Sainsbury is a semantic theorist of fiction concerned with giving a general account of how vacuous names can be referential expressions. He takes fictional names to be a particular species of vacuous ones. Thus he adopts a strict non-objectualist approach to fiction, which he calls the “irrealist” view. He claims that there is no sense in which the objects of fiction exist and that sentences in which such objects figure are typically paraphrases of sentences about the works themselves. So, Sainsbury accepts both the Rule and the Postulate.

The next chapter discusses the work of Kendall Walton.<sup>13</sup> Walton’s account of fictionality accepts both the Rule and the Postulate. He denies the existence of fictional objects as well as the idea that there is any sort of metaphysical account on which we can say that there “are” fictional objects. Walton’s is a No-Object approach to fiction. He interprets the Rule and the Postulate in ways that make it hard to say that no object is an object of fiction, that there is no sense of “object” according to which Anna Karenina is an object. True, Anna is an object of fiction but being so is not a way of being an object. I take this as encouraging, as one of the virtues of Walton’s account is the skill with which he develops his theory in the face of the difficulties posed by these assumptions. Even so, there are problems with the account. The theory that Walton erects under the guidance of the Rule and the Postulate is not the best theory that they admit of.

Accordingly, mine is a twofold task. I must expose the difficulties that inhere in Walton’s theory. I must show that these problems are not necessitated by his adoption of Parmenides’ Rule and the Non-Existence Postulate and his rejection of Frege’s Dismissal. In order to do this, I will consider recent work done in experimental psychology. Looking at the experimental data will provide necessary insight into the nature of our responses to fiction.

Some of those working in philosophy of mind and aesthetics have explored this avenue.<sup>14</sup> Specifically, they have considered the implications of simulation theory and theory

---

<sup>12</sup> R.M. Sainsbury, *Reference Without Referents* (New York: Oxford University Press, 2005); and R.M. Sainsbury, *Fiction and Fictionalism* (London: Routledge, 2009).

<sup>13</sup> Kendall Walton, *Mimesis as Make-Believe: On the Foundations of the Representational Arts* (Cambridge MA: Harvard University Press, 1990).

<sup>14</sup> S. Nichols, and S. Stich, *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding of Other Minds* (Oxford: Oxford University Press, 2003); and Shaun Nichols, ed. *The Architecture of the Imagination: New Essays on Pretence, Possibility, and Fiction* (Clarendon Press, 2006); Alvin Goldman, “Imagination and Simulation in Audience Responses to Fiction,” in *The Architecture of Imagination* (2006): 41-56; and Alvin I. Goldman, *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (Oxford: Oxford University Press, 2006).

theory to better understand the nature of imagination, which is taken to be the mental state that we are in when we engage with fiction. Both of these theories attempt to explain how we attribute thoughts to others based on their behaviour – that is, mindreading. On this approach, in order to understand imagination we must understand the mechanism that underlies it and this mechanism is the same as what underlies mindreading, since like mindreading imagination requires that we take up the perspectives of others. However, because this literature treats imagination as a kind of mindreading, it does not provide direct evidence for why and how we engage with fiction as we do.

Instead of taking this approach, the literature that I engage with seeks to understand our responses to fiction more directly.<sup>15</sup> It seeks to determine whether and how reading fictional narratives influences our beliefs. A better understanding of the effect of fiction on our beliefs will provide clearer guidance with respect to the many issues of interest to philosophers working on fiction.

The *subject-matter paradox* in general and the *paradox of fiction* in particular both arise because of the peculiar nature of our responses to fiction. They raise the question of how we can respond to what we do not believe, whether those responses are intentional, doxastic or affective. However, the empirical data supports the claim that fictional narratives impact our beliefs and attitudes about both the fictional and the actual worlds and shows that we do in fact accept and act as though fictional statements are true, even when we are aware of their falsity. It further suggests that this acceptance is an automatic pre-reflective response, which induces real responses. As a result, I deny the *subject-matter paradox* and propose a new way of understanding how we can genuinely have such responses to fictions while knowing that the events they portray did not actually occur. I call this the *double-aspected* nature of our responses to fiction.

In order to accomplish this, I propose a general characterization of the explanandum for such a theory; a schema that can be filled out in order to characterize a wide range of responses that we have to fictions. This explanandum must be articulated and understood by philosophers attempting to develop a comprehensive theory of fiction as well as those who are working to

---

<sup>15</sup> Melanie C. Green, Jennifer Garst, Timothy C. Brock, and L.J. Shrum. “The Power of Fiction: Determinants and Boundaries,” in *The Psychology of Entertainment Media: Blurring the Lines Between Entertainment and Persuasion*, ed. Karen E. Dill (Mahwah, NJ: Lawrence Erlbaum, 2004) 161-176; Richard J. Gerrig, *Experiencing Narrative Worlds: On the Psychological Activities of Reading*. (New Haven, CT: Yale University Press, 1993); and D.T. Gilbert, R.W. Tafarodi, & P.S. Malone, “You Can’t Not Believe Everything You Read,” *Journal of Personality and Social Psychology* 65 (1993): 221–233.

effectively address issues that arise as the result of the peculiarities of fiction within various diverse areas of philosophy. In either case, the characterization of such an explanandum should provide a common framework within which to solve their respective problems. Taking the No-Object view as the ontological starting point from which to develop this characterization, I propose that the correct explanandum for any theory of fiction is the double-aspected nature of our responses to fictional narratives.

# 1 Objects Not Included

## 1.1 The No-Object View

In the introduction to this thesis I announced my intention to work toward a theory of fiction that respects Parmenides' Rule and the Non-Existence Postulate. Recall that the Rule asserts that there is nothing that does not exist, and the Postulate that the individuals and events of fiction do not exist. The Rule is a general metaphysical principle and the Postulate a particular one. The joint acceptance of these assumptions implies a no-object view of fiction, which in its most basic sense denies that there is any sense of "object" in which the objects of fiction are objects. On all the standard readings, a No-Object theory with respect to a purported class of objects provides that there is no reference to its purported members and, accordingly, that sentences purportedly about them aren't true.<sup>16</sup>

A No-Object theory of fiction must account for a good deal of behavioural and linguistic data that make it deeply counterintuitive. One striking thing about the Sherlock Holmes stories is that we seem to have absolutely no difficulty understanding them. We know quite a bit about Holmes, Watson, their many adventures and the London in which they take place. We can distinguish Holmes and Watson from one another as well as from real people with ease. This takes no training or technical savvy; a small child can tell Watson from her own paediatrician. We know who we are talking about when we talk about Holmes, Watson and likewise Mrs. Hudson. Since we know these things about them, there are many things that are true of them. "Holmes lived at 221B Baker Street" is true and "Holmes lived in New York" is false. "Holmes patronized Lestrade" is true and "Holmes patronized his brother Mycroft" is false – indeed it is laughable. This is all in virtue of the fact that Doyle tells us so in the particular way authors of stories have of so doing.

The empirical data reflected in our linguistic behaviour supports a number of claims:

1. Fictions have objects. Fictions are about things. They are about fictional characters.
2. We refer to, individuate, have knowledge of, make true claims about, and draw correct inferences about the individuals and events that figure in fiction.
3. Fictions come about because of the author's creative endeavour, and it is ultimately to this that propositions (1) and (2) owe their truth.

---

<sup>16</sup> Assuming the standard objectual interpretation of the quantifiers.

Doubtless, some readers will judge that the Rule and the Postulate place excessively harsh restrictions on a theory of fiction – that a theory developed under these constraints will be severely undermined by them. After all, (1) and (2) seem to contradict the Rule and the Postulate outright. Why develop a theory of fiction that takes them as their starting point? The Rule and the Postulate cost us the truths of (1) and (2). Is this not too heavy a cost to bear? Would it not be perverse to pay it?

It would be quite wrong to leave the impression that the Rule and the Postulate lack for takers, that they are declarations of such extremity as to attract only slight and intermittent support. Of course, the opposite is true. The Rule and the Postulate are hegemonic. They have had a dominant place in the philosophy of language from Frege and Russell to the present. The Rule and the Postulate are mainstream, and the theorists who have embraced them over the years have not scrupled to pooh-pooh the likes of (1) and (2) in the manner, for example, of Frege's Dismissal.

I find myself in an interesting position. Concerning the Rule and the Postulate, I am at one with Frege and Russell, with Strawson, Donnellan, Searle and the others.<sup>17</sup> Concerning (1) and (2), I am at one with the objectualists, with Meinong, Parsons, Jaquette, Thomasson and Van Inwagen.<sup>18</sup> At least, I am one-half with them. I am with them in the sense of thinking that any suitably general philosophy of natural language that fails to provide for (1) and (2) in some principled form of rescue is a less than adequate theory. So I am indeed oddly positioned. Given the Rule and the Postulate are true (1) and (2) are false, but falsity be hanged, I have now pledged myself to their rescue. What could this possibly mean?

In "Sense and Reference," Frege claims that declarative sentences express thoughts. When we consider the meaning of a sentence, we ask for both the thought it expresses and the truth-value of that thought.<sup>19</sup> He believes that the reason for this is that it is our quest for the truth that compels us to determine not just the cognitive significance of a sentence, but also whether or not it expresses a truth. However, Frege explains that we do not always have to ask for the truth-value of a sentence. One case in which we need not attempt to determine the

---

<sup>17</sup> Peter Frederick Strawson, "On Referring," *Mind* 59 (235) (1950): 320-344; Keith S. Donnellan, "Speaking of Nothing," in *Contemporary Research in Philosophical Logic and Linguistic Semantics* (Springer Netherlands, 1975) 93-118; and John R. Searle, "The Logical Status of Fictional Discourse," *New Literary History* 6 (2) (1975): 319-332.

<sup>18</sup> Meinong, "The Theory of Objects," 76-117; Terence Parsons, *Nonexistent Objects* (New Haven: Yale University Press, 1980); Dale Jaquette, "Intentional Semantics and the Logic of Fiction," *The British Journal of Aesthetics* 29 (1989): 168-176; Amie L. Thomasson, *Fiction and Metaphysics* (Cambridge University Press, 1999); and Peter van Inwagen, "Creatures of Fiction," *American Philosophical Quarterly* 14 no. 4 (1977): 299-308.

reference of a sentence is when we are considering sentences that contain non-referring, or vacuous, names; the class of singular terms that includes fictional names. The reason that Frege thinks we need not attempt this is that he does not consider the question of the referents of fictional names to be a question that theories of reference need address. It is not a serious question because fictional discourse is not itself serious, hence Frege's Dismissal.

Frege considers the example of the name "Odysseus" and explains that, since it is likely that this name does not have a reference, the sentence "Odysseus was set ashore at Ithaca while sound asleep" is also unlikely to have a reference. In this way, Frege incorporates his treatment of fictional names within his general theory of reference. In such fictional cases, he explains, the purpose of the sentence is not to convey information, and so we need not be concerned by the lack of reference. Rather, the purpose of the sentence is to express a particular idea in a certain light, and so we are only concerned with the thought that it expresses. Further, in the case of fiction and poetry, Frege explains that all we are interested in beyond the aesthetic value of the work are the thoughts and emotions that it evokes. He claims that questions of truth would serve to detract from our enjoyment of that experience, because they serve to deconstruct the artifice. Hence, the sentences in which such names figure have no truth-value.

Russell respects Frege's Dismissal, but rejects Frege's theory of reference because he does not think that it provides an adequate solution to the problems that it faces, chief among them, that of vacuous names and non-referring expressions. He rejects the claim that sentences containing non-referring expressions have no truth-values in favour of the view that they are false.<sup>20</sup> He uses the example "The present King of France is bald" to demonstrate this. This sentence expresses the proposition "There is exactly one individual such that it is the present King of France, and that individual is bald." If we take the sentence to express that proposition, we understand that it must be false, since France presently has no king. This is because the sentence asserts a conjunction, of which one of the conjuncts is false. This sentence must be false, since it falsely asserts the existence of the individual denoted by the "the present King of France" and if that individual does not exist then he cannot be bald.

Russell's view of names provides a neat solution to the problem of vacuous names and non-referring expressions. The solution is to acknowledge the simple fact that, when we use a

---

<sup>19</sup> Gottlob Frege, "Sense and Reference," in *The Philosophical Review* 57 no. 3 (1948): 209-230.

<sup>20</sup> Russell, "On Denoting," 479-93.



name, unless we are directly acquainted with the individual in question, we do not know for sure that the name has a real-world extension. Neither do we know whether the variable of the existential quantifier has a value. So, when determining the reference of a name, when we say that there is at least one individual that fits the description, we must add the caveat that this is only the case if it exists. So, in the example “Sherlock Holmes lived at 221B Baker Street” we understand the sentence as expressing the following proposition: “there is at least one detective who is a brilliant deducer, and he lived at 221B Baker Street.” The name “Sherlock Holmes” is meaningful because it is associated with the description “detective who is a brilliant deducer.” Even though no one is directly acquainted with such an individual, a regular speaker of our language can grasp the meaning of the name. Moreover, because of our acquaintance with the constituents of such a description, we can grasp the proposition that is being asserted and ascertain its truth-value. When we do this, we determine that the proposition “there is at least detective who is a brilliant deducer” is false. Since Sherlock Holmes does not exist, there is no such individual that fulfills this description, and so the entire proposition “Sherlock Holmes lived at 221B Baker Street” is false. So, Russell respects Frege’s Dismissal while accepting a no-object view of fiction. In fact, he argues directly against Meinong’s non-existent entity view in “On Denoting.”<sup>21</sup>

The No-Object view becomes yet more puzzling if we reject Frege’s Dismissal as it denies a univocal theory of reference to account for our discourse about fiction. Frege and Russell would not deny that their subscription to the tough principles that constitute a No-Object view do violence to our intuitions about Holmes. However, they think that this can be accommodated by the fact that fiction does not matter. Fictional discourse is not serious discourse, after all. Moreover, neither Frege nor Russell think that natural language is reliable, so they are warranted in dismissing all kinds of facts about our behaviour and language use including those that motivate a theory of fiction. This makes the Rule and Postulate considerably less demanding since their theories both satisfy them, but only at the price of dismissing fictional discourse. Both Frege’s and Russell’s views are motivated by the belief that the underlying language of philosophy is classical first order logic. So it is hardly surprising that any treatment of fiction in the long tradition ensuing from their work will be constrained by them. But this violence is contraindicated by the competent linguistic behaviour of us all. This makes the question of whether to accept or reject Frege’s Dismissal a terribly important one. If we accept

---

<sup>21</sup> Russell, “On Denoting,” 491.

it, we should not feel philosophically obliged to honour the facts about our behaviour and language use, but if we take Frege's Dismissal seriously enough to reject it, we have to deal with some serious problems. It is not just that the Rule and the Postulate ride roughshod over our linguistic intuitions. There is also a rich body of linguistic intuitions embedded in our competent discursive behaviour that the No-Object view overrides. What our language use and behaviour in response to fiction also show, however, is that the very people whose intuitions it tramples have other commitments that make these intuitions true. Do we not deny Holmes' existence in the same conversation in which we discuss his address? So, we must pause a while. We must consider the possibility that a No-Object theory of fiction rejecting Frege's Dismissal is so deeply counterintuitive as to be a non-starter.

We ourselves – the average, competent, intelligent speakers of English – have a lot to say about Holmes. We have about as much to say about him as we do about Palin. We say “Holmes was a detective” and “Palin is a politician.” In doing so, we take ourselves to be stating facts about both, to be saying things that we know to be true. In each case we speak in ways that carry the clear suggestion that there is something referred to, individuated and talked about. When talking about Sarah Palin there is no additional question – beyond the usual questions that we must ask about the conditions on successful reference and assertion – that we must ask about how we manage to do this. That we manage to talk about Palin (or Paris or Peirce or Plato) is not at all surprising, except to philosophers of language. Even they would agree that talking about Palin is not surprisingly surprising. But it is not this way with Holmes. Talking about Holmes is deeply puzzling. The name “Sarah Palin” refers because, in substantial part, there is a particular individual in the world who bears it. Equally, what is true or false of her is a matter of how the world is. We draw inferences about her based on our knowledge of her together with other more general facts about the world. Holmes is different. There is nothing in the world named by “Sherlock Holmes,” hence nothing contributing to the fact that he lived at 221B Baker Street. For if nothing in the world is named by “Sherlock Holmes,” how can it be the case that he lived there? There is nothing in the world that counts as grounding inferences about him. What we can appeal to, however, is the fact that Doyle tells us all sorts of things about Holmes. An account of our linguistic behaviour will rely at least in part on Doyle's generating a certain class of sentences.

## 1.2 Ambiguation

Ambiguation is the standard way out of inconsistency and it looks as if it is available to us here. That is, we can highlight and embrace the ambiguities that lead to these apparent inconsistencies. There is a sense in which one set of our linguistic intuitions is true and another sense in which the contravening set of intuitions is also true. Our actual behaviour is largely at odds with our theoretical commitments, which are themselves tracked by our linguistic behaviour. It is the job of any philosopher developing a theory of fiction to identify those senses and verify their presence in this apparent contradiction. We are further constrained by an obligation to account for that ambiguity.<sup>22</sup>

One way to preserve the ambiguity is to locate it in the truth predicate; to say that we are equivocating on 'true'. If this is the case then there are two truth predicates: real truth and truth-in-fiction. "Sherlock Holmes lived at 221B Baker Street" is not really true but it is true-in-fiction.

Another way to preserve the ambiguity is to claim that our discourse about fiction contains a systematic ambiguity. This is an ambiguity between, say, "Sherlock Holmes lived at 221B Baker Street" and " $f$ (Sherlock Holmes lived at 221B Baker Street)".<sup>23</sup> If this is the case then the ambiguity arises because sentences about fiction take the form  $f(\phi)$ , where " $f$ " is a sentence operator indicating fictionality. The ambiguity arises because, in general,  $\phi$  is false while  $f(\phi)$  is true.  $\phi$  is in the scope of  $f(\phi)$ , so we explain away any inconsistency by saying that our intuition denying the truth of "Sherlock Holmes lived at 221B Baker Street" arises from a recognition of the falsity of the unprefix sentence and our intuition acknowledging its truth is really a recognition of the truth of " $f$ (Sherlock Holmes lived at 221B Baker Street)." If this is the case, then a semantic theory of fiction will have two ends - first, to tell us the meaning of  $\phi$  and, second, to tell us the meaning of  $f(\phi)$ . When we read a story, what we understand is the meaning of  $\phi$ . In effect we ignore the  $f$ -operator and read it with the same understanding that we would sentences about Palin. When we discuss a story, what we understand is the meaning of the prefixed sentences, tacitly acknowledging the  $f$ -operator's presence. Thus, it becomes necessary

---

<sup>22</sup> Of course, subscribing to Frege's Dismissal solves this problem as it allows us to dismiss the vast majority of our discourse about Holmes and his like, thereby removing the contradiction.

<sup>23</sup> Lewis, "Truth in Fiction," 37-46; and Gregory Currie, "Fictional Names," *Australasian Journal of Philosophy* 66 no. 4 (1988): 471-488.

for a semantic theory of fiction to do two things. First, it must produce a semantics for  $\phi$ . Second, it must produce a semantics for  $f(\phi)$ .

I am well aware of the difficulties inherent in a No-Object account of fiction. I cut myself off from some plausible avenues of explanation of linguistic practices. Views which reject the Rule or the Postulate may strike us as offering better explanations of the object-involving nature of literary fictions and our discourse about them. If you reject the Rule, you can say that there are things that do not exist. This would explain why we have no problem talking about Holmes despite his non-existence. In fact, we have no problem talking about his non-existence. He is an object, just not an existent one. If you reject the Postulate, you can say that the objects of fiction exist. This leaves open the possibility that the objects of fiction are a special kind of object – existent but abstract, or existent but theoretical, or who knows what. With these options available, why would we adopt so austere an ontology for fiction? One possible answer is that the constraints are ontologically conservative. Developing a plausible theory of fiction without helping oneself to these entities would in itself be a significant achievement. But this is hardly my motivation. The main consideration driving my commitments is equally straightforward: accepting the Rule and the Postulate and rejecting the spirit of Frege's Dismissal facilitates an optimal theory of fiction, a better theory of fiction than is attainable with these constraints removed. They are a positive good for our best theory. We have independent philosophical reasons for being minimalistic in our ontological commitments.<sup>24</sup> Additional ontological baggage is not justified unless it is absolutely necessary for our theories. Ontological minimalism is an important theme in contemporary philosophy, thanks notably to the efforts of Quine, but it is not my focus here.

### 1.3 Sainsbury

Mark Sainsbury also develops a No-Object theory of fiction. He does not believe that it is necessary to posit fictional entities in order to account for our robust linguistic practices with respect to fiction. How then does Sainsbury account for these basic data about fiction, while remaining a strict non-objectualist? Sainsbury believes that we should treat "Holmes" in the same way that we treat "Palin." This is because part of Sainsbury's project is to develop a general semantic theory, one that accounts for how we understand names and sentences that do not

---

<sup>24</sup> For example, those put forward by Quine. W.V.O. Quine, "On What There Is," in *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1980), 1-19.

refer to anything. He thinks that this will explain how it is that we can have understanding in the face of reference failure and therefore predicate attribution failure. Moreover, he believes that the theory that tells the story about fictional names will be the same as the one that tells the story about names that refer to actual individuals. In developing such a theory, he classes fictional names with other non-referring expressions, claiming that there is no semantic difference between them. For example, “Sherlock Holmes” and “Vulcan” receive the same treatment from Sainsbury. Sainsbury is also committed to developing a theory of fiction which is ontologically minimalistic. He believes that one of the conditions on developing a good semantic theory is that it be univocal across all referring expressions.

On the face of it, this is quite problematic. Unlike Sherlock Holmes, we have nothing to say about Vulcan. There is nothing to know about Vulcan. There are no truths about Vulcan. Treating Sherlock Holmes in this same fashion is counterintuitive. We have no problem individuating Holmes and being secure in our knowledge of many truths about him, but we have no way of individuating Vulcan and we know nothing about it. This shows a very serious disparity between our actual use of and beliefs about the referents of two terms which receive the same treatment on Sainsbury’s view. Moreover, it shows that our actual behaviour is at odds with the theory that accounts for them. The theory treats them uniformly; our behaviour does not.

Why then concern ourselves with what Sainsbury has to say about fiction? Certainly, Sainsbury’s work on fiction does not dominate the literature. However, his work in logic and the philosophy of language is well known and well regarded. His recent work towards a theory of fiction is informed by a long career of historical scholarship and cutting edge research on a wide variety of semantic and logical problems.

More importantly for our purposes, Sainsbury’s work towards a univocal semantic theory is instructive for those who would take the ambiguity route out of the difficulties that arise from accepting the Rule and Postulate while also affirming (1) and (2). Sainsbury’s semantics provides an approach that accounts for our fundamental assumptions, despite being problematic with respect to our language use and behaviour. Sainsbury expressly acknowledges the ambiguity thesis reflected in the distinction between  $\phi$  and  $f(\phi)$ , and I believe that he is headed in the right direction about the semantics of  $\phi$ , while he says very little about the semantics of  $f(\phi)$ .<sup>25</sup>

---

<sup>25</sup> Sainsbury develops his account of the scope of  $f(\phi)$  in *Reference Without Referents*. In this book, he discusses the

Moreover, we share the same metaphysical commitments. So, a good theory of fiction may well preserve some aspects of his theory while eliminating others.

Sainsbury has a positive semantics which is wholly general and not restricted to non-referring expressions. Indeed, it does not distinguish between those with and without referents. Rather, his approach is to generate a general theory of reference which will account for the whole class of singular terms in English. This includes “Sarah Palin,” “Sherlock Holmes,” “Vulcan” and “the number two.” What the theory will do, in part, is explain how it is we understand sentences in which these terms occur in subject position. In order to develop such a general and univocal theory that can deal with vacuous names, Sainsbury adopts a free logic.

Why worry about the correct semantics for  $\phi$ ? Within the scope of  $f(\phi)$ , it seems as though our understanding of fiction isn’t much affected. So, it is important to motivate our discussion of Sainsbury’s account. The reason is that despite needing a separate semantics for  $\phi$ , we still need to account for the truth of sentences like “Sherlock Holmes doesn’t exist” and explain how they are generated, initiated and transmitted. So, even though within  $f$  “Sherlock Holmes” has a referent, there will be cases where the name does not figure in prefixed discourse, but is still intelligible. It is these cases that Sainsbury can account for.

### 1.3.1 Negative Free Logic

Coined by Lambert, “free logic” is the abbreviated term for “logic free of existence assumptions with respect to its terms, singular and general.”<sup>26</sup> It allows for the use of extensionless names and predicates. This is a radical departure from classical logic which requires singular terms to have values assigned. Classical logic is particularly problematic for discourse about the non-existent as it requires that we, at least tacitly, affirm the existence of just those things the existence of which we would deny. Thus, to say anything about, say, Sherlock Holmes in a classical language requires that we quantify over Holmes. “Sherlock Holmes is a detective”, for example, is expressed in classical logic as “ $\exists x[(x=Holmes)\&Detective(x)]$ .” So, we cannot talk about Holmes without denying that which we know to be true – we must deny that he does not exist. Indeed, the proposition that there is no such thing as Holmes, or more accurately

---

semantics of referring expressions at length while only briefly touching on how we are to understand them in the context of fiction. While he does take up this issue in *Fiction and Fictionalism*, the project there is metaphysical. As such, any semantic account of the fictionality operator remains underdeveloped.

<sup>26</sup> Karel Lambert, “The Definition of E! in Free Logic,” in *Abstracts: The International Congress for Logic, Methodology and Philosophy of Science* (Stanford: Stanford University Press, 1960).

“Sherlock Holmes does not exist,” is expressed as  $\sim\exists x(x=\text{Holmes})$ . However, in order for this to be meaningful in classical logic there must be an individual or object named by the singular term “Holmes” in the domain of quantification. Therefore, such a statement is false. Not only is the sentence denying Holmes’ existence – already deemed invalid due to its inconsistency – a problem for classical logic, but the seemingly unproblematic – albeit false – claim that Holmes is a detective is likewise invalid as there is no individual in the domain which it names.

Free logics do not require the existence of a referent and so allow for the truth of sentences containing empty singular terms. While there are a number of different free logics – including positive, neutral and negative – they share one basic feature in common.<sup>27</sup> That is, the traditional rules of quantification only apply to sentences containing referring singular terms. These rules are often reformulated as follows:

*Universal Instantiation:* From  $\forall xAx$  and  $\exists x x=t$  infer  $A(t/x)$  (where ‘ $A(t/x)$ ’ is the formula which results from ‘ $Ax$ ’ by replacing every occurrence of ‘ $x$ ’ by ‘ $t$ ’).

*Existential Generalization:* From  $A(t/x)$  and  $\exists x x=t$  infer  $\exists xAx$ .<sup>28</sup>

Thus, in free logics, singular terms only carry existential import if they are non-empty.

Sainsbury’s particular approach is to adopt a negative free logic in which simple sentences that contain non-referring singular terms are false.<sup>29</sup> This is contrasted with positive free logics, in which simple sentences containing empty singular terms can be true, and neutral (or, as Sainsbury calls them, Fregean) free logics, in which simple sentences containing them have no truth-value. Sainsbury’s motivation for adopting a negative free logic is that it allows us to speak meaningfully, indeed speak at all, about both what does and does not exist, while still getting the truth-values of sentences with empty names right. While he believes that a general semantics should allow for the logical truth of the sentences in which empty terms figure, he does not believe that it should allow for the truth of simple sentences containing them. So, for any empty singular term, the predication of any property to it will turn out to be meaningful but

---

<sup>27</sup> Rolf Schock, *Logics Without Existence Assumptions* (Stockholm: Almqvist & Wiksell, 1968); and Timothy Smiley, “Sense without denotation,” *Analysis* 20 no. 6 (1960): 125-135. Lambert’s was a positive free logic, while Schock developed the first negative free logic and Smiley was the first to suggest a neutral free logic.

<sup>28</sup> Sainsbury, *Reference Without Referents*, 65.

<sup>29</sup> Sainsbury defines a simple sentence as “one constructed by inserting  $n$  referring expressions into an  $n$ -place predicate”. *Reference Without Referents*, 66, fn. 9.

false. The reason for this, quite unlike that of classical logic, is not that these sentences presuppose existence, but rather that there is no term in the sentence which denotes an individual for a property to be predicated of.

The upshot of this is a semantics according to which sentences such as “Holmes was a detective” and “Sherlock Holmes does not exist” are meaningful and the former is false while the latter is true. Moreover, it does this in a way that respects both the Rule and Postulate. As a result, it serves well as a model for the semantics of  $\phi$ . What it does not respect are claims (1) and (2) above, but this is not surprising. Sainsbury set forth to develop a general and univocal semantics, and as I have granted above, no such general account will honour the Rule and Postulate while providing for the truth of (1) and (2). Sainsbury himself acknowledges as much, but is untroubled by it.

Sainsbury is not troubled because he does not believe that the correct way of dissolving the sorts of inconsistencies that we are concerned with is to recognize the ambiguity between  $f$ -sentences and their scopes.<sup>30</sup> This is because he is concerned about sentences incorporating singular terms from two distinct fictions, such as “Anna Karenina is more intelligent than Emma Bovary.” Sainsbury’s worry is that the  $f$ -operator is not sufficiently flexible to account for the truth of such claims. This claim is puzzling, as he seems to suggest just such an approach in his later work.<sup>31</sup>

### 1.3.2 Proper Names

The primary motivation for Sainsbury’s project is his desire to do justice to intelligible empty names. In developing his semantics, he adopts a broadly Kripkean approach.<sup>32</sup> Accordingly, he takes names to be rigid designators whose referents are fixed by an initial baptism.<sup>33</sup> However, in order to reconcile his stated goals with his desired account of the nature and origin of proper names, his account must diverge from Kripke’s in a number of small but important ways. First, Sainsbury claims that all proper names, including empty ones, designate rigidly. Second, he focuses on the name-using practices that baptisms originate, rather than the causal chains that perpetuate the use of a name after these baptisms.

---

<sup>30</sup> Sainsbury, *Reference Without Referents*, 214.

<sup>31</sup> Sainsbury, *Fiction and Fictionalism*, 150.

<sup>32</sup> Saul Kripke, *Naming and Necessity* (Cambridge, MA: Harvard University Press, 1980).

<sup>33</sup> Kripke’s account of proper names is well known and well rehearsed in the philosophy of language and metaphysics literature. As such, I will only talk about it in passing, to the extent that it serves as a point of departure for Sainsbury’s own account.



Sainsbury claims that allowing for the rigidity of non-referring singular terms does not distort Kripke's notion so much as it modifies it. This is because, in keeping with Kripke's account, it ensures that a name will have the same referent at every world where the referent exists (and none where it does not). Whereas, for Kripke, a baptism necessarily associates the name that it originates with exactly one object, for Sainsbury, a baptism associates the name with at most one object.<sup>34</sup> Sainsbury admits that empty names are not the norm, but explains that this does not mean that they are any less names (or any less meaningful). While it is true that, in general, not only are baptisms intended to associate a name with an object but they also succeed in doing so, this is not universally true of baptisms. Nor is it universally true of names, by their very nature, that they have referents. Just as non-empty names necessarily refer to the same object at all possible worlds, empty names necessarily fail to refer to any object at any worlds. Indeed, they not only fail to refer, but they fail to refer in the same way for the same reason.

To account for the fact that a name can be meaningful despite this failure of reference, Sainsbury introduces the notion of a name-using practice. This allows him to distinguish between two related, but conceptually distinct, success conditions for baptisms. The first is the association of a name with the object that it is intended to introduce. The second is the origination of a name-using practice. The first success condition is in line with the traditional Kripkean picture of reference. Since on his account a baptism associates a name with an object, some object must be introduced. One way in which the baptism may fail is if it inadvertently associates the name with an object or individual other than the one that the baptism is intended to introduce. However, for Kripke, such a failure will still succeed in introducing a name. The other way in which the baptism may fail is if it fails to associate the name with any object at all. For Kripke, such a baptism fails to introduce a name. The second success condition is in line with both Kripke's and Sainsbury's picture of reference. According to Kripke, when a name is used, it refers to its bearer because the user has the intention of using it to refer to the same object as previous speakers. So, a successful baptism will initiate a causal chain in which all speakers in the community use the name in the same way to refer to the same object. This occurs because learning the referent of a name requires that speakers use the name in the same way as the person who introduces them to it. It is in virtue of this causal relation that the speaker refers to just that object introduced by the baptism.<sup>35</sup> On this view, the semantic significance of

---

<sup>34</sup> Sainsbury, *Reference Without Referents*, 106.

<sup>35</sup> Kripke, *Naming and Necessity*, 96.

the name is the object introduced by the baptism. Since on Sainsbury's account, a baptism originates a name-using practice, as long as the baptism originates a name that is used in the same way by all those initiated in the practice, the baptism has been successful. Typically – in the case of non-empty names – both success conditions are achieved; the goal of a baptism is to introduce a name which will then be used by those participating in the practice to refer to that object.<sup>36</sup> However, this is not universally true. In cases where a name-using practice is successfully initiated but an object is not successfully associated with the name, intentionally or not, empty names are introduced.

How is this possible? What difference between the traditional Kripkean view and Sainsbury's allows the intelligibility of empty names? According to the former, to understand a name is to know who the bearer is. According to the latter, to understand a name is to know what it would be for someone to be the bearer. One ties the meaningfulness of a name to the existence of a referent, the other to reference conditions. This is because the picture that focuses on the causal chain between baptism and speaker presupposes that as speakers we stand in a causal relation to a referent. By contrast, the picture that focuses on name-using practices, presupposes only that as members of a practice we are engaged in a practice that associates the name with a reference condition and use it accordingly.

### 1.3.3 Fiction

Sainsbury develops his views about the proper interpretation of discourse in and about fiction, beginning in *Reference Without Referents* and continuing in *Fiction and Fictionalism*. While the former is focused on semantic issues and the latter is ostensibly concerned with ontology, the approach introduced in *Reference Without Referents* is taken up and expanded upon in *Fiction and Fictionalism*. The interpretive strategy that he adopts is one of presupposition and replacement. That is, while he believes that most of the sentences in which names of fictional characters, objects and events figure are – strictly speaking – false, a proper understanding of them can be gained by the acceptance of certain presuppositions on the parts of both speaker and listener. Alternatively, a correct interpretation and evaluation of such sentences may rely on the understanding that such sentences need not be accepted as true and that their seeming truth relies on the fact that we can replace them with sentences which express what the speaker intends that they convey. Thus, for Sainsbury, the difference between the sentence as presented

---

<sup>36</sup> Sainsbury, *Reference Without Referents*, 91-92

and as properly interpreted is pragmatic rather than semantic. It is not that sentences in and about fiction have a special semantic character, but rather that what we use them to convey is something other than their literal meaning.

Sainsbury takes what he calls an “irrealist” approach to fiction and its ontology. This is, in essence, a non-objectualist approach. He denies the existence of fictional entities and does not believe that there are exotic, or abstract nonexistent, objects. As such, he accepts both the Postulate and the Rule. He also wants to account for the apparent truth of (1) and (2), to explain how it is that we speak and act as though fictions have objects and how we are able to individuate, have knowledge of, make true claims about and draw correct inferences regarding the objects of fiction. Moreover, he wants to do all of this while denying that there is any special semantic category of fictional sentences or fictional names.<sup>37</sup> That is, he rejects the adoption of an *f*-operator as an ambiguation strategy for fictional discourse.<sup>38</sup> He does so because, as stated above, his is a wholly general semantics.

According to Sainsbury, one way of explaining our understanding of and discourse about fiction relies on presupposition, or the acceptance of certain premises (which we know to be false or indeterminate) as true, for the purposes of evaluating the propositions in question.<sup>39</sup> In the case of Sherlock Holmes, the sentence “Holmes lived at 221B Baker Street” presupposes that Holmes exists as well as all of the information imparted to us about Holmes by the stories. We understand the sentence under this presupposition and take the sentence to be true relative to the presupposition. Important to Sainsbury’s account of presupposition-relative truth and understanding, is the claim that it is no sort of truth at all and that in accepting the presupposition and understanding the proposition with respect to it we do not actually come to believe it. In order to do this however, he must claim that what the lay reader thinks is truth about the story or character – call it truth-in-fiction or the truth of the relevant *f*-sentences – is really not truth at all, but rather fidelity to the story or presupposition-relative truth. Given these constraints, how can he account for our ability to engage with fiction as we do?

Sainsbury claims that one of the advantages to this approach is that it accounts for the seeming truth of many sentences which he does not believe can be implicitly prefixed by “in the

---

<sup>37</sup> “Semantics will recognize no special category of fictional sentences or fictional names. Everything will proceed just as for non-fictional regions of language”. Sainsbury, *Reference Without Referents*, 202.

<sup>38</sup> It is interesting to note that, in certain circumstances, he does accept the implicit prefixing of fictional sentences with phrases such as ‘In the fiction’.

<sup>39</sup> Sainsbury, *Fiction and Fictionalism*, 145.

fiction.” That is, while he rejects the *f*-operator which semantically encodes the information communicated by such a prefix, he accepts that this information may sometimes be pragmatically imparted. However, he believes that the inference that the sentence is so prefixed is only valid in the case in which no actual objects or individuals figure – that do not express propositions attributing attitudes (propositional or otherwise) about the fictional to actual individuals or those comparing the fictional to the actual. That is, we can take “Holmes lived at 221B Baker Street” to be prefixed, and we can similarly take “Holmes was much cleverer than Inspector Clouseau” to be, but we cannot take “Sarah Palin admires Mrs. Bennett” to be so prefixed. In the case of multiple fictions, we can prefix the sentence under consideration as follows: “In the *Sherlock Holmes* stories and *Pink Panther* movies, Holmes was much cleverer than Inspector Clouseau.” So we have a mechanism by which we can adequately accommodate the prefixing of multiple fictions in order to meaningfully discuss their objects. However, he does not believe that we can prefix a sentence in which an actual individual figures without implying that the claims made in the sentence are merely fictionally, rather than actually, true. That is, we cannot prefix “Sarah Palin admired Mrs. Bennett” with “In *Pride and Prejudice*,” as to do so would be to claim that it is merely fictional than Palin admires her.

Although Sainsbury believes that prefixing is an acceptable interpretive strategy for sentences in which characters from multiple fictions figure, he suggests that it may not be desirable, expressing a preference for presupposition in such cases. Thus, in the case of Holmes and Clouseau, we ought to accept the presupposition that the individuals in question exist – despite the fact that we do not actually believe them to – which, Sainsbury claims, will allow us to genuinely assert the comparison.<sup>40</sup> He prefers this strategy because while we can properly understand the sentence as referring to both stories, the sentence is intuitively true, but actually false. The Conan Doyle’s stories do not mention Clouseau or imply anything about him, nor do the *Pink Panther* movies mention Holmes. While it is clear that Holmes is quite clever and Clouseau quite dim, according to neither story (nor their agglomeration) is Holmes compared to Clouseau. So, neither the sentence in question nor its prefixed version are literally true, while the sentence is true under the presupposition that these individuals exist and that the stories describe them accurately.<sup>41</sup>

Presupposition and implicit prefixing are just two of a variety of interpretive strategies

---

<sup>40</sup> Sainsbury, *Fiction and Fictionalism*, 125.

<sup>41</sup> Sainsbury, *Fiction and Fictionalism*, 122-123.

that Sainsbury suggests. In the case of sentences like “Sarah Palin admires Mrs. Bennett,” he explains that while presupposition can account for their truth, replacement is also an option. In this case, the sentence is literally false. However, we can replace it with one that communicates the same information without committing us to the existence of exotic objects.<sup>42</sup> For example, while “Sarah Palin admires Mrs. Bennett” must be false, “Sarah Palin admires some characteristics and actions ascribed to Mrs. Bennett” might be true.<sup>43</sup> In this case, while the truth of the sentence in question could be accommodated under the presupposition that Mrs. Bennett exists, the replacement sentence would be straightforwardly true if correct. As Sainsbury believes that fictional truth should be reduced to actual truth, if at all possible, he claims that replacement is preferable in such cases.

The interpretive strategies offered by Sainsbury are all attractive. Moreover, for each of the sentence types above, some or all of them seem to be effective. So why look any further for an explanation of fictional discourse? Why seek a semantic solution when there are readily available pragmatic interpretations?

First, it is important to note, that for the most part the development of these strategies occurs in *Fiction and Fictionalism*, a work not primarily concerned with semantics, but with ontology. Sainsbury’s aim is to show how we can account for our understanding of and intuitions about fiction and discourse about it without taking an objectualist stance. Thus, his discussion of how best to interpret this discourse is not intended to provide a systematic schema by which to determine either the propositions expressed by the sentences in question or their truth conditions. Rather, it is to show that given our understanding of fiction as a mode of artistic discourse, we have at our disposal a number of tools which are effective for accounting for the seeming truth of statements in and about it without adding to our ontological commitments. While Sainsbury does in places refer to the semantic project of *Reference Without Referents* as support or explanation for his claims, his approach does not require the acceptance of it.

Second, the motivation for the semantics laid out in *Reference Without Referents* is to provide a univocal account of referring expressions. This is premised in part on the claim that our semantics should be descriptive and not revisionary. Sainsbury argues that the theoretical categories of a good semantic theory will reflect the natural categories of the language it

---

<sup>42</sup> Sainsbury, *Fiction and Fictionalism*, 119.

<sup>43</sup> Sainsbury, *Reference Without Referents*, 208.

describes.<sup>44</sup> This is chiefly used to defend the claim that the interesting theoretical distinction is not between empty and non-empty names, but between the types of success conditions that are placed on their baptisms. However, it further serves as justification for the claim that, “Semantics will recognize no special category of fictional sentences or fictional names. Everything will proceed just as for non-fictional regions of language.”<sup>45</sup> While this claim seems to follow from Sainsbury’s commitment to a descriptive semantics, it ignores the double-aspected nature of fiction, which is exactly what he is accounting for in *Fiction and Fictionalism*. We can accept the fact that our ontological minimalism commits us to the straightforward falsehood of discourse about so-called fictional objects and individuals, while seeking a systematic account of their apparent truth.

The question of whether our behaviour and language use with respect to fiction is best explained by a semantics that captures the systematic ambiguity inherent in it or a pragmatic account that guides our interpretation is an important one. However, in order to answer it, more work must be done. I believe that the answers to this question lie in our cognitive responses to fiction rather than an account of them as linguistic phenomena. It is not until we better understand why we talk, think and reason about fiction in the way that we do that we will be able to decide how to account for our linguistic behaviour with any confidence. While this approach may provide useful guidance in our decision to treat these issues as semantic or pragmatic ones, it may also leave the question open. So, the task of determining how best to treat questions regarding language use must be delayed until we have a clearer understanding of the nature of our responses to fiction.

Kendall Walton takes such an approach. He begins by examining the nature of our engagement with representations in general and with fictions in particular. He then uses his account of engagement to explain an array of responses that we have towards the fictional, including the affective, intentional and doxastic. In the next chapter, I critically discuss his view and consider the lessons that we can learn from it. Ultimately, however, I believe that careful attention to the empirical data generated by experimental psychological studies of these responses will suggest a theory that can provide a treatment of fiction in a way that both systematically and accurately captures their peculiarities.

---

<sup>44</sup> Sainsbury, *Reference Without Referents*, 95.

<sup>45</sup> Sainsbury, *Reference Without Referents*, 202.

## 2 A Tale of Two Fictions

### 2.1 Representation

In the introduction to this thesis, I decided to forego the spirit of Frege's Dismissal. I said that fiction's distinctive features call for theories which are purpose-built for them. If this is right, we should be hesitant to integrate an account of our behaviour and language use with respect to fiction with established semantic theories, the development of which were in no way motivated by the special problems that fiction presents. It is instructive to examine Walton's contributions against the background of these claims. There is a certain extent to which his views conform to them, but he does not accept them outright. That is, while rejecting Frege's Dismissal, Walton does not think that theories of fiction must be purpose-built. His treatment of fiction should be seen in relation to a more comprehensive theory. The more comprehensive theory is a theory of artistic representation, and Walton's account of fiction is somehow subsumed or comprehended by it. We have it, then, that Walton proposes to tell us what novels are like by extension from the representational status of all works of art – what he will tell us about music, painting, photography and so on.

It is a further peculiarity of his view, and a complication of it, that Walton uses the term "fiction" in a quite technical way, according to which *any* work of representational art is a fiction. The *Mona Lisa* is a fiction, Beethoven's late sonatas are fictions, Jackson Pollock's *Number 5* is a fiction, John Cage's *4'33''* is a fiction, and of course *Anna Karenina* is a fiction. Walton's purpose is to construct a general theory of fictions in this technical sense, and then to account for fictions in the literary sense by making the requisite qualification. It would be going too far to suggest that literary fictions are catered as a mere after-thought of the more capacious theory, but neither is it quite true to say that Walton's account of literary fictions was purpose-built for them.

Walton's technical use of "fiction" invites a confusion which we would do well to avoid. To this end, I will use the term "fiction<sub>L</sub>" to refer to literary fictions (that is, fiction in the ordinary language sense of the term) and "fiction<sub>W</sub>" to refer to Waltonian fictions (that is, representational art works of all kinds). *Anna Karenina* is accommodated not in a stand-alone theory of fiction<sub>L</sub>, but in a theory of fiction<sub>W</sub> extended to the particular requirements of fiction<sub>L</sub>. These are necessary qualifications.

“Fiction<sub>w</sub>” is a term of art for Walton (if the pun might be excused), according to which “representation” and “fiction<sub>w</sub>” are synonymous. “Representation,” in turn is also given a quite different meaning from that we find in common usage. His own description of the project that he is undertaking is telling:

My decisions about how to shape the category of representations and the reasons for them will emerge gradually as our theory develops. Indeed to construct a theory, to achieve an understanding of things, is in large part to decide how best to classify them, what similarities and differences to recognize and emphasize. Determination of the scope of our investigation will thus be largely a result of it. We won’t know what, beyond my initial examples, the theory is a theory of until we have it in hand.<sup>46</sup>

I think that Walton is accurately describing his own project and we will see that this is the case as we continue to explore his theory.

Common usage reflects a fairly well-established sense of “representation” answering to the following conditions:

(1) If  $x$  is a representation, then  $\exists y(x \text{ represents } y)$ .

More generally, for any representation,  $\langle \exists y, z \dots (x \text{ represents } y \text{ and } z \text{ and } \dots)$ .

(2) If  $x$  is a representation of  $y$ , then  $x$  is a representation of  $y$  as  $z$ .

That is, representations depict their objects as having certain features or qualities.

Walton’s representations and fictions are one and the same.<sup>47</sup> Representations in his sense include the *Mona Lisa*, Beethoven’s sonatas and Pollock’s *Number 5*. But for these ordinary meanings, some of these works of art are representational while others clearly are not. The *Mona Lisa* and *Anna Karenina* are representational while *Number 5* and *4’33’’* are not.

Anybody with a nodding acquaintance of the arts knows that there is a common sense distinction between representational and non-representational art, but there is no hint of its presence here. It is difficult to mark the divide in any precise and principled way. For example, is the ballet *Giselle* representational? Is Marta Carrasco’s *J’arrive...!*? We know that some sculpture, such as Michelangelo’s *David*, is representational, but what about works of abstract sculpture like Henry Moore’s *Oval with Points*? Tchaikovsky’s *1812 Overture* represents Napoleon’s defeat – or

---

<sup>46</sup> Walton, *Mimesis as Make-Believe*, 3.

<sup>47</sup> Nelson Goodman, *Languages of Art: An Approach to a Theory of Symbols* (Indianapolis, IN: The Bobs-Merrill Company, 1968).



the battle of Borodino. John Cage's *4'33"* represents nothing. For present purposes, it is enough to note the deep and common sense consensus that there are many works of art that are not representational. So we should emphasize the non-standardness of Walton's technical appropriation.

Care should be taken with the common sense conditions of two paragraphs ago. On the face of it, they violate my (and Walton's) No-Object assumption, itself derived from Parmenides' Rule and the Non-Existence Postulate. Condition (1), for example, asserts that *Anna Karenina* is representational fiction in the sense of "representational" only if there *exists objects* which *Anna Karenina* represents. This is problematic. Since it is true by the lights of everyday intelligent speech but false according to my theoretical stipulations, I cannot use Walton's refusal of these theoretical falsehoods about representation as occasion for complaint. For why rebuke a philosopher for evading concepts laden with what, by my own insistence, have theoretically impermissible consequences?

We are in the grip of an important problem for any No-Object theorist posed by any form of discourse embedding quantificational clauses and/or referential singular terms. This is the place to note this difficulty, not solve it. My point at present is that, whatever we may think of the ordinary object-laden sense of "representation," Walton's representation is not it. When we bear in mind the extraordinary difficulty – if not impossibility – of purging the language of our ontologically austere theories of these merely purported objects in languages containing such devices, the risk is that neither my nor anyone else's no-objectism will be a consistently stateable thesis.

We will need another notational device to avoid confusion: I will use "representation<sub>R</sub>" to refer to representations in the generally established sense and "representation<sub>W</sub>" to refer to Waltonian representations. For "representation<sub>W</sub>," Walton imposes the following conditions: *x* is a representation<sub>W</sub> only if

- (1) *x* has the function of serving as a *prop* in a game of make-believe,<sup>48</sup>
- (2) *x* has the function of *prescribing* certain *imaginings*,<sup>49</sup>
- (3) *x* has the function of *generating fictional truths*,<sup>50</sup>

---

<sup>48</sup> "Representations, I have said, are things possessing the social function of serving as props in games of make-believe." Walton, *Mimesis as Make-Believe*, 69.

<sup>49</sup> "A prop is something which... mandates imaginings." Walton, *Mimesis as Make-Believe*, 69.

<sup>50</sup> "Props are generators of fictional truths, things which, by virtue of their nature or existence, make propositions fictional." Walton, *Mimesis as Make-Believe*, 38.

- (4)  $x$  makes certain *propositions fictional*<sub>w</sub>,<sup>51</sup>  
 (5)  $x$  makes certain propositions ‘*true in some fictional world*’.<sup>52</sup>

For ease of reference, I will speak of the italicized expressions in conditions (1) to (5) as the analytical vocabulary of Walton’s theory of representation<sub>w</sub>. The *Mona Lisa*, Beethoven’s sonatas, *4’33”*, *Number 5* and *Anna Karenina* are all representations<sub>w</sub>. Not all of them are counted traditionally as representational, or representations<sub>R</sub>. It is safe to say that the *Mona Lisa* and *Anna Karenina* are both representations<sub>R</sub>. Both *depict* or *portray* individuals and events of a kind that occur concretely in nature. Women occur concretely in nature. The *Mona Lisa* is a painting of a woman. She is seated and smiling enigmatically. She is represented as such. *Anna Karenina* is a story about a woman and the events which take place leading to her gripping demise. The deaths of women are events that occur concretely in nature. Tolstoy’s story represents Anna in these ways. Both these representations have *objects*, and this we might say is an important commonality among representations. They have objects not in any metaphysically loaded sense, but rather in the sense that it is within our intuitive grasp to identify what these representations are representations of. In short, we know that *Anna Karenina* is, among other things, about Anna Karenina; in the case of Beethoven's sonatas this is not so clear. Certainly, the music is expressive and evocative, but in what sense does it represent anything?<sup>53</sup> Similarly, we can ask of abstract works such as *Number 5* what they represent and how. On our working definition of “representation<sub>R</sub>”, these are not representational works. However, Walton includes them in the class of representations<sub>w</sub>. Why?

In an effort to dissipate some of the tension that arises from his rejection of the common sense distinction between representational<sub>R</sub> and non-representational<sub>R</sub> art works Walton calls on the distinction between “figurative” and “non-figurative” works. Walton says that a figurative work of art is a work that “prescribes imaginings” about the world and not the work, while a non-figurative work prescribes imaginings about the work itself. Georges Seurat’s *La Grande Jatte* “prescribes imaginings” about people and places, as well as itself; Kasimir Malevich’s *Suprematist Painting* only prescribes imaginings about itself. *Suprematist Painting* is a two-dimensional work, but the imaginings it prescribes include certain spatial relations between its

<sup>51</sup> Walton, *Mimesis as Make-Believe*, 38.

<sup>52</sup> “Fictional worlds are associated with collections of fictional truths; what is fictional is fictional in a given world.” Walton, *Mimesis as Make-Believe*, 69.

<sup>53</sup> Many people think that Tchaikovsky's “1812 Overture” represents the defeat of Napoleon by the Russians and that there is a lot of narrativity in Beethoven's scores.

elements, including that there is a yellow rectangle in front of a green trapezoid.<sup>54</sup> Walton thinks that both types fall under the broader class of representations<sub>w</sub>, once again departing from the common sense understanding of the term.<sup>55</sup> For Walton, all art is representational art.

How are we to understand this notion of prescribing imaginings? Imagination is the sort of thing that is susceptible causal stimulation and causal inhibition. But *prescription* is a sort of rule-setting or order-giving. In virtue of what does *Suprematist Painting* exercise the function of giving orders? What conditions would have to be met before such instructions were complied with? These are all open questions and insofar as Walton attempts to address them, it is by thinking about children's games of make-believe. Walton tells us that his theory of fiction<sub>w</sub> is based on children's games of make-believe, so we would do well to look for answers there. To what extent does the appeal to these games answer these questions?

Walton models the "games of make-believe" we play with representational artworks on some well-known children's games, such as Cops and Robbers and Tea-Party. He here introduces a number of load-bearing technical terms from his analytical vocabulary. Together these technical terms are offered as the underpinnings of his account of fiction<sub>w</sub> and allow him to offer solutions to a number of problems that must be accounted for by a theory of fiction<sub>I</sub>. They include conditions on what is often called "truth in fiction" and the conditions enabling psychological and emotional responses to what these sentences represent. What are these features? Again, they are the terms of Walton's analytical vocabulary: *imagination*, *make-believe*, *props*, *principles of generation* and *fictional truths*.

## 2.2 Imagination

The notion of "imagination" is the foundation on which Walton's theory is built. Introduced without definition, imagination animates all of the other technical notions upon which the theory of representation<sub>w</sub> relies.<sup>56</sup> Walton claims a natural analogy between belief and truth, on the one hand, and imagination and fictionality, on the other. I myself am not so sure about the postulated analogy. When we have evidence that a proposition is true, we are

---

<sup>54</sup> Walton, *Mimesis as Make-Believe*, 54-57.

<sup>55</sup> Walton's definitions of figurative and non-figurative art are generally cashed out in terms of the visual arts. But he does draw an analogy between figurative painting and musical works which prescribe that we imagine certain happenings outside of the music (e.g. David slinging a stone at Goliath in Johann Kuhnau's Biblical Sonatas) and between nonfigurative painting and works which prescribe things of themselves (e.g. Bach's Art of the Fugue). See Walton, *Mimesis as Make-Believe*, 335-336.

<sup>56</sup> Walton, *Mimesis as Make-Believe*, 19.

warranted in believing it, but that does not mean that the world actually is as the belief indicates. But, on Walton's account a proposition is fictional<sub>w</sub> just in case there is a *prescription to imagine* it, and in imagining it is so, it becomes *fictional that* it is so. So what determines what we are supposed to imagine? Is there something about the actual world which determines what we are to imagine and thus what is fictional?

In the language of worlds, it amounts to this: When we have a *warrant* to believe that *p*, the world may be such that  $\sim p$ . But when it is prescribed that we imagine that *p*, *p* holds in the requisite *fictional world*. The difference is illustrated in the falsity of (1) and the truth of (2):

- (1) "We are justified in believing that *p*"  $\rightarrow$  "*p* holds in the actual world"
- (2) "There is a prescription to imagine that *p*"  $\rightarrow$  "There is a fictional world in which *p* holds"

This is an important difference. The falsity of (1) tells us something of substantive importance about belief, namely that even *justified believing* does not make it so. (2) tells us something tautological about imagining, namely, that imagining it makes it so in a fictional world. The truth of (2) embeds the unexplained notion of "a prescription to imagine that *p*." What is this?

It will make things clearer to define the concepts that Walton cashes out in terms of imagination rather than attempting to define imagination itself. I begin with the notion of "make-believe." If imagination is the foundation of Walton's theory, then make-believe is its cornerstone, a species of imaginative activity exemplified by children's games and characterized by the role that props play in these games. In building a theory of fiction<sub>w</sub> modelled on children's games of make-believe, Walton believes that he can deal with a wide variety of questions and problems inherent in our understanding of fiction<sub>w</sub>.

The model that he develops based on children's games is not a perfect one, in no small part because the relationship between children's games and our engagement with fictions<sub>w</sub> is not well characterized. He variously claims that it is one of analogy or one of similarity in kind. On the one hand, he explains that he takes "seriously the association with children's games.... We can learn a lot about novels, paintings, theatre and film by pursuing the analogies with make-believe activities like these."<sup>57</sup> The words "association" and "analogy" suggest that while there

---

<sup>57</sup> Walton, *Mimesis as Make-Believe*, 4.

are certain similarities between them, they are of different kinds. However, he also states that, “The activities in which representational works of art are embedded and which best give them their point are best seen as continuous with children’s games of make-believe,”<sup>58</sup> which suggests that children’s games and engagement with representational<sub>w</sub> works are of a kind. This leads to a confusion about how we are to understand his claims. Given the weight that the theory puts on this relationship, it is particularly problematic. Is Walton building a model, or is he explaining one member of a class (engagement with art works) in terms of another (children’s games)? It will help to consider an example of a children’s game and see the extent to which the theoretical terminology appropriated by Walton captures our intuitive understandings of such games.

Consider the example of Billy and Jimmy playing Cops and Robbers. The game is a conventional solution to a coordination problem and a child who does not understand or abide by the conventions will fail at the game. Fortunately, most children know how to play make-believe from a very early age.<sup>59</sup> They pretend and act out different scenarios (make-believe<sub>M</sub>). And they understand, for example, the implications of the various roles within the game. Billy is the cop, Jimmy the robber. So the players make-believe<sub>M</sub> that Billy is the hero and Jimmy the villain. In doing so, Jimmy might imagine that he is a gangster in the twenties. However, knowledge of the conventions does not determine the outcome of the game. In fact, the outcome is underdetermined by the rules, leaving a lot of opportunity for creative endeavour. Justice may prevail, with Billy catching Jimmy, or the forces of evil may win, when Jimmy escapes with the loot. This outcome will depend on how the players choose to play out the scenario and shows the extent to which children’s games of make-believe are self-directed. Once the players establish the scenario that they will act out,<sup>60</sup> they will often use various toys and other objects to play the role of different object. A pillowcase may stand in for a bag of cash or for jewels. A toy gun may represent<sub>R</sub> a real one, but a twig or a child’s extended finger might fill this role. These may be called “props” in the traditional sense (props<sub>p</sub>). To this end, I will distinguish the technical notation “prop<sub>p</sub>” to denote the common usage from “prop<sub>w</sub>” for the Waltonian appropriation. Traditionally, the term “prop” comes from its use in theatre and film,

---

<sup>58</sup> Walton, *Mimesis as Make-Believe*, 11.

<sup>59</sup> Alan M. Leslie, “Pretense and Representation: The Origins of ‘Theory of Mind’,” *Psychological review* 94 no. 4 (1987): 412-426.

<sup>60</sup> In the case of Cops and Robbers, this is done by convention. In cases where the game is not well-established, this may be determined by stipulation.

where the production's property department keeps portable objects to use on set. Those objects are known as props and are used to represent<sub>R</sub> other objects.

When Walton characterizes these familiar games, he appropriates and rewrites this standard description of a children's game of make-believe<sub>M</sub> for his own purposes, making it "true" that games like Cops and Robbers are characterized by "prescriptions to imagine." Walton gives what can only be described as a theoretical redescription of games of make-believe in which the vocabulary standardly associated with these games drops out and a new vocabulary emerges. This is the same technical vocabulary that forms Walton's analytical vocabulary for this theory of fiction<sub>W</sub>. This redescription likens games to works of art, particularly theatrical performances, while claiming that this in fact shows that artworks are like games.

For example, Walton claims that the use of "props" is an essential feature of games of make-believe. For Walton, "prop," is a term strikingly alienated from this original meaning, in the manner of "fiction" and "representation." So here, too, we should take care to distinguish between the ordinary and Waltonian uses of this term.

A "prop<sub>W</sub>" is defined as follows:

- (1)  $x$  is a prop<sub>W</sub> only if  $x$  has the function of generating fictional<sub>W</sub> truths.<sup>61</sup>
- (2)  $x$  is a prop<sub>W</sub> only if  $x$  has the function of prescribing imaginings.<sup>62</sup>

Walton writes: "A fictional truth consists in there being a prescription or mandate in some context to imagine some thing."<sup>63</sup> The defining feature of representations<sub>W</sub> is just that they are props<sub>W</sub>. That is, they prescribe imaginings and generate fictional<sub>W</sub> truths; likewise for fictions<sub>W</sub>. It is the very fact that props<sub>W</sub> prescribe the imaginings that they do that makes them representations<sub>W</sub> for Walton. In building a theory of fiction<sub>W</sub> modelled on children's games of make-believe, Walton believes that he can deal with a wide variety of questions and problems inherent in our understanding of fiction<sub>W</sub>.

Recall the example of Billy and Jimmy playing Cops and Robbers. Walton thinks the toy guns (or twigs or outstretched fingers) that they use in their game are straightforwardly props<sub>W</sub>. So when Jimmy sneaks up on Billy and holds the "gun" to his back, it is fictional<sub>W</sub> that the

---

<sup>61</sup> Walton, *Mimesis as Make-Believe*, 37.

<sup>62</sup> "It is by mandating the imagining of propositions that props generate fictional truths." Walton, *Mimesis as Make-Believe*, 42.

<sup>63</sup> Walton, *Mimesis as Make-Believe*, 39.

robber sneaks up on the cop and holds him at gunpoint. Does this change at all if the gun that Jimmy is playing with is a real one? Not so. In using the real gun in the game, Walton thinks that the children still use it to prescribe certain imaginings. So, he thinks real objects can still act as  $\text{props}_w$ , prescribing certain imaginings about themselves within the game. In this case, the gun prescribes that we imagine that the robber has it trained on the cop. Now consider the case in which there is a struggle and Billy kicks Jimmy's gun out of the way. If Jimmy runs towards Billy and pretends to punch him, what  $\text{prop}_w$  prescribes the imagining that the robber has punched the cop? Is it the fist that Jimmy swings? What is the right way of understanding how such imaginings are prescribed? On a Waltonian interpretation of these events, to say that the fist is a  $\text{prop}_w$  seems to get something wrong. Rather, he would say that Billy and Jimmy are themselves both  $\text{props}_w$  in the game; their actions and interactions prescribe certain imaginings within it. Walton thinks that what makes it  $\text{fictional}_w$  that the robber punches the cop is that Jimmy performs a certain action towards Billy. Within the game, Jimmy is a  $\text{prop}_w$  that prescribes that we imagine that there is a robber, Billy is a  $\text{prop}_w$  that prescribes that we imagine that there is a cop, and that Jimmy swings at Billy prescribes that we imagine that the robber has punched the cop.

Walton says that the imaginings mandated by a given  $\text{prop}_w$  will depend on the principles of generation at play in a given game of make-believe. That is, the rules that determine which fictional truths are to be generated by a given  $\text{prop}_w$  in a given context. In the case of children's games these rules are seldom made explicit, as they are generally either familiar or obvious. As such, we are often unaware of the principles of generation at play. Walton thinks that we know intuitively how to use and respond to toy guns as  $\text{props}_w$  in games of Cops and Robbers.

So, Walton has taken a practice that is generally seen to be quite natural, one that we all understand intuitively, and redescribed it in quite a striking way. He has erected a theoretical framework around it and, in doing so, is establishing the groundwork for a much more comprehensive theory, redescribing the standard notion of games of make-believe<sub>M</sub> using his own analytical vocabulary. Using it to motivate his own account of how it is we engage in the social practice of playing pretend in order to use it as the basis for his explanation of the more complex practices of engagement with representational art. Deeply embedded in the whole notion of the  $\text{representationality}_w$  of these works is the analytical vocabulary employed in Walton's theoretical account of children's games.

However, the claim that our engagement with works of fiction<sub>w</sub> is analogous to these games of make-believe is at best a promissory note. We can only appreciate this analogy once we have considered the theory as a whole. The analogy does no work to further our understanding of our engagement with fiction<sub>w</sub> until we have a theory in place. The trouble is that Walton's use of the notion of a "game of make-believe" has no explanatory power when considered independently of its use within his theory. So like the other terms that Walton appropriates into his analytical vocabulary, "games of make-believe" is divorced from its ordinary meaning. Accordingly, we should not fall into the trap of false familiarity when considering the claim that when we appreciate a representation<sub>w</sub> we are using it as a prop in a "game of make-believe" in the ordinary sense. True, "game of make-believe" has an ordinary meaning. But if it is through it that we seek an understanding of the nature of our appreciation of works of art, it becomes manifest that such appreciation is not like games of make-believe<sub>M</sub> in a number of important ways.

First, children's games of make-believe are largely self-directed. The players establish the scenario that they will act out and only then use props<sub>p</sub> in ways that are appropriate to their game. When Jimmy and Billy are playing, they decide that they will play Cops and Robbers and then use their toy guns accordingly. Moreover, the same props<sub>p</sub> can serve as different objects for different games. A toy gun used one day as a handgun in a game of Cops and Robbers can play the role of a ray-gun in a game of Space Cowboys the next. Contrary to Walton's claims it is not a prop<sub>w</sub> that dictates their imaginings (whatever they are supposed to be), but the nature of the games that dictates how the props<sub>p</sub> are to be used. Second, a single game can have several props<sub>p</sub>, each playing different roles within the game. Not only might Jimmy and Billy's game involve a toy gun, but a whistle might play the role of a siren. True, once the role of a prop<sub>p</sub> has been established, it may enable further still further uses. However, it is the child's own understanding of Cops and Robbers that starts this process and determines the role that the toy gun is to play. This is not the case with representations<sub>w</sub>. As a prop<sub>w</sub>, a representation<sub>w</sub> prescribes what is to be imagined and there is a much narrower range of games authorized for it than is suitable for the props used in children's games. In these games of make-believe, the course of the game is determined by the prop<sub>w</sub> and the range of imaginings open to the appreciator is severely limited. This will become apparent as we examine the role that this concept plays in our engagement with various sorts of art works.



## 2.3 Fictional Worlds

Walton asks two fundamental questions about representations<sub>w</sub>: (1) what do these works represent? (2) in virtue of what do they do so? To answer these questions, he offers a general account of representation<sub>w</sub> and fiction<sub>w</sub>, modelled on his account of children's games of make-believe. The account provides that when engaging with an artwork, appreciators engage in games of make-believe of a particular sort. In response to question (1), we can now answer that what these works represent is of what the work prescribes that we imagine. This is because Walton thinks that a representation<sub>w</sub> is a prop<sub>w</sub> in a game of make-believe, and so a device for the prescription of certain imaginings. *La Grande Jatte* is a prop<sub>w</sub> which prescribes that we imagine that it is a beautiful Sunday afternoon in a park by the water with a couple strolling down the hill and boats on the water, and so this is what it represents. In response to question (2), we can say that it is in virtue of its function as a prop<sub>w</sub> in a game of make-believe that *La Grande Jatte* represents what it does. So, *La Grande Jatte* represents a beautiful Sunday afternoon in a park by the water because it is a prop<sub>w</sub> in a game of make believe. As a prop<sub>w</sub>, the painting prescribes certain imaginings, and so generates certain fictional<sub>w</sub> truths, including that it is a beautiful afternoon, that the park is by the water, that there is a couple strolling down the hill and that there are boats on the water. Walton says that fictional<sub>w</sub> truths are indexed to the works that generate them. The fictional<sub>w</sub> truths generated by *La Grand Jatte* are not just fictional, but *La-Grande-Jatte-fictional*. This is because these are the propositions made fictional<sub>w</sub> by the painting; the propositions that are true in the world of *La Grande Jatte*.

Walton's idea of work-indexed fictional truths is tied to his notion of fictional worlds. Walton says that there are two sorts of fictional worlds associated with artworks: the work world and appreciator worlds. Whereas the world of *La Grande Jatte* is a work world, the worlds of the games that we play with it are appreciator worlds. Appreciators are "self-reflexive" props<sub>w</sub> in games of make-believe. Our interactions with works make us props<sub>w</sub> in broader, appreciator games of make-believe. This is because, in interacting with representational<sub>w</sub> works, we prescribe to ourselves certain imaginings about ourselves. So, for example, when it is I who sees *La Grande Jatte*, I imagine of myself that I see the scene depicted before me. In looking at *La Grande Jatte*, it is fictional<sub>w</sub>, that is, it is fictional<sub>w</sub> of me, that I see a park by the water, the couple strolling down the hill and the boats on the water.

The notion of appreciators as self-reflexive props<sub>w</sub> is central to Walton's theory of our *psychological* engagement with fictions<sub>w</sub>. Walton thinks that although we cannot have genuine

emotional responses to fictional objects and events, it can be fictional that we do so. When we experience the psychophysical responses that characterize or are symptomatic of certain emotions (which he calls quasi-emotions), Walton says that this mandates that we imagine that we feel the emotions themselves. We do not feel these emotions. We do not feel any emotions. But it is *fictional<sub>w</sub>* that we do.

Consider the powerful example of Picasso's *Guernica*, which depicts the destruction caused by the bombing by air of Guernica by German and Italian planes in the Spanish Civil War of 1936 to 1939. It is *Guernica*-fictional that a woman stands grieving for a dead child, that a horse has been run through with a spear and that a man is trapped by a fire. It is also fictional<sub>A</sub>, that is, fictional in an *appreciator's* game played with *Guernica*, that the viewer witnesses these horrors. Further, as self-reflexive props<sub>w</sub>, if we gasp or begin to weep when viewing the painting, it is fictional that we experience the emotions of shock and grief. This is because gasping and weeping are the quasi-emotional states associated with shock and grief. (Note, the subscript "A" is a variation of the subscript "W.") In this way, Walton's theory of representations<sub>w</sub> purports to account for our psychological and emotional responses to works of art. Of course it does not, as Walton acknowledges. What it does – or may do – is account for our psychological and emotional responses<sub>w</sub> to those works of art.

One important feature of Walton's account is his use of examples. Rather than provide clear definitions of his technical terms, Walton attempts to sketch their meanings by using them to illustrate certain features and functions of various types of works of art. In some cases, his reflections are quite robust, as with his discussion of figurative paintings, particularly *La Grande Jatte*. In others it is less so, as with that of non-figurative paintings such as *Suprematist Painting*. Others still are merely dealt with *en passant*; notably, musical works and sculpture. Some are not dealt with at all, for example, dance and photography.

When he discusses works of art as representations<sub>w</sub>, Walton tends to focus on figurative painting in general and *La Grande Jatte* in particular. While he uses many other examples of figurative works to illustrate specific claims, only *La Grande Jatte* is used as a sustained example of a work of art *qua* representation<sub>w</sub>. That he takes figurative painting as paradigmatic of representations<sub>w</sub> is evidenced by the fact that, rather than discuss the non-figurative work *Suprematist Painting* as he does *La Grande Jatte* – explaining how it functions as a prop<sub>w</sub> in order to explain such ideas as what it is for a proposition to be fictional and what it is that makes it so – Walton only uses the example of *Suprematist Painting* to illustrate the differences between

figurative and non-figurative paintings. Without further guidance, how are we to understand the representational<sub>w</sub> nature of non-figurative works? What sorts of props are they and how do they prescribe imaginings? Moreover, how are we to understand the representational<sub>w</sub> nature of other art forms when Walton cherry picks examples from them to support individual claims that he makes throughout his account?

## 2.4 Bootstrapping

Had Walton provided sustained examples, and used them throughout his account to illustrate the representational<sub>w</sub> nature of a variety of art forms, he could have provided a better understanding of the theory as a whole by applying the analytical vocabulary and conceptual framework to a number of artworks. This would have given the reader a convincing understanding of how the various parts of the theory work together in their interlocking applications to those real things. This would also give the reader grounds for evaluating the theory's ability to explain our engagement with a work of art by considering how well the theory, as a whole, captures both the nature of the artwork and our engagement with it. Instead, Walton employs many examples as one-offs, using them to highlight certain aspects of his view while downplaying the fact that they do not fit well with others. Whether by design or inadvertence, this sparsity of data makes it very difficult to both decipher the definitions of the theory's technical vocabulary and evaluate the strength of the theory as an account of artistic representation<sub>w</sub>.

At this stage, an important question must be asked about Walton's theory of representations<sub>w</sub> in general as it applies broadly to works of art. In evaluating any theory, it is important to ask whether the theory correctly and adequately captures the nature of the phenomena under consideration. So the question arises: is there any reason to think that this account is right?

The way Walton has forged his theory makes this an uncommonly difficult question to answer. Recall that Walton has appropriated a number of common sense terms divorced from their ordinary meanings to serve as technical terms in his theory. He claims that their technical definitions will emerge as the theory develops. At the beginning of his first chapter, Walton asks "What is it to imagine?" He answers that "Fortunately, an intuitive understanding of what it is to imagine, sharpened by the observations of this chapter, is sufficient for us to proceed in our

investigation.”<sup>64</sup> He ends the chapter by reviewing the analytical vocabulary. “This, in brief outline,” says Walton “is the skeleton of my theory. Let’s flesh it out and see what it can do for us.”<sup>65</sup> However, in order to evaluate the theory we must first know whether we understand it, and we cannot determine this without knowing what he means by the technical terms that it employs. That is, to come to an understanding of the whole theory is to achieve an understanding of its otherwise undefined analytical vocabulary. However, it might strike some readers that it is not possible to understand the theory without first having been told what its key technical terms mean.

## 2.5 Implicit Definition

Walton gives us reason to think that his technical terms are defined *implicitly*, as it is they depend on their use within the theory that furnishes their meanings. In order to explore this possibility, we must consider how this implicitness works. If we turn to the philosophical literature for guidance, there are at least three ways in which Walton’s theory might be understood and thereby evaluated. First, it might be that the theory is an implicit *definition* of its terms. Second, the terms employed by the theory might be understood on the model of *semantic holism*, in the manner of Quine. Third, it might be best understood on Wittgenstein’s account of *meaning as use*. It will be helpful here to consider and evaluate each of these options in turn.

### IMPLICIT DEFINITION

To say that a theory  $T^*$  is an *implicit (stipulative) definition of  $X$*  is to say that  $X$  is governed by the definition

$$\phi =_{\text{Df}} \text{The True,}$$

where  $\phi$  is the conjunction of the members of  $T^*$ . (If  $T^*$  is infinite then a stipulation of the above form will be needed for each sentence  $\psi$  in  $T^*$ .)<sup>66</sup>

---

<sup>64</sup> Walton, *Mimesis as Make-Believe*, 19.

<sup>65</sup> Walton, *Mimesis as Make-Believe*, 69.

<sup>66</sup> Anil Gupta, “Definitions,” *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2009/entries/definitions/>>.

This means that for a formal language which includes sentences held to be self-evident within the theory, the technical terms included in it can be defined as what is true according to those axioms. Terms can thus be defined by their use within the theory, so long as the theory in question is *admissible*. A theory is admissible only if (1) it meets the eliminability criterion such that for any sentence containing  $\phi$ , it can be restated without using  $\phi$ , and (2) it respects the conservativeness criterion such that for any sentence which can be proved to be true using  $\phi$ , the same sentence can be proved to be true without using  $\phi$  (i.e. the definition does not add any new information to the theory).<sup>67</sup> So, if the technical terms contained in Walton's theory of representation<sub>w</sub> are implicitly defined by the theory, then their meanings are provided by their use in the axioms of the theory. Moreover, any sentences that can be stated or proved using those terms can be stated or proved without using them. However, Walton's theory cannot be understood as an implicit definition of its terms in this sense. This is because this notion of definition holds only for formal languages, while Walton's theory is formulated in natural language. More importantly, it is not clear that any of the sentences of the theory can be reformulated without appeal to the technical terms it employs. As the meanings of these terms are interdependent, sentences of the theory cannot be formulated without appeal to themselves. So I conclude that Walton's implicitness claim for his theory of artistic representation cannot be interpreted so as to make the theory an implicit definition of its technical terms

Quine adopts a semantic holism which requires that the meaningfulness of language reflect our rich empirical reality however it might be experienced. And "The unit of empirical significance is the whole of science."<sup>68</sup> While the truth predicate is affixed to a theory sentence by sentence, what is confirmed or disconfirmed is the theory as a whole. Confirming evidence applies to a theory *in toto*, not to its sentences one-by-one. Thus, we have space to decide which of its particular sentences to give up. This is because we cannot determine the truth-value of one statement in isolation from our other theoretical commitments.<sup>69</sup> Quine does not acknowledge meanings, but rather is concerned with the distribution of the truth predicate. However, there is no evidence that Walton would welcome or accept this brand of holism. Walton is concerned

---

<sup>67</sup> "Definitions," The Stanford Encyclopedia of Philosophy.

<sup>68</sup> W.V.O. Quine, "Two Dogmas of Empiricism," in *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953), 41.

<sup>69</sup> Peter Hylton, "Willard van Orman Quine," *The Stanford Encyclopedia of Philosophy* (Fall 2010 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2010/entries/quine/>>.

with meanings. Regardless of how we characterize his approach, he is attempting to provide his readers with the meanings of  $\text{fiction}_w$ ,  $\text{representation}_w$ ,  $\text{make-believe}_w$ ,  $\text{prop}_w$  and so on.

The question of whether Walton is an implicit definitionist in the Wittgensteinian sense is not an easy one to answer, not the least because Wittgenstein's doctrine of meaning as use is not well understood. According to Wittgenstein, "For a *large* class of cases — though not for all — in which we employ the word "meaning" it can be defined thus: the meaning of a word is its use in the language."<sup>70</sup> So, for Wittgenstein, in order to determine the meaning of a word, we must consider the various ways that it is used within a language. To understand what a word means is to understand the way that it functions within a discourse. However, as words can be used in many ways that are related while not identical, the meaning of most words will rely on a "family resemblance" or cluster concept<sup>71</sup>.

Walton's use of artworks as examples supports a Wittgensteinian interpretation of his claims. He relies on these examples to direct the reader's understanding of his claims. In this way Walton points to the particular features that works of art have in common and that are most obvious in his chosen examples, to show us how these terms are to be used. Further, there is a trend in the literature on analytic aesthetics that rejects the notion that art can be defined in terms of necessary conditions. Instead, aestheticians take seriously the suggestion that art is a cluster concept.<sup>72</sup> Again, Walton's appeal to example as well as his reluctance to explicitly define the technical terms essential to his theory, suggests that his account is best understood on a Wittgenstein account of meaning. Unfortunately, if all we have is appeal to a family resemblance, the sparseness of Walton's examples leaves us with too little data to fully grasp the meanings of these terms.

## 2.6 Literary Fiction

There is reason to think that Walton's theory of  $\text{fiction}_w$  may do better as a theory of  $\text{fiction}_l$  than as a theory of artistic representation. I have already explained in the introduction that amongst the virtues of Walton's account as a theory of  $\text{fiction}_l$  are the fact that it is both a No-Object and a psychology first theory. So how does the theory fare as a theory of  $\text{fiction}_l$ ?

---

<sup>70</sup> Ludwig Wittgenstein, *Philosophical Investigations*, eds. G.E.M. Anscombe and R. Rhees, G.E.M. Anscombe trans., (Oxford: Blackwell, 1953) Remark 43.

<sup>71</sup> Anat Biletzki and Anat Matar, "Ludwig Wittgenstein," *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2010/entries/wittgenstein/>>.

<sup>72</sup> Berys Gaut, "'Art' as a Cluster Concept," in *Theories of Art Today*, edited by Noel Carroll (Madison: University of Wisconsin Press, 2000), 46.

Walton thinks that fictions<sub>L</sub> are representations<sub>W</sub>. They are said to prescribe certain imaginings on the part of the reader. Among them, that what he is reading is a true account. For example, Walton says that, *Gulliver's Travels* prescribes that the reader imagines that he reads the journal of a ship's physician. In this same way, all works of fiction<sub>L</sub> are reflexive; they prescribe that readers imagine about themselves.<sup>73</sup> In the case of fictions<sub>L</sub>, much of what we are said to be prescribed to imagine is the truth of certain propositions. Specifically, we are prescribed to imagine that the statements that figure in the fiction<sub>L</sub> are true. As such, these propositions are fictional<sub>W</sub>. For example, in *Gulliver's Travels* it is fictional "that there is a society of six-inch tall people called Lilliputians."<sup>74</sup> Walton says that the fact that a work of fiction<sub>L</sub> prescribes certain imaginings warrants us in making corresponding assertions independently of whether we have reason to believe that these claims are true. Walton writes "the sentences of *Gulliver's Travels* warrant the assertive utterance 'A war was fought over how to break eggs' quite apart from whether they give us reason to think such a war actually was fought."<sup>75</sup> That is, in making such propositions fictional, the novel warrants their assertion within the context of the game of make-believe in which the fiction<sub>L</sub> serves as a prop<sub>W</sub>. For Walton, this is just what "truth in fiction" amounts to.

Essential to this account of representation<sub>W</sub> is the fact that there are different sorts of fictional<sub>W</sub> worlds associated with a work. Recall that for all artworks there is a given work world and there are usually multiple appreciator worlds. The work world is the world of the story, separate from the specific games of make-believe that readers play with it. An appreciator world is the world of a reader's particular game of make-believe. So, while it is fictional in the world of *Gulliver's Travels* that the book is a journal, it is not fictional in that world that I am reading that journal. However, when I read the book, I am prescribed to imagine that I am reading a journal. So it is fictional in my appreciator world that I do so. Walton says that this is because I am a self-reflexive prop<sub>W</sub>; I prescribe this imagining of myself in virtue of the fact that I read the book.

Walton uses the idea of appreciators as self-reflexive props<sub>W</sub> in an effort to explain our emotional responses to fictions<sub>W</sub>. His discussion of quasi-emotions and the fictional experience of emotions focuses on fictional<sub>L</sub> cases. This is a particularly important issue for our engagement

---

<sup>73</sup> Walton, *Mimesis as Make-Believe*, 117.

<sup>74</sup> Walton, *Mimesis as Make-Believe*, 36.

<sup>75</sup> Walton, *Mimesis as Make-Believe*, 71.

with fictions<sub>L</sub> because we need to explain how it is that we feel, for example, pity for Anna Karenina despite the fact that no such person exists or ever did exist or will. As he explains:

To allow that mere fictions are objects of our psychological attitudes while disallowing the possibility of physical interaction severs the normal links between the physical and the psychological. ... We do indeed get “caught up” in stories; we frequently become “emotionally involved” when we read novels or watch plays or films. But to construe this involvement as consisting of our having psychological attitudes towards fictional entities is to tolerate mystery and court confusion.<sup>76</sup>

Walton must provide an explanation of the fact that we experience ourselves as having such responses, even though, he thinks, they are not *emotional* responses in any conventional sense of the term. His reservation arises from the action-guiding character of real emotions. For a psychological state to count as a genuine emotion, we must have the ability to act on it. But we cannot act on our responses to fiction. We cannot rescue Little Nell. We cannot spirit Anna to safety.

Based on his account of appreciators as self-reflexive props, Walton argues that we do not actually have emotional responses to the objects and events of fiction, but rather that it is *fictional* that we do. We have real responses to fictions and those responses are not emotions, but are merely quasi-emotions – the psychophysical responses associated with them.<sup>77</sup> But because we are props in our own appreciator games of make-believe, this makes it fictional that we experience the emotion in question. So, when we read *Anna Karenina*, it is fictional that we learn about Anna’s plight and so fictional that we believe that she suffers. We experience quasi-pity as a result of realizing that fictionally Anna suffers. This makes it fictional that our quasi-pity is caused by a belief that she is suffering, and hence fictional that we pity her.<sup>78</sup>

Walton takes narrative fiction<sub>L</sub> as paradigmatic and the novel *Gulliver’s Travels* as a sustained example with which to illustrate his claims. This is quite useful for his discussion of fictions<sub>L</sub> as representations<sub>w</sub>, in view of the importance of revealing the nature of such works as props<sub>w</sub> in games of make-believe<sub>w</sub>. With the novel, he has an example that fits comfortably into

---

<sup>76</sup> Walton, *Mimesis as Make-Believe*, 196.

<sup>77</sup> Simo Saatela, “Fiction, Make-Believe and Quasi Emotions,” *The British Journal of Aesthetics* 34 (1994): 25-34.

<sup>78</sup> Walton, *Mimesis as Make-Believe*, 245. This is a paraphrase of Walton’s account of Charles’ fear of the slime. It is important to note that this adaptation for the case of Anna is not unproblematic. Charles’ fictional fear of the slime is triggered by physiological cues in the movie that he is watching which are not present when we read a novel like *Anna Karenina*.



such an account. If the novel as a whole is a  $\text{prop}_w$ , then once we are prescribed to imagine that it is a true account of events, we are further prescribed to imagine the truth of the sentences contained within it. This warrants our assertions of both the sentences themselves and the inferences that can be drawn from them, within the games of make-believe in which they serve as  $\text{props}_w$ .

Central to his discussion of  $\text{fiction}_L$  is his discussion of our psychological participation in the games of make-believe that we play with  $\text{fictional}_L$  works as  $\text{props}$ . By “psychological participation,” Walton means our emotional responses to representations. In discussing this, he draws exclusively from examples of our responses to novels, movies and plays – *Anna Karenina*, *Death of a Salesman* and *Othello*. This suggests that psychological participation is predominantly a problem for  $\text{fiction}_L$ , rather than for  $\text{representation}_w$  as a whole. However, instead of employing a single sustained example bridging between the discussion of  $\text{fictions}_L$  as  $\text{props}_w$  and that of the relationship between quasi-emotions and  $\text{fictional}_w$  emotions, he employs select examples to suit his specific purposes in each case. As a result, it is unclear whether the claims that he makes using the example of *Gulliver’s Travels* generalize to these other cases.

It is rather striking that, while these examples are used liberally throughout the discussion of psychological participation, the main example is not of any actual work of  $\text{fiction}_L$ , but rather one of Walton’s own construction. This is the example of Charles and the slime, a hypothetical scenario devised to aid in the discussion of our emotional responses to fiction. In this example, Charles is watching a horror movie in which “a terrible green slime... oozes slowly but relentlessly over the earth, destroying everything in its path.”<sup>79</sup> Charles’ heart begins to race and his muscles tense. He experiences quasi-fear as a result of realizing that, fictionally, he is in danger. It is fictional that “his quasi-fear is caused by a belief that the slime poses a danger, and hence that he fears the slime,”<sup>80</sup>

The example is problematic for a number of reasons. Film is quite a different medium from novels and short stories or even plays. Film presents the fictional by presenting us with images of real things. Even in cases where we see a monster, some real thing is playing the role of the monster. The audience’s quasi-emotions are triggered not only by the story, but also sensory input to which we have deep seated adaptive responses. This is demonstrably different from the case of novels and short stories, which present the fiction via linguistic representation.

---

<sup>79</sup> Walton, *Mimesis as Make-Believe*, 196.

<sup>80</sup> Walton, *Mimesis as Make-Believe*, 245.

Whereas a movie relies on actual sounds and images to evoke these psychophysical responses, novels and short stories do not. By using a film as his primary example in his account of our psychological responses to fiction<sub>L</sub>, Walton is relying on experiences as of the actual world. In the case of novels, and to a lesser extent plays, our experiences are quite different. We know what is true in the fiction and respond appropriately to it, but it is not clear how this triggers quasi-emotions, as his account of these responses relies on quite a different type of experience.

Not only is the horror movie example paradigmatic of Walton's account of our psychological participation with respect to fiction<sub>L</sub>, it is paradigmatic of his account of such participation with respect to representation<sub>W</sub> in general. Here too, he relies heavily on the narrative structures of fictional<sub>L</sub> works to account for why it is that we respond in the ways that we do. It seems then that this account is not an account of our engagement with fiction<sub>W</sub>, but of fiction<sub>L</sub>. In claiming that this account applies to all representations<sub>W</sub> (or fictions<sub>W</sub>), Walton is equivocating on "fiction," using features characteristic of fictional<sub>L</sub> works to make claims about representational<sub>W</sub> works broadly construed. But why should this be a worry? There are two reasons. First, in equivocating on "fiction," Walton presupposes that any account that he gives of psychological participation will hold for our participation in games involving all representations<sub>W</sub>, regardless of the medium, while presenting an account that is designed to cover our psychological participation with fiction<sub>L</sub>. As a result, he fails to account for the fact that we have far less access to the fictional<sub>W</sub> truths of representations<sub>W</sub> such as paintings and symphonies than we do to those of novels and movies. While it does seem plausible that our emotional responses to works of fiction<sub>L</sub> arise from our knowledge and understanding of what is true in them, this does not seem to be the case in for most representational<sub>W</sub> works. Second, Walton takes his account of fiction<sub>L</sub> to be an instance of the broader account. However, since he equivocates on "fiction," this cannot be the case. Instead, it seems that his account of fiction<sub>L</sub> is an adaptation of the more general case.

## 2.7 Adaptation vs. Instantiation

Walton works out his account of fictions<sub>L</sub> in his more comprehensive theory of fictions<sub>W</sub> - in his general theory of artistic representation. This is important. It commits Walton to the view that getting at the truth about literary fictions is impossible without getting at the truth about artistic representation. There is a dependency relation embedded in this claim, and therefore a relation *R* between the theory of literary fictions and the theory of artistic

representation in virtue of which this dependency obtains. Given the purported dependency, this account is not complete as long as we do not know what  $R$  is.

It is not that such  $R$ s are unheard of in the philosophical literature. Two come to mind at once – the relations in question are *instantiation* and *adaptation*.

Walton claims that it is difficult to fully spell out what his theory of representation<sub>w</sub> is a theory of, but points to such examples as *La Grand Jatte*, *A Tale of Two Cities*, *Hedda Gabler* and *The Telltale Heart*, noting that he takes these examples “to be paradigms of representational art” as well as “central instances of works of ‘fiction.’”<sup>81</sup> Further, he notes that he resists “the implication that our category is a species of a larger class of ‘representations,’ understood to include ‘nonfictional’ as well as ‘fictional’ ones.”<sup>82</sup> He does not take himself to be presenting a theory of fiction<sub>L</sub> as separate from his theory of representation<sub>w</sub>. Therefore, it is fair to read him as claiming that any theory of fiction<sub>L</sub> that he is proposing is an instantiation of a broader theory of representation<sub>w</sub>. Let me say, as a first pass, that a theory  $T$  instantiates a theory  $T^*$  *only if*

- (1)  $\Phi$  is a true sentence (proposition) of  $T$  *if and only if* it is a true sentence (proposition) of  $T^*$ ;
- (2) Every predicate in the analytical vocabulary of  $T_{FL}$  *entails* its syntactic counterpart in the analytical vocabulary of  $T^*$ . (For example, “prop” in  $T_{FL}$  entails “prop” in  $T_{FW}$ , even though the extension of “prop” in  $T_{FW}$  is a proper superclass of its extension in  $T_{FL}$ ).<sup>83</sup>

Of course, there is more to a full definition than the necessary conditions furnished by (1) and (2). But it is easy to see that there are features of Walton’s theory of fiction<sub>L</sub> ( $T_{FL}$ ) that clash with these conditions in relation to his theory of fictions<sub>w</sub> ( $T_{FW}$ ).

The trouble with ascribing to Walton’s theory the instantiation-interpretation of  $R$  is that the distinctiveness of fictions<sub>L</sub>, especially with respect to psychological issues, requires that there be some predicates in the analytical vocabulary of  $T_{FL}$  that have no occurrence in  $T_{FW}$ , hence satisfying (2) only vacuously and falsifying (1) outright. It is therefore more nearly true to say that  $T_{FL}$  is a *non-conservative* instantiation of  $T_{FW}$ , in the following sense:

---

<sup>81</sup> Walton, *Mimesis as Make-Believe*, 1.

<sup>82</sup> Walton, *Mimesis as Make-Believe*, 3.

<sup>83</sup> Recall that a theory’s analytical vocabulary is made up of the terms, technical or otherwise, that drive the theory’s definitions and other lawlike connections, if any. In Walton’s case, the analytical vocabulary is “prop”, “imagination”, “make-believe”, “prescription”, “fictional truth”, and “fictional world”.

For theories  $T$  and  $T^*$ ,  $T$  is a *non-conservative instantiation* of  $T^*$  *only if*

- (i)  $T^*$ 's analytical vocabulary is wholly contained in the analytical vocabulary of  $T$ ,
- (ii) The analytical vocabulary of  $T$  contains some terms not contained in the analytical vocabulary of  $T^*$ ,
- (iii) If  $\Phi$  is a sentence of  $T$  whose predicates are common to the analytical vocabularies of  $T$  and  $T^*$  then  $\Phi$  is true in  $T$  if and only if it is true in  $T^*$ ,
- (iv) There are true sentences of  $T$  not expressible in the analytical vocabulary of  $T^*$ ,
- (v) A term in the analytical vocabulary of  $T$  entails its counterpart term, if any, in the analytic vocabulary of  $T^*$ .

Note here too that I have advanced only necessary conditions on the relation of non-conservative instantiation. Sufficiency could not be claimed in the absence of the specification of other commonalities in the respective languages of  $T$  and  $T^*$ . What is more, the idea of non-conservative instantiation as developed here is itself an adaptation of the set theoretic idea of non-conservative restriction. This is a well-defined notion for formal languages, but a correspondingly unruly one for natural languages. To take just one example, there is a substantial, albeit not perfect, consensus about how the generic idea of entailment in formal languages should be interpreted. But there is nothing remotely close to a similar consensus about how entailment works for natural language predicates. Still, I offer this partial account in an attempt to elucidate Walton's claim about the containment of  $T_{FL}$  in  $T_{FW}$ . Moreover, it highlights the fact that there is still important work to be done in order to determine the meanings of the otherwise undefined terms of its analytical vocabulary of his broader theory of fiction<sub>w</sub> before we can say what sort of relation  $R$  is.

What, now, of the possibility that adaptation is a better interpretation of  $R$ ? It would be remiss of us to ignore an established usage of "adaptation" in the philosophical literature. This is the sense in which Hintikka's epistemic logic, say, is an adaptation of Lewis' modal system S4.<sup>84</sup>

---

<sup>84</sup> Vincent Hendricks and John Symons, "Epistemic Logic," *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2009/entries/logic->

Hintikka's epistemic logic is an adaptation of S4 because not only did he write new axioms for S4's modal operators, he also introduced new symbols for agents. He also preserved the distinguishing axiom of S4,  $\Box\Phi$  only if  $\Box\Box\Phi$ , except that the boxes are now the knowledge-operator K. Accordingly, what is  $\Box\Phi \vdash \Box\Box\Phi$  in S4 is  $\Box_A\Phi \vdash \Box_A\Box_A\Phi$  in Hintikka's logic. However, if the adaptation-interpretation of Walton's R is correct, it is unlikely that *this* is the sort of adaptation-relation that holds between Walton's theories of fiction<sub>L</sub> and representation<sub>W</sub>. This is because, once again, our present notion of adaptation works well for a formal system, but Walton's theory of fiction and representation is not a formal system but a natural language account.

It is more likely, then, that in saying that T<sub>FL</sub> is an adaptation of T<sub>FW</sub>, we mean that the truths of the former are similar in some way to the truths of the latter. That is, there is some sort of analogical tie between the two theories. The two theories share quite a bit in common, and Walton's account of representation<sub>W</sub> is supposed to illuminate fiction<sub>L</sub>. This is an interesting suggestion, but there is no clear account of such a relation in the philosophical literature and little prospect of working out what such an adaptation relation would amount to given the sparseness of Walton's account.

If Walton's account of fiction<sub>L</sub> is an adaptation and not an instantiation of his general theory of representation<sub>W</sub>, then we must return to the problem raised in our earlier discussion of that theory. We are either to understand fiction as a technical term divorced from its ordinary meaning, with its definition emerging by its use within the theory (fiction<sub>W</sub>) or we are to understand it according to its ordinary meaning, contrary to Walton's claims (fiction<sub>L</sub>).

If it is the former, then are we to suppose that, in assimilating fictions<sub>L</sub> to representations<sub>W</sub>, Walton is providing a semantic elucidation of the term "literary fiction" or that he is answering the question "In what sense are fictions<sub>L</sub> are representational?" Divorced of its ordinary meaning and with little guidance provided by the broader theory from which the literary case is an adapted, it cannot be an elucidation of the term "literary fiction" as we know little more about what it is to be a fiction<sub>L</sub> than we do about what it is to be a representation<sub>W</sub>. As a result it cannot account for the representational<sub>W</sub> nature of fictions<sub>L</sub> any more than it can for any other artworks. As such, it is impossible to evaluate it as a theory of fiction<sub>L</sub>. If it is the latter,

then a yet larger problem arises: the theory must simply be false. Fictions, while representations, are no more props in games of make-believe than are any other artworks.

Walton's general strategy is not what is at issue here. In fact, by taking an approach that is primarily concerned with our psychological engagement with fiction, he provides useful methodological guidance regarding how to proceed in developing a systematic and comprehensive theory of fiction which can explain the wide range of responses that we have to it. However, as I state in the introduction, I believe that it will be useful to consider empirical data which directly concern the nature of these responses, instead of attempting to explain our engagement in terms of or by analogy to other imaginative activities.

### **3 We Can Believe It for You Wholesale (Or, Total Acceptance)**

In the previous two chapters, I discussed two diametrically opposed theoretical approaches to fiction. Walton's approach privileges our psychological engagement. That is, he is concerned with explaining why it is that we appear to have emotional responses to and beliefs about fictional works and the individuals and events that figure in them. Sainsbury's approach privileges our semantic engagement. That is, he is concerned with explaining why it is that we appear to refer to and say true things about these same works, individuals and events. Despite these very different motivations and approaches, Walton and Sainsbury have one very important thing in common; they are No-Object theorists, like myself. As such, both develop theories of fiction based upon the foundational assumptions that there is nothing that does not exist (Parmenides Rule) and that the objects of fiction do not exist (the Non-Existence Postulate). Taken together, these assumptions commit us to the position that no object is an object of fiction.

Both Walton and Sainsbury make important contributions to the cause of non-objectualism, and it is worth asking what we can take from each in the aid of developing an account of fiction that is ontologically minimalistic while accounting for the peculiarities of the phenomenon under consideration.

I hold that Walton's primary contribution is that he expounds a psychology first approach to theorizing about fiction. Since his primary motivation is to provide an account of our engagement with fiction, he first examines why it is that we seem to have emotions and beliefs about the fictional. He then uses his explanation of our psychological engagement to ground an explanation of how we individuate, know about and talk about fictional individuals, objects and events. This is an important contribution because it highlights an important insight: if we are to develop a theory of fiction that truly reflects how we come to understand and use the names, sentences and information that fiction introduces to us, we must first understand how it is that we engage with these stories at the most basic level.

However, one of my primary points of departure from Walton is how we resolve the apparent inconsistency in our engagement with fiction. Walton's approach is to explain away such inconsistency, saying that it does not in fact exist. That is, he does not think that we have emotional responses to, beliefs about or make reference to fictional characters, but rather that it is fictional that we do. My approach is to embrace ambiguity and so to say that the inconsistencies do in fact exist and that it is the job of a theory of fiction to account for both

horns of the seeming dilemma.

By adopting the general semantic framework put forward by Sainsbury, we can explain our intuitions against the reality of these sorts of relations to the fictional. This framework accounts for our resistance to the truth of claims about fictional characters, objects and events. Such claims include those which express our relations to them. Moreover, as a general semantic theory, Sainsbury's *Reference Without Referents* can be used as part of a more robust philosophical theory that accounts for our behaviour and language use with respect to fiction.

The crucial methodological question that must be dealt with is how we can use insights about our psychological engagement with fiction to understand this linguistic behaviour. It seems that our experiences of emotional and doxastic responses to fiction are more immediate and thus provide us with greater insight into our engagement with fiction. Pretheoretically, our assertions that we are devastated by the fate of Anna Karenina and that we believe that Sherlock Holmes is exceedingly clever are unimpeachable in a way that our assertions that 'Anna' refers to Anna Karenina and 'Holmes' refers to Sherlock Holmes are not.

Further, there is a great deal of experimental data available in the psychological literature that examines our ability to acquire beliefs through fiction. I believe that this data will be helpful in explaining our engagement with fiction and that the insights gleaned from it will be useful for explaining why we ascribe truth or falsity to sentences in and about fiction, why we are able to individuate the objects and individuals that are referred to in both fiction and discourse about the fictional, and why we are able to infer from the sentences of fiction further facts about the fictional. I believe that the answers to these questions lie in the nature of our cognitive responses to fictions the peculiarities of which are due to the peculiarities inherent in fiction itself.

Among the psychologists working on issues to do with narratives (both fictional and otherwise), a number of studies have examined their ability to influence our beliefs about the actual world. Of them, some examine the extent to which certain types of content presented in fiction are compartmentalized as purely fictional (or beliefs purely about fiction), while other types are more readily incorporated into our overall store of beliefs.

There are a number of important parallels between the philosophical and psychological literature on fiction. Both disciplines seek to understand the nature of our engagement with fiction. Psychologists refer to this experience as transportation, the metaphorical phenomena by which we find ourselves immersed in or absorbed by what we read. Both also strive to characterize the mental state we are in when we engage with narratives, particularly of the



fictional variety. A number of psychologists have examined the relationship between comprehension of the narrative and acceptance of its propositional content. Moreover, both psychologists and philosophers have attempted to explain how it is that we come to form beliefs about both the actual world and the fictional world, even in those instances in which we have anterior knowledge that overrides the information presented by the narrative at hand. In what follows, I will provide a brief summary of the psychological literature on each subject, drawing parallels with the philosophical landscape as needed. I will then make a case for the relationship between these phenomena and our emotional responses to fiction. Finally, I will consider a number of consequences regarding the mental states associated with fiction and, in light of these findings, how our engagement with fiction might best be characterized.

Until very recently, there has been very little empirical psychological work done on the effects that narratives in general, and fictions in particular, have on our beliefs. Indeed, in a recent review article Green, Garst, and Brock note that “An authoritative reference in the field of persuasion and attitude change contains more than more than 60 pages of references but has no mention of the impact of narratives and fiction on attitude or belief change.”<sup>85</sup> Fortunately, the authors cited above, as well as such researchers as Richard Gerrig, Deborah Prentice and Keith Oatley have made headway into examining this issue in experimental contexts.<sup>86</sup> This work, while often pursued without consideration of the interesting and informative work done by philosophers, would seem to be a helpful supplement to those philosophers who are concerned with just these issues. Although introspection and intuition are useful philosophical tools and experimental philosophy is beginning to take off, we would be remiss if we did not at least attempt to align our theories with the results being produced by empirical researchers such as these. This thesis is more centrally concerned with examining the causal effects of fictional narratives on our attitudes and emotions about the fictional world, rather than the actual. However, I believe that the results being produced by psychologists on the effects of narratives, fictional and otherwise, on our real world attitudes can provide valuable insights into these issues

---

<sup>85</sup> Melanie C. Green, Jennifer Garst, Timothy C. Brock, and L.J. Shrum, “The Power of Fiction: Determinants and Boundaries,” in *The Psychology of Entertainment Media: Blurring the Lines Between Entertainment and Persuasion*, ed. Karen E. Dill (Mahwah, NJ: Lawrence Erlbaum, 2004) 161-176; and Eagly, Alice H., and Shelly Chaiken, *The Psychology of Attitudes*, Harcourt Brace Jovanovich College Publishers, 1993.

<sup>86</sup> Richard J. Gerrig, *Experiencing Narrative Worlds: On the Psychological Activities of Reading*. (New Haven, CT: Yale University Press, 1993); D.A. Prentice, & R.J. Gerrig, “Exploring the boundary between fiction and reality,” in *Dual-Process Theories in Social Psychology*, edited by S. Chaiken & Y. Trope (New York: Guilford, 1999) 529–546; and K. Oatley, “Why Fiction may be Twice as True as Fact: Fiction as Cognitive and Emotional Simulation,” *Review of General Psychology* 3 (1999): 101-117.

as well. To this end I will discuss a number of research avenues being pursued by psychologists on the effects of narratives on our doxastic states. I will then consider how this is suggestive for philosophers working towards an understanding of the psychological relations in which we stand to the fictional.

### 3.1 So Close and Yet So Far?

Green and Brock identify three phenomena as being central to the persuasive effects of narrative: transportation, proximity and comprehension. Correctly determining the nature of these three phenomena is crucial to understanding the nature of our psychological engagement with fiction. Transportation, in the psychological literature, is analogous with what Walton refers to as psychological participation, or what we might more generally think of as engagement. It is the extent to which the reader “suspends disbelief” or is “caught up in the story.” Proximity is a separate, but importantly related concept. The proximity of a story is the degree to which a story conforms to our own experience of the actual world. So, the greater the similarity between the world as represented in the story and the world as it actually is, the greater the proximity as experienced by the reader. This affects the extent to which the reader scrutinizes the information presented in the story, and so affects the level of transportation experienced by the reader. Lower levels of scrutiny result in higher degrees of transportation. Comprehension, the extent to which we understand a story and so achieve uptake, is an important notion in the psychological literature on narrative. One question that arises is the relationship between belief and understanding. That is, is believing what we read constitutive of understanding it, and so is belief a necessary component of comprehension?

Related to these three phenomena are three distinctions that arise repeatedly in the experimental literature. These distinctions are between fact and fiction, context details and context-free assertions, and compartmentalization and integration. Each of these distinctions plays a role in both the extent to and manner by which we come to acquire beliefs as the result of reading narratives.

The distinction between fact and fiction would seem, at least intuitively, to be a crucial one when considering the psychological effects of different narrative types. Intuitively, we are aware of the fact that a lot of what we know about the world is knowledge in virtue of being told that that is how the world is. Similarly, we know the difference between fiction and reality and would like to think that we do not believe things in virtue of reading them *in a work of fiction*.

However, with respect to belief acquisition, the data show that our responses do not subdivide in ways that we might expect.

The distinction between context details and context-free assertions is harder to capture. Prentice and Gerrig introduced these terms to highlight a difference between two types of information that can be introduced by a narrative. Context details are those “aspects of the story that were specific to the particular world.”<sup>87</sup> Context-free assertions are those that “have the potential to transcend their fictional origins.”<sup>88</sup> These are assertions that, whether expressly stated in or implied by the text, make more general claims about, for example, human nature – the morals of the story. Prentice and Gerrig demonstrated that, to a certain extent, the difference between these two types of information mediates the likelihood that the reader will acquire it as a belief. In their studies, they treated them as two distinct types of information. However, this effect is mediated by such considerations as proximity and the contents of reader’s prior beliefs. For example, a reader not familiar with pertinent facts about the setting of a realist novel may not know which claims are merely fictional and specific to the depicted world and which have been taken by the author to be commonplace claims about the world, depending on knowledge of the setting to reinforce them. Gerrig himself notes that “the two types of information are best thought of as anchoring a continuum,” but that “for [their] initial experiments, however, [they] made a sharp methodological contrast.”<sup>89</sup> Thus, while the initial experimental results are suggestive, difficulties in drawing such distinctions clearly as well as with replicating proximity effects, as will be discussed below, calls their significance into question.

Where the two above distinctions do seem to bear out is in a third, between compartmentalization and integration of the beliefs acquired as a result of our exposure to information in narratives. It seems that while being told that a certain narrative is fictional does not affect the extent to which the reader will acquire new beliefs as a result of the information in the text, this information does influence whether those beliefs are integrated into the reader’s overall store of beliefs about the real world or compartmentalized from it in a belief set specific to that narrative, or at least narratives of the same kind. Further, experiments such as those described below suggest that a further determinant of whether story information will be compartmentalized or whether they are integrated is the extent to which the reader perceives it

---

<sup>87</sup> Gerrig, *Experiencing Narrative Worlds*, 216.

<sup>88</sup> Gerrig, *Experiencing Narrative Worlds*, 216.

<sup>89</sup> Gerrig, *Experiencing Narrative Worlds*, 217.

to be a context detail or a context-free assertion. Thus, while being exposed to a piece of information in a story that is labelled as fiction may predispose a reader to compartmentalize his corresponding belief, if he takes the assertion of that fact to be context free, it is more likely that he will integrate it into his overall set of real world beliefs. Thus, even if understanding entails belief, it seems that many factors determine the nature of those beliefs.

## 3.2 Fiction by the Numbers

### 3.2.1 You Can't Not Believe Everything You Read

Gilbert considered two different models of information processing. The first, based on a simplification of Descartes, is a dual process view.<sup>90</sup> On this view, when we are exposed to new information, we merely comprehend, withholding either acceptance or rejection until we assess it. The second, based on a simplification of Spinoza, is a single process view according to which acceptance is part and parcel of comprehension. On this view, when we are exposed to new information, in order to comprehend it we must also accept it. Of course, upon further assessment, we may come to revise our beliefs and so reject the information under consideration. Indeed this often happens very soon after the initial acceptance-entailing comprehension. However, studies have shown that when subjects are prevented from entering the assessment phase of belief formation, it is not the case that they withhold judgement, neither accepting nor rejecting the information with which they are presented. Rather, they continue to accept the presented information until an opportunity for assessment arises.<sup>91</sup>

One such study, conducted by Gilbert and his collaborators was designed to determine how the belief formation process functions by attempting to induce results consistent with the Spinozan model that could not be accounted for on the Cartesian model.<sup>92</sup> They attempted to do this by interrupting the comprehension processes before subjects could have an opportunity to evaluate the information presented in order to determine whether beliefs were formed prior to evaluation.

The experiment was conducted using 71 female University of Texas at Austin undergraduate students as subjects. Each of the subjects was asked to read two experimental

---

<sup>90</sup> D.T. Gilbert, D.S. Krull, & P.S. Malone, "Unbelieving the Unbelievable: Some Problems in the Rejection of False Information," *Journal of Personality and Social Psychology* 59 (1990): 601-613.

<sup>91</sup> D.T. Gilbert, "How Mental Systems Believe," *American Psychologist* 46 (1991): 107-119.

<sup>92</sup> D.T. Gilbert, R.W. Tatarodi, & P.S. Malone, "You Can't Not Believe Everything You Read," *Journal of Personality and Social Psychology* 65 (1993): 221-233.

narratives in the form of crime reports. Subjects were told that each report contained both true and false statements and that they were reports of two unrelated crimes that had recently occurred in Austin. In both cases the colour of the text varied with its truth or falsity: true statements appeared in black text and seven false statements appeared in red text. Subjects were instructed that the false statements were taken from reports of crimes that were not related to the crimes that the true statements were about and so a statement labelled as true could not be taken as evidence against a statement labelled as false. The subjects were divided into two roughly equal groups; each of which was given two of four reports: one in which the false statements exacerbated the severity of the crime, the other in which they mitigated it. Subjects were instructed that they would be acting as the trial judge, evaluating the information presented in the reports and so should pay careful attention to the details about both the circumstances of crimes and those accused of committing them. The reports were presented to the subjects in the form of text crawling across the screen at a rate of approximately sixteen characters per second.

Additionally, as the reports were presented, strings of digits presented in blue text would crawl across the screen directly below them. This occurred while the false statements, presented in red text, crawled across the screen and the digits occupied the same amount of space on the screen as the text for each false statement. After each report ended, a string of fifteen blue digits would crawl across both lines. The subjects in each group were further divided into two roughly equal groups: one group was asked to ignore the numbers, and the other was given a hand-held counter and asked to record each time the digit '5' appeared, so that the group with the counters were both processing the information presented in the reports and performing the digit search task at the same time. The digit '5' appeared on the screen forty-four times in the course of the experiment and the subjects who performed this task, failed to detect it about nine percent of the time. This suggested to Gilbert and his team that the task, while not overwhelming, was quite demanding.

After reading their assigned reports, each subject completed a number of tasks. One asked them to consider the facts of each crime and recommend a prison sentence between zero and twenty years for the accused. A second asked them to answer questions on a questionnaire indicating, according to three nine-point scales their feelings about each of the accused (where '0' indicated 'neutral' and '9' indicated 'extreme dislike'), whether they thought counselling would be effective to rehabilitate the accused (where '0' indicated 'no help at all' and '9' indicated 'a great deal of help') and how dangerous they thought the accused was (where '0' indicated

‘slightly’ and ‘9’ indicated ‘extremely’). Finally, each subject was shown a set of thirty sentences and asked to indicate of each sentence whether it was a true statement from the first report, a false statement from the first report or did not appear in the report. They were asked to do the same task with a second set of thirty sentences with respect to the second report. Each of these sets of thirty sentences contained four sentences that had appeared as true statements in the report in question, seven sentences that had appeared as false statements in the report in question and nineteen sentences that had never appeared in the report in question.

Gilbert and his team predicted that the false statements, whether they exacerbated or mitigated the crimes described, would only effect the number of years that the subjects who performed the digit-search task chose as sentences for the accused in both reports. Taking the group that did not perform that task as the control group, Gilbert and his collaborators analyzed the data to determine the effect of interruption on the impact of the false statements.

When analyzing the data, the researchers separated the results of the experimental group from those of the control group. Within both the experimental group and the control group, for each report, they further broke down the data according to whether the false statements were exacerbating or mitigating. Moreover, within both the experimental group and the control group, they calculated the difference in the scores for both the sentencing task and the questionnaire task by subtracting the ratings given by the group who had read the report with the mitigating false statements from the ratings given by the groups who had read the report with the exacerbating false statements. In the case of the control group, the differences were marginal. In the case of the experimental group, the differences were significant in most cases. That is, while there was only a marginal difference between the sentences recommended by subjects in the control group who read the mitigating false statements and those recommended by subjects who read the exacerbating false statements, subjects in the experimental group who read exacerbating false statements recommended a sentence that was nearly twice as long on average as those recommended by those who read the mitigating false statements. Further, while there was only a marginal difference between the scores of those who read the mitigating and extenuating false statements within the control group with respect to likability and dangerousness, subjects in the experimental group who read exacerbating false statements rated the accused as significantly less likable and significantly more dangerous on average than those who read the mitigating false statements. Finally, while the differences in rating the benefit of counselling was quite small between those who read the mitigating false statements and the exacerbating false statements in

both the control and experimental groups, there was a significant difference in the latter that did not emerge in the former. This indicated to Gilbert and his team that when subjects were interrupted, they tended to act as though the false statements were true.

When analyzing the data from the third task – recalling whether sentences were presented as true in the report, false in the report or neither – the researchers compared the likelihood that subjects in the control group would recall sentences as true, false or neither with the likelihood of the same behaviour within the experimental group. Moreover, they further compared this for sentences that actually were presented as true, presented as false or not presented at all. The analysis found that for statements that were identified as true in a given report, not only was there no significant difference between the control group and experimental group in terms of the likelihood that subjects would recall them as true, but in both cases the likelihood that subjects would mistakenly recall them as false or neither was very small. For statements that were identified as false in a given report, the likelihood that subjects would mistakenly recall them as true was significantly higher in the experimental group than in the control group. Finally for statements that were not present in a given report, the likelihood that subjects would mistakenly recall them as true was significantly higher in the experimental group than in the control group. However the proportion of experimental subjects who mistakenly recalled false statements as true (21%) was considerably higher than that of experimental subjects who mistakenly recalled statements that did not appear at all as true (4%). This seems to indicate that when faced with a statement that they did not immediately recognize as true or false experimental subjects did not simply guess that it was true. If they did, then they would have done so for statements which did not appear in the reports as well.

Further analysis of individual results demonstrated that the subjects who mistakenly remembered more false exacerbating statements as true were reliably more likely to recommend longer prison sentences than those who mistakenly remembered fewer of these statements as true, and subjects who mistakenly remembered more false mitigating statements as true were reliably more likely to recommend shorter prison sentences than those who mistakenly remembered fewer of these statements as true. Gilbert and his team took this as evidence that these subjects actually believed these statements to be true, and did not just guess that they were true because they could not recall their truth value. If they were simply guessing then this would not have made a difference to the task of deciding an appropriate sentence, which they performed prior to the task of recalling whether a statement was true, false or neither. So, before

they were asked to reflect on whether these statements were true, subjects were already acting as though they believed them. The researchers concluded that these results supported the hypothesis that when presented with information we come to believe it in the process of comprehending it and only after we have done so – and only if given the opportunity – may we revise our beliefs.

### 3.2.2 Fictional Narratives Change Beliefs

Prentice, Gerrig, and Bailis set out to examine how readers' personal experiences and pre-existing beliefs mediate the extent to which information encountered while reading fictional narratives affects their real world belief set.<sup>93</sup> They focused their study on what they call "participatory responses" or "p-responses," which they characterize as non-inferential responses that readers have in light of various sorts of attitudes or dispositions to a given text. They assert that these responses are voluntary and may vary among readers. They wanted to study the influence of p-responses because they are among a cluster of response types believed by persuasion researchers to mediate how our attitudes (presumably of the propositional variety) may be influenced or affected by persuasive messages.

The main goal of fiction is not generally taken to be persuasion, and the structures of fiction tend to be very different from those of persuasive messages such as political speeches, public service announcements or philosophy papers. In light of this, Prentice, Gerrig and Bailis sought to determine how p-responses affect how fictional texts influence our attitudes about the actual world. They predicted that the extent to which a reader is motivated to scrutinize and critically evaluate the information to which they are exposed in the course of reading a work of fiction is mediated by the story's proximity to the reader's own experience. That is, a reader will be both more willing and more able to be critical of the information presented in a fictional narrative the more that the world as presented in the story conforms with the world as experienced by the reader. This prediction was motivated by the fact that previous studies had demonstrated that people are more motivated and more able to assimilate new information when that information holds personal relevance for them. Prentice and Gerrig predicted that this would hold in the context of fiction: "when readers encounter a claim about a world with which they are familiar in a work of fiction, they will evaluate the information given what they know to

---

<sup>93</sup> D.A. Prentice, R. J. Gerrig, & D.S. Bailis, "What Readers Bring to the Processing of Fictional Texts," *Psychonomic Bulletin & Review* 5 (1997): 416–420.



be true.”<sup>94</sup>

To test their hypothesis, Prentice, Gerrig and Bailis conducted two identical experiments, one at Yale and one at Princeton. Each experiment used two versions of the same short story, one set at Yale and one set at Princeton. So, each experiment had a “home-school” and an “away-school”. Each version had “mild unsupported assertions concerning well-known facts about the real world.”<sup>95</sup> That is, each had context-free assertions about the actual world that were either obviously true or obviously false. In neither case did the stories contain enough support for the assertions to stand up under scrutiny. The hypothesis was that students would be more likely to be persuaded by assertions made in the away-school stories, regardless of whether they were actually true or how much support was given for them. This would support the view that scrutiny increases with proximity.

In the first of these experiments, twenty-nine undergraduate students at Yale participated in this experiment. Nine students were used as the control set and given a story which was roughly the same length as the two experimental stories (twenty single spaced pages) but did not contain any information that was relevant to the experiment.<sup>96</sup> Ten were given the home-school story. Ten were given the away-school story. The plot of each experimental story included the kidnapping of a professor and a group of students. Each story involved the discussion of sixteen topics which contained statements about the actual world. In each experimental story, half of the assertions were true and half were false. For each topic of discussion, if one of the experimental stories made actually true statements about it, the other story made actually false statements about it. That is, neither story made actually true statements about the same topic. With the exception of the eight false statements in each story, all other information that was provided by each was actually true.

After reading their assigned stories, each participant completed a couple of tasks. One asked them to rate the story that they had just read in terms of interestingness, quality of prose and their perceptions of the characters. The other was irrelevant to the story and was simply used as filler to occupy them for roughly ten minutes.

After completing these tasks, each student filled out a questionnaire that contained 32 statements, sixteen of which were related to the information in the two experimental stories.

---

<sup>94</sup> Prentice, Gerrig, & Bailis. “What Readers Bring,” 417.

<sup>95</sup> Prentice, Gerrig, & Bailis. “What Readers Bring,” 417.

<sup>96</sup> The story was “The Raven” by Robert Twohy, published in *Ellery McQueen’s Mystery Magazine*.

Two versions of the questionnaire were used. In each questionnaire, half of the experimental statements were consistent with the assertions made in each experimental story and half were inconsistent. In each questionnaire, half of the experimental statements were actually true and half were actually false. For each topic, if one of the questionnaires contained an actually true statement about it, the other questionnaire contained an actually false statement about it. Participants were asked to indicate their level of agreement with each statement by rating it on a nine-point scale, where “1” indicates “strongly disagree” and “9” indicates “strongly agree.”

When analyzing the data, in order to account for the fact that readers were generally more likely to agree with assertions that were actually true, the researchers subtracted the average rating given by the control subjects for each statement from the ratings given by each of the experimental subjects. Thus the analysis was not of the ratings themselves, but of the differences between an experimental subject’s rating and the average rating by the control subjects. Where experimental subjects agreed with these statements less than the control group did the result was a negative score, whereas when they agreed more than the control group the score was positive. They then took the average of the difference of each participant’s ratings of statements in each of four categories: actually true and story consistent, actually true and story inconsistent, actually false and story consistent and actually false and story inconsistent. The subjects who read the home-school story were not significantly influenced by the assertions of the story. That is, they generally agreed with statements that were actually true and disagreed with statements that were actually false, regardless of whether or not their assessment was consistent with the story. The subjects who read the away-school story were more influenced by the assertions of the story, agreeing more with actually true and consistent statements as well as actually false and consistent statements.

Thirty undergraduate students at Princeton participated in the second experiment. The control story and two experimental stories were each read by different groups of ten students. The design and methodology of this experiment was the same as the Yale experiment, except that in this case the Yale story was the “away-school” story and the Princeton story was the “home-school” story.

The data collected from this experiment was analyzed in the same way as in the first experiment. The results from the Princeton experiment were even more significant than those from the Yale experiment. Subjects who read the home-school story generally resisted agreeing with any story consistent assertions whether actually true or actually false. Moreover, they tended

to resist agreement with actually true statements that were story consistent more than with actually true statements that were story inconsistent. Story consistency did not influence the level of agreement with actually false statements, which was quite low. Subjects who read the away-school story in this experiment had results consistent with those who read the away-school story in the first experiment. That is, they agreed more with story consistent statements than with story inconsistent ones, whether they were true or false.

The researchers then performed a further analysis of the combined results of the two experiments. The goal was to determine the extent to which students agreed with story consistent assertions, regardless of actual truth. In both experiments, subjects who read the away-school story were more influenced by story assertions – agreeing more with story consistent statements than story inconsistent statements – than subjects who read the home-school stories. However, whereas subjects in the Yale experiment who read the home-school story tended to agree with actually true story consistent statements, subjects in the Princeton experiment tended to disagree with all story consistent statements, whether actually true or actually false. Despite these discrepancies, this led the researchers to conclude that as proximity decreases, the influence of context-free assertions on our beliefs increases as does the level of integration of the corresponding beliefs.

### **3.2.3 Replications with Mixed Corroboration**

Wheeler, Green, and Brock set out to replicate the results of Prentice, Gerrig and Bailis in order to confirm that, in the absence of motivation or ability to scrutinize information presented in a narrative, readers tend to accept and integrate that information into their own beliefs, but that when they are motivated or able to scrutinize, as in cases of increased proximity, they are not inclined to integrate that information.<sup>97</sup> In order to do this they first adapted the Yale and Princeton experiments to be conducted at Ohio State University (OSU) using Yale as the away-school.

#### ***3.2.3.1 Initial Adaptation***

The experiment was conducted using 220 OSU undergraduate students as subjects. Sixty-three of the participants were used as control subjects and given the same control story as in the Prentice, Gerrig and Bailis experiments. Each of the remaining 157 subjects was given one

---

<sup>97</sup> S.C. Wheeler, M.C. Green, & T.C. Brock, "Fictional Narratives Change Beliefs: Replications of Prentice, Gerrig, and Bailis (1997) with Mixed Corroboration," *Psychonomic Bulletin & Review* 6 (1999): 136–141.

of either of the experimental stories. The away-school story was the same as the Yale story used in both of the Prentice, Gerrig and Bailis experiments.<sup>98</sup> The home-school story was the same as the Yale story with the setting changed to OSU. In each story, half of the experimental story assertions were actually true and half were actually false. For each topic, if a story assertion about that topic was actually true in the home-school story then the story assertion about that topic was actually false in the away-school story.

After reading their assigned stories, each participant completed a couple of tasks. As in Prentice, Gerrig and Bailis, the first asked them to rate the quality of the story which they had just read. For the second, two-thirds of the participants were asked to complete the same irrelevant task as in Prentice, Gerrig and Bailis. However, in this case, the other third of participants was asked to complete a portion of the verbal SAT and to complete the NC (need for cognition) scale.<sup>99</sup> The NC scale is used to gauge a subject's "propensity to engage in effortful thought."<sup>100</sup>

After completing these tasks, each student filled out a questionnaire that contained 32 statements, fifteen of which were related to the information in the two experimental stories. Two versions of the questionnaire were used. In each questionnaire, half of the experimental statements were consistent with the assertions made in each experimental story and half were inconsistent. Also, in each questionnaire, half of the experimental statements were actually true and half were actually false. For each topic, if one of the questionnaires contained an actually true statement about it, the other questionnaire contained an actually false statement about it. Participants were asked to indicate their level of agreement with each statement by rating it on a nine-point scale, where "1" indicates "completely disagree" and "9" indicates "completely agree." Additionally, each subject was given a test designed to measure their recall of basic plot information consisting of nine questions.

If the hypothesis and results of Prentice, Gerrig and Bailis were valid, subjects who read the away-school story should have been more influenced to change their beliefs to conform to story consistent statements than those who read the home-school story. As with Prentice, Gerrig and Bailis the data that was analyzed was the difference between the results for each

---

<sup>98</sup> The one difference between the stories in this study and those in Prentice et al. is that one topic was excluded in this case because it pertained only to Ivy League schools and so was deemed irrelevant to OSU students. Wheeler, Green, and Brock, "Fictional Narratives Change Beliefs," 141.

<sup>99</sup> J.T. Cacioppo, R.E. Petty & C.F. Kao, "The Efficient Assessment of Need for Cognition," *Journal of Personality Assessment* 48 (1984): 306-307.

<sup>100</sup> Wheeler, Green, and Brock, "Fictional Narratives Change Beliefs," 136.

experimental subject and the average agreement rating given by the control subjects for each statement. If the assertions made in the story influenced subjects' beliefs, experimental subjects should agree more with story consistent statements than the control group did and so have a positive score. Likewise, if subjects were not so influenced, a negative score should result for story inconsistent statements. Both predictions should have held regardless of the actual truth of the test statements. Thus, if Prentice, Gerrig and Bailis were correct, and agreement was moderated by proximity, the ratings given by home-school subjects should have shown little variation from those of control subjects, whereas those given by away-school subjects should have varied significantly. That is, scores should have been significantly higher in the case of story consistent statements and significantly lower in the case of story inconsistent statements, regardless of actual truth. However, the scores of both home- and away-story subjects were the same, with positive scores for story consistent statements and negative scores for story inconsistent statements. This indicates that while story assertions did influence readers' beliefs, this was not moderated by proximity.

In order to determine whether proximity effects were moderated by a subjects' need for cognition, Wheeler, Green and Brock further analyzed the results of those experimental subjects who completed the NC scale test. This was to determine whether the reason for the lack of proximity effect was lack of inclination to scrutinize information presented in narratives. The hypothesis was that subjects with a higher need for cognition (HNC) scrutinize story assertions equally in both home- and away-school stories, whereas those with a lower need for cognition (LNC) exercise less scrutiny in the away-school case than the home-school case, resulting in a higher level of influence of away-story assertions on the beliefs of LNC subjects. However, analysis of the results for HNC subjects showed no significant difference from that of the full experimental sample. This showed that a subject's score on the NC scale did not determine whether their level of scrutiny, and so their resistance to persuasion, increased with proximity.

#### *3.2.3.2 Need for Cognition and Personal Relevance*

Due to their failure to replicate the results obtained by Prentice, Gerrig and Bailis, Wheeler, Green and Brock considered a further possible factor to account for the difference between the results of the original study and their own initial results. They hypothesized that in order for subjects to engage in scrutiny of the story assertions they must perceive the presented narrative as personally relevant to them. They further hypothesized that whereas in Prentice,

Gerrig and Bailis home-school subjects perceived higher personal relevance than away-school subjects, this was not a factor in the case of the participants at OSU. The OSU students did not exhibit the same difference in scrutiny based on the difference in proximity. In order to test this hypothesis, the researchers conducted a further variation on Prentice, Gerrig and Bailis's initial experiments.

The experiment was conducted using 68 OSU undergraduates as subjects: twenty control subjects, nineteen home-school subjects and nineteen away-school subjects. The experimental design was the same as that for the initial adaptation with the following exceptions: (a) all subjects were asked to do the verbal SAT portion and NC scale test as their intervening tasks and (b) a new task was introduced after the questionnaire and recall test in order to determine the amount of scrutiny that each subject applied to the text. In this second task, subjects were asked to reread the story and circle all portions of the text that contained "false-sounding" information.<sup>101</sup> This task is referred to as Pinocchio circling.

As with the initial experiment, there was no significant difference between the scores of the home-school subjects and away-school subjects, nor was there any significant difference between HNC and LNC subjects. Further analysis was done to determine whether there was any interaction between proximity, need for cognition and scrutiny. This analysis was based on the results of the Pinocchio circling task.

In order to analyze this data, the researchers considered both how many lines of the story were circled by each subject as well as how many circles were made in total by each subject. Home-school subjects tended to both make more circles and circled more lines than did away-school subjects; although not all results were consistent with this trend. This suggests that scrutiny increases with proximity. Moreover, HNC subjects reliably both made more circles and circled more lines than did LNC subjects. This suggests that as need for cognition increases, so does scrutiny. What remains to be seen is the relationship between level of scrutiny and the degree to which a subject's belief is influenced by story information. Despite the differences in levels of scrutiny, there were no corresponding differences in levels of belief change between either home- and away-school subjects or between HNC and LNC subjects. What then could account for the discrepancy between the results of Prentice, Gerrig and Bailis and the results of Wheeler, Green and Brock?

---

<sup>101</sup> Wheeler, Green, and Brock, "Fictional Narratives Change Beliefs," 139.

### *3.2.3.3 Perceived Strength of Story Assertions and Levels of Scrutiny*

Wheeler, Green and Brock hypothesized that a further factor in these differences might be the difference between the perceived strength of the arguments made for the assertions in the story among subjects in Prentice, Gerrig and Bailis and the subjects in Wheeler, Green and Brock. That is, they considered whether what were intended to be weak arguments were taken as such by students at Yale and Princeton, but taken as strong arguments by students at OSU. If this was the case and there was a difference in level of perceived argument strength by subjects in the two studies, then it could be that arguments perceived as strong and subjected to greater scrutiny could lead to an increase in belief change. However, they hypothesized, if subjects perceived a mix between strong and weak arguments, then this difference might not be noticeable in the results, leading to no noticeable difference in belief change based on either proximity or need for cognition. This would explain why a difference in scrutiny did not seem to affect the degree of influence of story assertions on beliefs. In order to test this hypothesis, Wheeler, Green and Brock devised a further experiment designed to account for the perceived strength of arguments in support of story assertions.

The experiment was conducted using 201 OSU undergraduates as subjects: sixty-four control subjects, sixty-eight home-school subjects and sixty-nine away-school subjects. The experimental design was the same as that for the previous study with the following exceptions: (a) all actually true story assertions were removed, leaving only actually false-assertions and (b) a subset of 59 subjects was given stories containing false story assertions as well as four arguments which had been shown to be perceived as weak in previous studies. These arguments were meant to determine whether the current subjects perceived arguments to have the same strength as the previous subjects had. In order to test perceived argument strength, subjects were asked to record their thoughts and feelings after reading each argument. An argument was perceived as strong if it prompted the reader to record more positive thoughts and feelings than negative ones and perceived as weak if it prompted more negative responses than positive ones.

Despite the elimination of actually true story assertions and a replication of the results that showed more Pinocchio circling by both home-school subjects and HNC subjects than by away-school subjects and LNC subjects, indicating that they exercised greater scrutiny, the influence of story assertions on the subjects' beliefs remained the same as in the previous two experiments. Moreover, to account for the elimination of the actually true story assertions, Wheeler, Green and Brock further analyzed the results from the questionnaire administered to

the subjects of the supplemental experiment to only reflect degree of belief with respect to those story assertions that had been previously shown to be weak. That is, they analyzed the extent to which subjects' beliefs were influenced by only those story assertions generally perceived as weakly supported. Even with the elimination of those story assertions perceived as supported by strong arguments, there was no significant difference between the influence of story assertions on the beliefs among subjects. Thus, despite increased scrutiny of those story items perceived to be weakly supported, neither proximity, nor need for cognition affected the influence of the actually false story assertions on the subjects' beliefs. This suggests that the immediate influence of story assertions, whether actually true or false and whether perceived to be strongly or weakly supported is much more pervasive than Prentice et al. initially found. It seems that regardless of truth or perceived persuasiveness, Gilbert was correct in his conclusion that we cannot help but accept assertions presented in narrative, despite our ability to revise those beliefs upon later reflection and consideration.

What Gilbert's results demonstrate is that under experimental conditions, when given appropriate distractor tasks in the course of reading and comprehending a narrative, subjects tend to treat what they read as true. This leads him to conclude that treating what we read as true, which he associates with belief, is an essential part of the narrative comprehension process. However, this notion of belief is rather thin and fleeting. They are not considered opinions, but rather the result of comprehending narratives and then being asked about them without being given the opportunity to reflect upon them. This provides a snapshot of subjects' doxastic states immediately after comprehension has occurred. Wheeler, Green and Brock's results provide us with a justification for extrapolating these results to the case of fictional narratives as well. Indeed, they give us reason to believe that they hold independently of both proximity and need for cognition, which are both significantly mediators of levels of scrutiny. This suggests that the effect is really that of belief formation, which is a component part of the comprehension process.

This result is itself quite interesting in light of traditional philosophical views on belief acquisition. Our ability to acquire and act on beliefs as a result of reading and comprehending fictions, despite our knowledge that their contents tend not to be true, is an important example of the peculiarity of our responses to fictions. We want a theory of fiction that can account for these responses. In the following chapter, I will begin to develop such a theory by developing a schema for describing the nature of these responses. This schema will respect the data which the



experiments discussed above provide and capitalize on the conclusions that they suggest. But first, I will show that at the very least these results give us reason to deny the existence of the paradox of fiction, which is an instance of the subject matter paradox; the very problem for which my characterization of our responses to fictions is intended to account.

### 3.3 Philosophical Implications

There are a number of generalizations that we can draw regarding how we read and process information presented in fictional narratives. First, acceptance is the default response to comprehension. Rejection of information is a further process which requires additional reflection and can be prevented in cases where readers have no opportunity for the necessary reflection. Second, despite increased scrutiny in close proximity cases, readers tend to accept and incorporate fictional data regardless of proximity. This happens regardless of the perceived strength of the arguments provided and the extent to which individual readers tend to engage in effortful thought. This indicates that reading something is enough to incline a reader to assent to it, despite previous countervailing knowledge. So, the overarching message to be drawn from the data is that fictional narratives influence our beliefs in a way that is both automatic and independent of real world knowledge or reasoning.

One important philosophical issue related to the status of beliefs acquired via fiction is the so-called paradox of fiction. Simply put the paradox consists of an inconsistent triad according to which:

- 1) We are genuinely moved by fictions.
- 2) We know that what is portrayed in fictions is not actual.
- 3) We are only genuinely moved by what we believe is actual.<sup>102</sup>

It is generally accepted that this triad is inconsistent and so, while each premise is itself plausible, for any two of the premises to be true the third must be false. However, it is not clear that the three premises of the triad are logically incompatible. Rather, it seems that what is at issue here is a purported psychological incompatibility. Given the conclusions that can be drawn from the Gilbert's work together with the extrapolations that can be made to the fictional case as a result of the experimental results obtained by Wheeler, Green and Brock we are now in a position to account for the appearance of a paradox. In order to do this, it will be helpful to consider a number of strategies by which philosophers have attempted to defuse it.

---

<sup>102</sup> Radford, "How Can We Be Moved by the Fate of Anna Karenina?"

Walton argues that defusing the paradox requires the denial of (1).<sup>103</sup> This is one of his motivations for developing his theory of fictional emotion, described in section 2.6 above. Recall that on this view we do not experience genuine emotional responses to fiction. Rather, it is fictional that we experience them. This is because while we experience the psychophysical responses – or quasi-emotions – associated with these emotions, we do not have the corresponding psychological attitude which is a necessary constituent of emotion, namely belief. Rather, the psychological attitude that we take to the individuals and events that figure in fiction is make-believe. In so characterizing our affective responses to the fictional, Walton defuses the paradox.

Unfortunately, the empirical findings discussed above cast doubt on this solution for two important reasons. First, they call into question Walton's claim that the psychological attitudes that we take towards the fictional are not actually beliefs. Thus, they undermine his strategy of defusing the paradox on the grounds that our affective responses towards them do not constitute emotions. If belief is the necessary component for emotion that Walton believes our responses lack, and the empirical data suggests that this is a component that they do in fact possess, then further reasons must be given for distinguishing emotional responses to the fictional from emotional responses to the actual. Second, the findings discussed above call into question whether there is a genuine paradox which must be dissolved. In order to determine whether this is the case, it will be instructive to rehearse another one of the strategies for dissolving the paradox by denying (3), the claim that we are only genuinely moved by what we believe to be true.

The strategy under consideration is one advocated by Carroll, in what he calls the "thought theory."<sup>104</sup> On this view, we can entertain thoughts about the individuals and events that figure in fictions without actually believing them. Further, this entertainment of propositions regarding the fictional is sufficient for our experience of the associated emotions. Thus, belief, or even make-belief, with respect to these propositions is not necessary for having emotions. However, one of the key results of these studies is Gilbert and his collaborators' finding regarding our inability to comprehend statements while withholding belief. True, the results demonstrate that we can revise our beliefs upon reflection, but while in the process of

---

<sup>103</sup> Kendall L. Walton, "Fearing Fictions," *The Journal of Philosophy*, 75 (1978): 5-27.

<sup>104</sup> Noel Carroll, *The Philosophy of Horror; or Paradoxes of the Heart*, (New York, NY: Routledge, 1990): 88.

absorbing and assessing them, acceptance is the default. This suggests that whether our emotional responses occur before or after we are able to revise our beliefs in light of our knowledge that the contents of fiction are not actual, any account of our emotional responses to fiction cannot be developed on the hypothesis that we merely entertain their contents without forming beliefs, whether they are consistent with or contradictory to the contents of fiction. In the case of the former, we cannot deny premise (3), while in the case of the latter the paradox stands.

A final option for defusing the paradox is to deny (2) and claim that we do not know that the situations and individuals portrayed in fiction are not actual.<sup>105</sup> Traditionally, this has been attempted by arguing that while we are in the grip of the fiction we are under the illusion that the events recounted actually happened. On this view we are either under the illusion that the contents of fiction are real or we willingly suspend our disbelief. This view is generally seen as problematic because the very act of engaging with fictions as fictions presupposes knowledge of their falsehood. Thus we would be put in a position of ascribing contradictory belief states to the vast majority of readers who respond emotionally to what we read about in fictions. However, I do not think that this solution is entirely wrong, it merely requires a more nuanced treatment. On such a treatment, the explanation for why (2) is to be denied is not based on the idea that we are temporarily unaware of the reality of the situation, nor that we willingly ignore it. Rather, it is that this premise does not give us the whole story. Knowing that what we read is merely fictional does not prevent us from believing it and, if we have moving experiences in the process of reading, this seems like a sufficient distractor to prevent us from revising our attitudes and thus rejecting the emotion as unfounded.

For this reason, I believe that the best solution to the paradox is not to defuse it, but to deny it outright. Counterintuitive though it may seem, these three premises are not inconsistent; the paradox does not exist. We do feel genuine emotions towards the fictional and we do so because we cannot help but believe the contents of fiction in spite of our knowledge that they are merely fictional. Our knowledge of their fictional status does not override our emotional responses; rather the automaticity of initial belief in the contents of fictions overrides this

---

<sup>105</sup> Some discussions of this approach include Gregory Currie, *The Nature of Fiction*, (Cambridge: Cambridge University Press, 1990); and Aaron Meskin and Jonathan M. Weinberg, "Emotions, Fiction, and Cognitive Architecture," *The British Journal of Aesthetics*, 43 (2003): 18-34.

knowledge. What the data show is that these processes of belief and emotional response occur in a way that prevents this knowledge from coming into play before these emotions take hold.

In the next chapter, I will expand on this more nuanced treatment of the paradox. I will then use it as a starting point for a systematic account of a range of responses to fiction. In doing so, I will provide a schema for characterizing our responses to fiction in a way that adheres to the commitments of the No-Object view and so can provide a useful starting point for a solution to the subject-matter paradox. I believe that a proper characterization of these responses is necessary for any adequate philosophical theory of fiction.

## 4 The Curious Case of Double-Aspected Responses

In the first two chapters I discussed two very different philosophical approaches to fiction. Mark Sainsbury develops his views on fiction within a broader theory of reference. Beginning with a semantic theory that accommodates empty names, he then considers a pragmatic approach to interpreting fictional names when they are used in the context of discourse about fiction. Kendall Walton develops his views within a broader theory of representation. He gives an account of our engagement with the fictional, beginning with representational works in general and then focusing on the implications of his views for works of fiction in particular. Both Sainsbury and Walton adopt a No-Object view of fiction, accepting the fundamental assumptions that there is nothing that does not exist and that there are no objects that are the objects of fiction. However, they are divided on the matter of whether to accept Frege's Dismissal, which requires that a semantic theory of fiction only be an extension or adaptation of a more general theory.

In the third chapter, I gave an overview of recent work in experimental psychology on the impact of narratives in general and fiction in particular on our doxastic states. The result of this work has been to demonstrate that fictions have a very different influence on our beliefs than previously anticipated. I then made some suggestions regarding the philosophical implications of this work. In particular, I suggested that it calls the paradox of fiction into question. In what follows, I develop a double aspected approach to understanding our responses to fiction, which is supported by the psychological results. This approach adheres to the commitments of the No-Object view and provides useful guidance regarding our treatment of Frege's Dismissal. It also provides us with the tools that we need to properly evaluate the theories of both Walton and Sainsbury *as theories of fiction*. I will begin by defining double-aspectedness as a property.

Let DA mean "double aspected."<sup>106</sup> Then for some response, R, fiction, F, reader, A, and states, s and s\*, R is DA if:

- (1) R is a response to fiction, F,
- (2) A knows that F is a fiction,
- (3) R "involves" s,
- (4) s is cognitively compatible with (2), and

---

<sup>106</sup> John Woods and Jillian Isenberg, "Psychologizing the Semantics of Fiction."

- (5) there is another state  $s^*$  such that  $s$  and  $s^*$  are of some common phenomenological type,
- (6)  $s^*$  is cognitively incompatible with (2),
- (7)  $s$  and  $s^*$  are phenomenologically indistinguishable to  $A$  as members of the same type.

To say that  $R$  is a response to  $F$  is to say that  $F$  causes  $R$  in  $A$ . These responses include the epistemic, doxastic and affective. For example, readers are moved to tears when they read *Anna Karenina*.

To say that  $A$  knows that  $F$  is a fiction is to say that  $A$ , reading  $F$  with understanding, does not believe that the individuals and events that figure in  $F$  are actual. For example, the reader knows that Anna never existed and that the events described in Tolstoy's book never occurred.

To say that  $R$  "involves"  $s$  is to say that  $s$  is a necessary component of  $A$ 's experience of  $R$ . So, it is a component part of our double aspected response of being moved to tears by *Anna Karenina* that we are sad about what happens to her in the story, or that we believe that she loved Vronsky.

To say that  $s$  is cognitively compatible with  $A$  knowing that  $F$  is a fiction is to say that  $A$ 's experience of  $s$ , which takes  $F$  as its object is such that  $A$  need not believe that the individuals and events which figure in  $F$  are actual. For example, the sorrow that we experience while reading *Anna Karenina* does not entail a belief in the existence of Anna or Vronsky. Nor does it entail that we believe that any of the events that figure in the story actually occurred. There is a way of understanding what it is to be in this state that does not presuppose that the objects of the state are actual.

To say that  $s$  and  $s^*$  are of a type that has a common phenomenal character is to say that both  $s$  and  $s^*$  are experienced in ways that are typical of, for example, affective, doxastic or epistemic states. So, if  $s^*$  has the same phenomenal character as  $s$ , then if a reader's experience of  $s$  is typical of belief then so is his experience of  $s^*$ , if his experience of  $s$  is typical of sadness then so is his experience of  $s^*$ , etc.

To say that  $s^*$  is cognitively incompatible with  $A$  knowing that  $F$  is a fiction is to say that experiencing  $s^*$  as part of his response to fiction would cause cognitive dissonance in  $A$  in light of his knowledge that the individuals and events that figure in  $F$  are not actual. As such,  $A$  cannot have an experience of  $s^*$  and also not believe that the individuals and events that figure in  $F$  exist.

To say that  $s$  is phenomenologically indistinguishable from  $s^*$  to  $A$  is to say that, from  $A$ 's internal perspective, the experience of being in state  $s$  is no different from the experience of being in state  $s^*$ . From the perspective of the reader, being inconsolable about Anna's fate in a way that is cognitively compatible with knowing that Anna and Vronsky did not exist feels the same as being inconsolable about her fate in a way that is cognitively incompatible with knowing that they did not exist. When we experience either of these states, the phenomenological character of the state alone is insufficient to determine whether we are experiencing  $s$  or  $s^*$ . To us, they are both sorrow over Anna's fate.

This analysis raises a number of important questions, particularly with respect to the relationship between  $s$  and  $s^*$ . For one, how can both states be of a type with a common phenomenal character and yet one be cognitively compatible with (2) and the other not? For another, what are the objects of these respective states?

In order for a DA response to be an A-response to fiction, it must be a response to a text. As such, one state that is a component part of the response must be, for example, sadness, belief or desire that some state of affairs obtains in the text. This is cognitively compatible with the reader knowing that the text is a fiction, because the response is to the fictional text itself and not to any putative individuals or events which figure in it. So,  $s$  does not entail belief in the existence of such individuals or events.

However, our response to a fiction is also a response what is being presented in the text. When we read the text we cannot help but respond as though what we are reading is actual. Part of our response to the text is not yet sensitive to the knowledge that it is merely fictional. Simply by reading, we are put into a state of acceptance and any other states, such as  $s^*$ , which arise from that acceptance will be consistent with it. As a result, we might have both some  $s$  which takes the text as its object as well as some  $s^*$  which takes the individuals and events about whom the story apparently tells us. These states will each be a belief, emotion, desire or the like. But they will not be the same belief, emotion, desire or the like, because they are, in some sense at least, not caused by the same things nor do they have the same objects.

At the end of the preceding chapter I discussed a number of philosophical issues that are related to and can be better understood by bringing to bear the insights provided by the psychological literature on how fictional narratives influence actual beliefs. One such issue was the paradox of fiction. Recall that the paradox of fiction is an inconsistent triad, the premises of which state that (a) we are genuinely moved by fictions, (b) We know that what is portrayed in

fiction is not actual and (c) We are only genuinely moved only by what we believe is actual. The triad is taken to be inconsistent, as it is generally believed that no two members of the triad can be true without the third being false. Having rejected the going diagnoses of and solutions to this paradox, I briefly offered my own treatment.

My solution was to suggest that premise (b), which is generally taken to be the hardest of the three to deny, is in fact the one that leads to the appearance of a paradox. As I explained, it is not that (b) is false per se, but rather that it is presented so as to lead to the appearance of a paradox. This is because our beliefs about fictions and their contents are much more nuanced than they appear. It is not a simple matter of either believing or disbelieving. The contents of the beliefs which seem to be the same and so appear to cause inconsistent belief sets are in fact quite different and lead to beliefs that, when properly understood, are not at all inconsistent.

Assuming that what looks like a paradox is actually an instance of a DA response, we can redescribe the paradox of fiction in terms of the above analysis and consider what insight it gives us into what is really leading to our conflicting intuitions. Such a response to fiction is DA if

- (1) the response is caused by fiction, F,
- (2) the reader knows that the events and individuals that figure in F are not actual,
- (3) the response experienced involves an affective state which has a phenomenology typical of emotion,
- (4) being in that affective state is cognitively compatible with the knowledge that the objects and events that figure in the fiction are actual,
- (5) there is another affective state which also has a phenomenology typical of emotions,
- (6) the latter state is not cognitively compatible with the knowledge that the individuals and events that figure in the fiction are not actual, and
- (7) the two states are phenomenologically indistinguishable to the reader as emotions.

So, for example, the response is a response to *Anna Karenina*. The reader knows that *Anna Karenina* is a fiction. The response involves an emotional state – a state of sadness about Anna’s tragic fate, where sadness is a phenomenological type. This emotional state is cognitively compatible with the reader’s knowledge that *Anna Karenina* is a fiction. There is a state, sadness\*, of the same phenomenological type as sadness about Anna’s fate. Sadness\* is cognitively incompatible with the reader’s knowledge that *Anna Karenina* is a fiction. Sadness and sadness\* both have a phenomenology typical of sadness states.

How is the double aspected nature of these responses reflected by the paradox and how



can we use the insight that these responses are double aspected to understand not only why the appearance of inconsistency arises, but also why this appearance is misleading?

On this analysis, when a reader is devastated by Anna's fate, he has a response to the fiction, *Anna Karenina*. He has this response knowing that *Anna Karenina* is a fiction – that is, knowing that Anna, Vronsky and their fates are not actual. This response involves being in an emotional state; he might cry or lament or be sad or bemoan their fates. But this devastation – this sadness, crying, bemoaning and lamenting – is cognitively compatible with his knowledge that what he appears to be responding to and to whom it appears to have happened are not real. He is devastated, not by Anna and her circumstances, but by the novel, *Anna Karenina*, and the appearance that it carries for us of stating certain facts about the world. It is the appearance of the novel that leads him to respond to it as though the characters were real, to have the sorts of responses that he would have were they to exist. Distracted by the appearance of reality and the strong emotional response that he has toward it, the reader responds as though he believes that the events and individuals that figure in the novel are actual. Indeed, the psychological evidence presented in the preceding chapter strongly suggests that in that context, he does indeed believe what he reads about Anna, Vronsky and their ill-fated love affair. All this is separate from and cognitively incompatible with his knowledge that *Anna Karenina* is a work of fiction. Moreover, these affective responses are responses to the text on the one hand and the appearance that it presents on the other and the two are phenomenologically indistinguishable as sadness. Upon reading a text it feels the same to respond emotionally to a fiction as it would to respond emotionally to the same scenario had it actually occurred.

When considered as an interpretation and explanation of the traditional formulation of the paradox, we can see that while the premises of the paradox include much of the characterization provided by our analysis of double aspected emotional responses, there is key information missing. This information, while understandably overlooked, is necessary for a complete understanding of the interplay between belief and emotion at work. Without it, we have the appearance of a paradox:

- (a) we are genuinely moved by fictions,
- (b) we know that what is portrayed in fictions is not actual,
- (c) we are only genuinely moved by what we believe is actual.

Proposition (a) of the paradox reflects clauses (1), (2), (3), and (4). Proposition (b) of the paradox reflects clause (2). Proposition (c) of the paradox reflects clauses (1), (5), and (6).

What needs to be filled in is the fact that the two states are phenomenologically indistinguishable as sadness. Framed in this way both premise (b) and (c) are true because they entail different belief states for what appears to the reader and to ourselves as the same emotional state. Moreover, both are part of the same larger response, which is itself double aspected.

If this is in fact the case, then we would be well served by considering what the data yielded by work in empirical psychology suggests about the nature of these belief states and how certain beliefs about the contents of fictional narratives can be explained as instances of s states in the context of DA responses. The results of these studies, in which readers know that the narratives that they are reading are fictional, is that we cannot help but accept assertions presented in narratives, despite our ability to revise those beliefs upon later reflection and consideration. In the case of our emotional responses to fiction, this explains a great deal, as the initial acceptance induces in us emotional responses that seem, to many philosophers, puzzling in light of the fact that we possess the countervailing knowledge that the information to which we are responding does not reflect the actual state of affairs. This is a puzzle that is not only about emotions but also about beliefs, or rather about the interactions between the two. Hence, the paradox. The psychological data together with a correct analysis of what it is for a response to be double aspected illuminates the puzzle in a way that defuses the paradox. However, it may be that these responses are not only double aspected with respect to emotions, but also with respect to beliefs. In the case of the beliefs that we form in the process of comprehending narratives, which I used in the previous chapter to account for the doxastic element of the paradox, might providing a treatment of the data as instances of a DA response not bear similar fruit?

Recall that, in the studies cited in the preceding chapter, subjects were given a series of tasks in which they first read narratives that were clearly labeled as fiction, then were given distractor tasks which were unrelated to the content of the narratives, and finally were asked to complete questionnaires which reflected their attitudes about certain topics. The questions asked were intended to elicit responses about the actual world and not the fictional narratives. Moreover, the questions were about topics that were generally taken to be common knowledge. To take an example from one study, subjects were asked whether or not eating chocolate promoted weight loss. In each study, a significant number of subjects responded with answers consistent with the narrative, even when the information presented in the narrative was

inconsistent with what was true in the actual world. That is, they knew that the narratives were merely fictional and were asked questions about what they believed to be actually true. So, in the example from the study, after reading a narrative which asserted that chocolate promotes weight loss, subjects responded in ways consistent with that assertion. This was despite knowledge to the contrary and the knowledge that the narrative in which it was asserted was fictional. This knowledge involves countervailing beliefs. Such a response to fiction is DA if

- (1) the response is a response to a fictional narrative, F,
- (2) the reader knows that F is a fictional narrative,
- (3) the response involves a doxastic state with a phenomenology typical of belief,
- (4) the doxastic state is cognitively compatible with the knowledge that narrative is fictional,
- (5) there is another doxastic state which also has phenomenology typical of belief,
- (6) the latter doxastic state is cognitively incompatible with the knowledge that the narrative is fictional, and
- (7) the two doxastic states are phenomenologically indistinguishable to the reader as beliefs.

The response is a response to a narrative fiction presented in the context of a psychological experiment. The subjects know that the narrative is a fiction. The response involves a belief state – a belief about the effect that certain activities have on human health, such as that chocolate promotes weight loss or that direct exposure to sunlight prevents skin cancer. That belief is cognitively compatible with the knowledge that the narrative is a fiction. There is another state, belief\* of the same phenomenal type as the belief about the effect of these activities on human health. Belief\* is cognitively incompatible with the reader's knowledge that these narratives are fictional. Belief and belief\* are both phenomenologically typical of belief states.

On this analysis, when a subject responds to questions about the actual world in ways consistent with story assertions, he has this response knowing that the narrative is fictional. That is, knowing both that the individuals and events that figure in these narratives are not actual and that the assertions made within it do not purport to actual truth. As with the example from the study mentioned above, knowing that the claim that chocolate promotes weight loss does not purport to actual truth. This response involves being in a belief state – having a belief with certain content. The subject might assert, assent to or otherwise act in ways that are consistent

with the belief that chocolate promotes weight loss. However, having that belief – and associated behaviours that reflect it – is cognitively compatible with his knowledge that the narrative in question is fictional. That is, his awareness of the fact that there is no reason to, and in fact good reason not to, accept the information presented in the narrative as true in no way hinders him from forming this belief. Moreover, possessing both the knowledge that the narrative is fictional and the belief that is consistent with the content of the story, but at odds with his knowledge of the actual world, causes no cognitive dissonance.

Further, there are beliefs that are cognitively incompatible with the knowledge that the narratives are fictional. While we do not know anything about the anterior beliefs of the subjects, a few generic examples are ready at hand: the belief that the individuals and events that figure in the narrative are actual, the belief that the outcome of the story might change from one reading to another, or the belief that the claims made in the story purport to truth. The impulse might be to say that in the case of doxastic states consistent with claims presented in fictional contexts about the health effects of certain activities, what we are dealing with are belief-like states, since these attitudes are cognitively compatible with any number of beliefs entailed by the knowledge that the narratives are fictional. By contrast, one might claim that in the case of doxastic states regarding how fictional narratives relate to the actual world what we are dealing with are actual beliefs, since the attitudes are cognitively incompatible with the beliefs entailed by the knowledge that the narratives are fictional. However, from the perspective of the reader, in the context of the experience, there is no phenomenological difference between having what seems to be a belief in each case. Moreover, the behaviours exhibited by individuals in both of these doxastic states are indistinguishable from belief to both the subject and the examiner. In both cases subjects responded in ways that indicate belief, according to the criteria laid out by the experimenters and accepted by their peers.<sup>107</sup>

Given the generally accepted view in philosophy that we cannot believe what we know not to be actually true, how are we to explain these responses? That is, given that these experimental results are at odds with the received view regarding knowledge, belief and the internal consistency, or at least coherence of, our belief states, what story can be told to explain the phenomena exhibited by these results? It seems that we have two options at hand. The first

---

<sup>107</sup> Ruth Barcan Marcus, "Some Revisionary Proposals About Belief and Believing," *Philosophy and Phenomenological Research* 50 (1990): 133-153; and Neil van Leeuwen, "The Motivational Role of Belief," *Philosophical Papers* 38 (2009): 219-246.

option is conceptual analysis. That is, we can give an account according to which while the empirical results appear to show us that subjects form beliefs based on the contents of fiction, the states that these responses involve are merely belief-like. The second option is to consider the mechanism by which these responses occur – that is, to give a description of the mechanism that causes these seemingly inconsistent responses. The responses themselves are interesting and suggestive of the effects that reading fictional narratives have on our beliefs about the actual world. But in order to understand what these results can tell us about a broader range of responses to fiction, it would be helpful, for the present purposes, to consider why it is that logic seems to break down in different domains of reasoning. For example, we might ask why belief formation is not beholden to logical consistency.

One hypothesis that emerges from the literature in philosophy of mind and cognitive science relies on the notions of “informational encapsulation”<sup>108</sup> and “cognitive impenetrability.”<sup>109</sup> A system is informationally encapsulated just in case that system cannot, in the course of its processing, access information found in other components of the overall system. The only information accessible to the system are its proprietary inputs – the information being fed directly into that system – together with any information that might be stored in or part of the system itself. So, in the case of acquiring beliefs from fictional narratives, if the belief acquisition process is informationally encapsulated, then the anterior beliefs of the subject cannot be accessed by the responsible system in the course of the reading comprehension process. The only information accessible to the subject during this process is what is presented in the narrative. Further, an encapsulated system is impenetrable with respect to the other components of the overall system, since the processing of the system is insensitive to the information available elsewhere within the overall system. That is, the outcome of any processing that occurs is not affected by that information. So, if the process by which subjects comprehend fictional narratives is informationally encapsulated, then it is impenetrable with respect to their anterior beliefs, since the process is insensitive to them. That is, the outcome of the comprehension process is not affected by a subject’s anterior beliefs. In the fictional case, these outputs would be beliefs that result directly as a result of the comprehension and acceptance of the contents of the fictional narrative.

---

<sup>108</sup> J. Fodor, *Modularity of Mind*, (Cambridge, MA: MIT Press, 1983).

<sup>109</sup> Z. Pylyshyn, “Computation and Cognition: Issues in the Foundations of Cognitive science,” *Behavioural and Brain Sciences* 3 (1980): 111-132.

This cognitive impenetrability would then explain why, in the experimental case, acceptance is part and parcel of comprehension. When a subject reads a fictional narrative the process by which they comprehend it is informationally encapsulated and so insensitive to their anterior beliefs. The only information accessible to the system responsible for comprehension is what is presented by the fictional narrative. Returning to the example cited above, this includes the claim that chocolate promotes weight loss. Further, the comprehension process is impenetrable with respect to the subjects' anterior beliefs, since its processing is insensitive to them. As such, even if the reader already held beliefs about the metabolic effects of chocolate, these beliefs would not influence the comprehension and acceptance process. In this case the output of the process would be the belief that chocolate promotes weight loss.

However, the hypothesis of complete encapsulation and so cognitive impenetrability of the process by which attitudes are acquired as a result of reading fictional narratives cannot be correct. We know that even in the case of forming beliefs about the fictional – and not just beliefs about the actual world based on statements made in fictional narratives – certain anterior beliefs influence how we understand fictions and so what we believe about the individuals and events that figure in them. One way of putting this is that fiction is parasitic on reality.<sup>110</sup> That is, reading fiction with understanding requires that readers fill in the details using information contained in their existing set of beliefs about the actual world. We understand the story by accepting what is written by the author and filling in the details of what the actual world would be like under the counterfactual circumstances described in the text. So, the process by which we acquire any beliefs about the fictional in the process of comprehension must be penetrable by at least some of our anterior beliefs. In light of this, it must be the case that the process by which we acquire beliefs about the actual world in response to the claims made in these narratives must likewise be penetrable.

If this is the case, then there are still a couple of theoretical options on the table. The first states that when we read narratives, the process by which we comprehend them is selectively penetrable. That is, the process by which we comprehend narratives is penetrable to the extent that it is sensitive to information available elsewhere, such as the anterior beliefs necessary to fill in the details of the story, but not beliefs about the ontological status of the

---

<sup>110</sup> John Woods, "Fictions and Their Logic," in *Philosophy of Logic* edited by Dale Jacquette, vol. 5 of *The Handbook of the Philosophy of Science*, edited by Dov M. Gabbay, Paul Thagard and John Woods. Amsterdam: North-Holland, 2006; and David Lewis, "Truth in Fiction," *American Philosophical Quarterly*, 15 (1) (1978): 37-46.

objects and individuals who figure in it. On this model, the process by which we form attitudes as a result of reading fiction is distinct from that by which we do so as a result of interaction with and discourse about the actual world. Indeed, the attitude itself is distinct. In the case of reading fiction the system is selectively penetrable. Following Walton, the attitudes acquired during comprehension, whether about the fictional or the actual, are make-beliefs. By contrast, in the case of reading about or otherwise experiencing the actual world there is a separate process that is sensitive to the information stored in our existing set of beliefs, the outputs of which are also beliefs. This maintains the model on which comprehension and acceptance occur under one process, while rejection or integration occur separately and are delayed from it. On this model, both beliefs and make-beliefs are formed in the process of comprehension and the determination as to whether the new beliefs are to be accepted or rejected – as well as whether any make-beliefs will be imported as beliefs, remain make-beliefs or be rejected as neither – occurs after the initial process of attitude acquisition.<sup>111</sup>

This view maintains much of what is appealing about cognitive impenetrability as an explanation of our quick and easy comprehension and acceptance of information presented in fiction. Moreover, it recognizes that at least some information from outside the system is necessary for comprehension itself and so for our initial attitude acquisition, which occurs as part of the same process according to some of the experimental results. However, it is problematic because it undermines the key findings of later experiments. Specifically, it undermines the conclusion that the attitudes formed in the process of reading fictional narratives are beliefs. What the distraction tasks built into the experimental context allows us to see is a snapshot of what subjects believe about the actual world in these cases. These attitudes influence behaviours in ways consistent with belief and so the model is not consistent with the data.

While cognitive impenetrability and informational encapsulation are, at least initially, attractive and theoretically powerful, they are not the kind of explanation needed in this case. However, we still need an account of belief acquisition and fixation that explains how we can read a narrative knowing both that it is fictional and that the claims that it contains about the actual world do not purport to truth while we still form beliefs about the actual world as part of our comprehension of the narrative. If the process is not informationally encapsulated from, and

---

<sup>111</sup> Nichols, Shaun, "Imagining and Believing: The promise of a Single Code," *The Journal of Aesthetics and Art Criticism* 62 (2004): 129-139; and Nichols, Shaun. "Just the Imagination: Why Imagining Doesn't Behave Like Believing." *Mind and Language* 21 (2006): 459-474.

so is sensitive to information contained in other parts of the overall system, then we must explain what it is that allows us to form these beliefs in the first place. We must explain why the fact that they are inconsistent with a large and consistent set of our anterior beliefs – including, again, that the narrative is fictional and does not purport to truth – does not lead to their automatic rejection as candidates for belief. If there is no mechanism that prevents these beliefs from being acquired on the grounds of inconsistency and these beliefs are inconsistent with a wide range of entrenched well justified beliefs, then why are they not rejected out of hand?

The second, perhaps simpler but certainly less theoretically ambitious, option is to say the correct model is a single system of stable belief formation through competition. That is, accepting or rejecting beliefs newly formed as a result of reading fictional narratives is not a matter of determining whether those beliefs are consistent or inconsistent with our existing knowledge and belief set. Rather, as with learning as a result of reading testimony or having experiences, we form new beliefs quite easily. Some are rejected almost immediately because they are cognitively incompatible with a wide range of existing beliefs. Others are filtered out over time, because while initially appealing or useful in the short term, holding them does not serve us well epistemically. Still others stay with us over the long term and perhaps some long held beliefs are inconsistent with other long held beliefs. However, this inconsistency is not grounds for rejection. Indeed it indicates that beliefs can still be cognitively compatible despite being inconsistent. If this is the case, then it would explain why mistaken beliefs about the world that result from reading the narrative influence a subject's behavior in the way suggested by the data, in the short time after reading a fiction. This is especially the case in circumstances when the subject has no opportunity to reflect upon these newly acquired beliefs and in fact good reason to make use of them.

Recall that in the first chapter the No-Object view of fiction was introduced. This view is underwritten by two fundamental assumptions. The first is Parmenides Rule, which is a general metaphysical principle stating that there is nothing that does not exist. The second is the Non-Existence Postulate, which is particular to fiction and states that there are no objects which are the objects of fiction. Taken together, these assumptions commit us to the view that there are no objects that are the “objects of fiction.” Moreover, there is an important question regarding whether to adhere to Frege's Dismissal, which states that since fiction is of no scientific importance, it deserves no semantic notice beyond what is expressly fashioned for matters of science. This commits those who accept it to the view that any semantic account of fiction can



only be an extension or consequence of a more general semantic theory.

Further, recall that in the same chapter, I discussed two separate but importantly related projects undertaken by Sainsbury. In the first, *Reference Without Referents*, Sainsbury offers a unified semantic account of referring expressions. On this view, we ought not treat vacuous names and non-referring expressions as semantically different from those names and referring expressions that do in fact have referents. He includes fictional names in this category and maintains that “Semantics will recognize no special category of fictional sentences or fictional names. Everything will proceed just as for non-fictional regions of language.”<sup>112</sup> This is largely because he sees his semantic project as descriptive, rather than revisionary, reflecting only the “natural categories” of the language that it describes.<sup>113</sup>

According to *Reference Without Referents*, the intelligibility of empty names is the result of a name using practice having been established. As with the Kripkean model for non-empty names, the practice originates with a baptism. A causal chain is established with the introduction of the name into the language and carries on to each speaker who uses the name with the intention that it refer in the way established by the initial baptism. However, on Sainsbury’s model a successful baptism does not depend on a name being associated with its bearer. While there is a bearer with which the name is intended to be associated in the vast majority of cases, its existence is not required for the baptism to be successful. Rather, its success depends on the name being associated with the reference conditions for it. To understand a name is not to know who bears it, but rather to know what it would be for someone to bear it. Thus, when using a name referentially, speakers are engaged in a practice which associates a name with a reference condition and use it accordingly. Usually, this will involve picking out an individual that actually exists, but it need not. These name using practices allow empty names to be used without all empty names having the same referent, such as the empty set, because the reference conditions are different for each name. Thus, it is Sainsbury’s view that the fact that some names lack a referent does not warrant the recognition of a separate semantic category. Rather, it merely accounts for a difference in whether the reference conditions for a given name correspond to a bearer.

Sainsbury’s second project in *Fiction and Fictionalism* is to provide an account of the ontology of fiction. In doing so, he establishes himself as an “irrealist” with respect to fiction, in

---

<sup>112</sup> Sainsbury, *Reference Without Referents*, 202.

<sup>113</sup> Sainsbury, *Reference Without Referents*, 95.

essence declaring himself at one with the No-Object view discussed above. However, rather than attempt to extend the semantic project that he began in *Reference Without Referents*, he discusses the merits of a number of pragmatic strategies that can be useful in interpreting discourse about fiction. So, it seems that Sainsbury accepts Frege's Dismissal, incorporating any strictly semantic account of reference in or about fiction to a more general account of vacuous names and non-referring expressions which are themselves to be understood on the same model as those that have referents. Any further explanation of how names function in the fictional case is not a semantic question, but rather a pragmatic one. While interesting and useful, this is done in a rather ad hoc way that does not account for why these names are in fact intelligible and how we are to interpret discourse about fiction when it employs those same names. This is not to say that the project of interpreting discourse about fiction is beholden to any theory of negative existentials. Rather, if a theory of name-using practices has already been established, it seems likely that considering the impact of those practices on our discourse about fiction may bear interesting fruit.

Once again, a response to fiction is DA if

- (1) the response is a response to a fictional narrative, F,
- (2) the reader knows that F is a fictional narrative,
- (3) the response involves an intentional state with a phenomenology typical of reference,
- (4) the semantic state is cognitively compatible with the knowledge that narrative is fictional,
- (5) there is another intentional state which also has phenomenology typical of reference,
- (6) the latter intentional state is cognitively incompatible with the knowledge that the narrative is fictional,
- (7) the two intentional states are phenomenologically indistinguishable to the reader as acts of reference.

So, for example, the response is a response to the *Sherlock Holmes* stories. The reader knows that these stories are fictions. The response involves a referential state – using the name “Holmes” to make claims about one of the individuals who figures in the *Sherlock Holmes* stories. For example, “Holmes is a brilliant deducer.” That non-extensional referential use of the name “Holmes” is cognitively compatible with knowledge that the narrative is a fiction. There is another state, reference\* of the same phenomenal type as reference to Holmes. Reference\* is cognitively incompatible with the reader's knowledge that the *Sherlock Holmes* stories are fictional.

For example, “Holmes is a real person.” Reference and reference\* are both phenomenologically typical of referential states.

How is the double aspected nature of our these responses to fiction reflected by the disparate approaches that Sainsbury takes to the semantics of fictional names and his pragmatic approach to interpreting discourse about the fictional?

On this analysis, when a reader uses the name “Holmes” to talk about the main character in the *Sherlock Holmes* stories, he has this response knowing that the stories are fictional. That is, he knows that Holmes, Watson and Lestrade do not exist. This response involves being in an intentional state. He might make assertions, have thoughts or participate in discussions about Sherlock and his colleagues. However, engaging in such behaviours is cognitively compatible with his knowledge that the *Sherlock Holmes* stories are fictional. That is, the awareness that the stories are fictional does not give him a reason to abandon the discussion as meaningless.

Further, there are referential uses of “Holmes” that are cognitively incompatible with the knowledge that the stories in which he figures are fictional. However, when using the name “Holmes” to talk about either the stories or Holmes’s ontological status, it does not seem to the reader or his interlocutors that there is any deep distinction between how we use the name in either case. We take ourselves in both cases to be referring to Holmes, whether or not the claims that we make about him are correct. The difference is that there seems to us to be something wrong about talking about Holmes as an individual in the actual world that is not at all troubling when talking about him in the context of the fiction. This difference seems to be one about what we can say about him, rather than whether we are referring to him in one case but not in the other.

Taking up Sainsbury’s idea of a name using practice, we could say that the initial baptism of a fictional name takes place in the course of the author’s writing of the story. For example, in introducing the name “Holmes” Conan Doyle provides the reference conditions for it. The name is intelligible because we, as readers and speakers engage in the name using practice that this baptism established. It is through reading the stories that we know what it would be for “Holmes” to have a bearer. This name using practice further individuates the reference conditions for “Holmes” from those for the names of other fictional characters, such as “Watson” or “Clouseau.” It allows the reader to talk about Holmes in ways that distinguish him from them. Whether using these names in the context of discussing the stories or in talking about their relations to the actual world, the establishment of these name using practices is what

allows us to do so intelligibly. We must simply acknowledge that in some instances, such as “Holmes is a brilliant deducer,” the response is double aspected, compatible with but reflecting our knowledge that the stories are merely fictional. In others like “Holmes is a real person,” the response has a single aspect, that of reflecting our views about the status of the would-be bearer of the name in the actual world. True, as with almost all language use, some pragmatic interpretation is required, but this is not unique to fictional names. We not only interpret names in their context of use, but we also interpret the sentences in which they are used in the broader context of the sort of discourse for which they are employed. So, understanding name using practices as double aspected in fictional cases seems to obviate the need for ad hoc interpretive strategies.

If this is the case, then we may ask whether our understanding of these practices leads us to accept Frege’s Dismissal or to reject it. And indeed this is not entirely clear. It seems that Frege is right to say that fictional names are not semantically special. The name using practices for them are no different in kind from those for any other names, vacuous or not. However, if our responses to fiction are indeed double aspected, then the way in which we might say that there is no semantic difference between fictional names and any other names is at odds with the spirit in which Frege’s Dismissal was intended. This is because their reference conditions are special; they involve the fiction. The authors of fictions perform baptisms – and so establish reference conditions – in the course of their works. Knowing what the story tells us about the individuals that figure in it is what allows us to individuate the reference conditions for the various names used within the fiction thereby allowing us to use those names intelligibly in our discourse about it. When using the name “Holmes,” readers know that they are talking about a fiction and it is in using the name in ways that reflect an understanding of the stories that they properly engage in this name using practice. Knowing what it would be for the name “Holmes” to have a referent while also knowing that there is no one who is Holmes is exactly what makes responses involving the name referentially double aspected.

## Afterword

As stated previously, there is much to recommend the work done by both Walton and Sainsbury towards developing a philosophical theory of fiction. I believe that Walton and Sainsbury are both attempting to theorize about fiction in terms of our responses to it.

Sainsbury's discussions of both the semantic content of fictional names and the pragmatic interpretation which discourse about fiction in which these names figure reflect an important tension in our responses to fictional narratives. On the one hand, we know that these names lack referents and that any theory of their semantic contents must account for this. On the other hand, we use these names in ways that reflect the fact that we respond to them as though they do have referents. In fact, we respond to the contents of the narratives in which they figure as though we believe that they were true, using the names to individuate individuals, objects, and events that we know not to exist.

Walton's psychology first approach emphasizes the importance of accounting for the nature of our responses. In positing imagination as the mental state distinctive of our responses to fiction which is analogous to belief in response to reality, he also suggests an approach that is intended to explain how it is that we respond to fictions as though they are true despite our knowledge that they are not.

What both of these approaches lack is a well articulated, suitably general, common explanandum, the explanation of which might act as a starting point to bring the important insights provided and approaches developed together in the development of a unified approach for philosophical theorizing about fiction.

For Walton, what is needed is an explanandum which motivates his account of fiction in terms of imagination and make-belief. In attempting to first determine the correct mental state and then use it to characterize our response to the fictional as one of make-believe, he gives an account which is difficult to evaluate. However, by characterizing our responses to fiction as double-aspected and taking this to be what must be explained by an adequate theory of fiction, we are now in a better position to first determine whether make-believe is such a response and then to further determine whether imagination is the correct corresponding mental state. If these responses can be shown to account for the double-aspected nature of our responses so much the better, as it provides a reason for accepting a powerful theory. If not then this characterization of our responses gives us reason to reconsider Walton's account.

For Sainsbury, what is needed is an explanandum which can act as a starting point for a

systematic explanation of how the semantic and pragmatic concerns that he is addressing come together. This is because we need to understand why both the semantic and the pragmatic solutions he is providing are necessary to account for the seeming contradiction between our knowledge that fictional names lack referents and our discursive behaviour which treats them as though such referents exist. With this in place Sainsbury's seemingly disparate approaches come together to explain the very concerns that a characterization of our responses as double-aspected addresses. It is the characterization of these responses as double-aspected that accounts for the consistency between these two elements of his theory. Moreover, it might provide guidance regarding how to further develop his pragmatic account in a way that is less ad hoc and more systematic.

## Bibliography

- Adaval, R. & Wyer, R. S. Jr. "The Role of Narratives in Consumer Information Processing." *Journal of Consumer Psychology* 7 (1998): 207-245.
- Bach, Kent. "Failed Reference and Feigned Reference." *Grazer Philosophische Studien* 25 (1985/1986): 359-374.
- Bachrach, Jay E. "Fictional Objects in Literature and Mental Representations." *The British Journal of Aesthetics* 31 (1991): 134-139.
- Balaguer, Mark. *Platonism and Anti-Platonism in Mathematics*. Oxford: Oxford University Press, 1998.
- Biletzki, Anat and Matar, Anat. "Ludwig Wittgenstein." *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2010/entries/wittgenstein/>.
- Biocca, F. "The Evolution of Interactive Media: Toward 'Being There' in Non-Linear Narrative Worlds." In *Narrative impact: Social and cognitive foundations* edited by M.C. Green, J.J. Strange, & T.C. Brock. Mahwah, NJ: Lawrence Erlbaum Associates, 2002.
- Bogdan, Radu J. "Pretending as Imaginative Rehearsal for Cultural Conformity." *Journal of Cognition and Culture* 5 (1-2): 191-213,
- Braun, D. "Empty Names, Fictional Names, Mythical Names." *Noûs* 27 (2005): 596-631.
- Brock, S. "The Creationist Fiction: the Case Against Creationism about Fictional Characters." *Philosophical Review* 119 (2010): 337-64.
- Brock, S. "Fictionalism about Fictional Characters." *Noûs* 36 (2002): 1-21.
- Bruck, M. & Ceci, S.J. "The Suggestibility of Children's Memory." *Annual Review of Psychology* 50 (1999): 387-418.
- Bruner, J.S. *Actual Minds, Possible Worlds*. Cambridge, MA: Harvard University Press, 1986.
- Bruner, J. S. "What Is a Narrative Fact?" *The Annals of the American Academy of Political and Social Science* 560 (1998): 17-27.
- Busselle, R.W. & Greenberg, B.S. "The Nature of Television Realism Judgments: A Re-evaluation of Their Conceptualization and Measurement." *Mass Communication and Society* 3 (2000): 249-268.
- Cacioppo, J.T. Petty, R.E. & Kao, C.F. "The Efficient Assessment of Need for Cognition." *Journal of Personality Assessment* 48 (1984): 306-307.

- Caplan, B. "Creatures of Fiction, Myth, and Imagination." *The American Philosophical Quarterly* 41 (2004): 331-337.
- Carroll, Noel ed. *Theories of Art Today*. Madison: University of Wisconsin Press, 2000.
- Carroll, Noel. *The Philosophy of Horror; or Paradoxes of the Heart*. New York, NY: Routledge. 1990.
- Carruthers, Peter. "Review of 'Recreative Minds: Imagination in Philosophy and Psychology'" by Gregory Currie, Ian Ravenscroft. *Notre Dame Philosophical Reviews* 2003 (11)
- Cartwright, R. "Negative Existentials." *Journal of Philosophy* 57 (1960): 629-639
- Castañeda, H.N. *Thinking, Language, and Experience*. Minneapolis: University of Minnesota Press, 1989.
- Chaiken, S., Liberman, A. & Eagly, A.H. "Heuristic and Systematic Information Processing Within and Beyond the Persuasion Context." In *Unintended thought* edited by J.S. Uleman & J.A. Bargh. New York: Guilford, 1989.
- Clark, Ralph W. "Fictional Entities: Talking about Them and Having Feelings about Them." *Philosophical Studies* 38 (1980): 341-349.
- Coplan, Amy. "Empathic Engagement with Narrative Actions." *Journal of Aesthetics and Art Criticism* 62 (2004): 333-342.
- Currie, G. *The Nature of Fiction*. Cambridge: Cambridge University Press, 1990.
- Currie, Gregory. "On Being Fictional." *Journal of Aesthetics and Art Criticism* 55 (1997): 425-427.
- Currie, Gregory. "Fictional Names." *Australasian Journal of Philosophy* 66 (1988): 471-488.
- Currie, Gregory. "Characters and Contingency." *Dialectica* 57 (2003): 137-148.
- Currie, Gregory. "Imagination as Motivation." *Proceedings of the Aristotelian Society* 102 (2002): 201-16.
- Currie, Gregory. "Desire in Imagination." In *Conceivability and Possibility* edited by Tamar S. Gendler & John Hawthorne. Oxford: Oxford University Press, 2002.
- Currie, Gregory. "Imagination as Simulation: Aesthetics Meets Cognitive Science." In *Mental Simulation* edited by Martin Davies & Tony Stone. Oxford: Blackwell, 1995.
- Currie, Gregory. *The Nature of Fiction*. Cambridge: Cambridge University Press, 1990.
- Dauer, Francis W. "The Nature of Fictional Characters and the Referential Fallacy." *Journal of Aesthetics and Art Criticism* 53 (1995): 31-38.



- Deighton, J., Romer, D. & McQueen, J. "Using Drama to Persuade." *Journal of Consumer Research* 16 (1989): 335-343.
- DeFattore, J. "Controversial Narratives in the Schools: Content, Values, and Conflicting Viewpoints." In *Narrative Impact: Social and Cognitive Foundations* edited by M.C. Green, J.J. Strange, & T.C. Brock. Mahwah, NJ: Lawrence Erlbaum Associates, 2002.
- Doggett, Tyler and Egan, Andy. "How We Feel About Terrible, Non-existent Mafiosi." *Philosophy and Phenomenological Research* 84 (2011): 277-306.
- Donnellan, K. "Proper Names and Identifying Descriptions." In *The Semantics of Natural Language* edited by D. Davidson and G. Harman. Dordrecht: Reidel, 1972.
- Donnellan, Keith S. "Speaking of Nothing." In *Contemporary Research in Philosophical Logic and Linguistic Semantics*. Springer Netherlands, 1975
- Dorsch, Fabian. *The Unity of Imagining*. Munich: Ontos Verlag, 2012.
- Eagly, A.H. & Chaiken, S. *The Psychology of Attitudes*. New York: Academic Press, 1993.
- Egan, Andy. "Pretense for the Complete Idiom." *Noûs* 42 (2008): 381-409.
- Evans, G. *The Varieties of Reference*. Oxford: Clarendon Press, 1982.
- Everett, A. "Against Fictional Realism." *Journal of Philosophy* 102 (2005): 624-649.
- Fazio, R.H. & Zanna, M.P. "Direct Experience and Attitude-Behavior Consistency." In *Advances in Experimental Social Psychology* vol. 14 edited by L. Berkowitz. New York: Academic Press, 1981.
- Field, Hartry. *Science Without Numbers*. Princeton: Princeton University Press, 1980.
- Fodor, J. *Modularity of Mind*. Cambridge, MA: MIT Press, 1983.
- Frege, Gottlob. "Sense and Reference" in *The Philosophical Review* 57 (May, 1948): 209-230.
- Friend, Stacie. "Fictional Characters." *Philosophy Compass* 2 (2007): 141-156.
- Funkhouser, Eric & Spaulding, Shannon. "Imagination and Other Scripts." *Philosophical Studies* 143 (2009): 291-314.
- Gendler, Tamar Szabó. "Imaginative Contagion." *Metaphilosophy* 37 (2006): 183-203.
- Gendler, Tamar Szabó. "On the Relation Between Pretense and Belief." In *Imagination Philosophy and the Arts* edited by Matthew Kieran & Dominic McIver Lopes. London: Routledge, 2003.

- Gerrig, R.J. *Experiencing Narrative Worlds: On the Psychological Activities of Reading*. New Haven, CT: Yale University Press, 1993.
- Gerrig, R.J. & Prentice, D.A. "The Representation of Fictional Information." *Psychological Science* 2 (1991): 336-340.
- Gilbert, D.T. "How Mental Systems Believe." *American Psychologist* 46 (1991): 107-119.
- Gilbert, D.T., Krull, D.S. & Malone, P.S. "Unbelieving the Unbelievable: Some Problems in the Rejection of False Information." *Journal of Personality and Social Psychology* 59 (1990): 601-613.
- Gilbert, D. T. Tafarodi, R.W. & Malone, P.S. "You Can't Not Believe Everything You Read." *Journal of Personality and Social Psychology* 65 (1993): 221-233.
- Goodman, Nelson. *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis, IN: The Bobs-Merrill Company, 1968.
- Goldman, Alvin. "Imagination and Simulation in Audience Responses to Fiction." *The Architecture of Imagination* (2006): 41-56.
- Goldman, Alvin I. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press, 2006.
- Green, M.C., Garst, J., Brock, T.C. & L.J. Shrum. "The Power of Fiction: Determinants and Boundaries." In *The Psychology of Entertainment Media: Blurring the Lines Between Entertainment and Persuasion*. edited by Karen E. Dill, 161-176. Mahwah, NJ: Lawrence Erlbaum, 2004.
- Green, M.C. & Brock, T.C. "The Role of Transportation in the Persuasiveness of Public Narratives." *Journal of Personality and Social Psychology* 79 (2000): 401-421.
- Green, M.C. & Brock, T.C. "In the Mind's Eye: Transportation-Imagery Model of Narrative Persuasion." In *Narrative impact: Social and cognitive foundations* edited by M.C. Green, J.J. Strange, & T.C. Brock. Mahwah, NJ: Lawrence Erlbaum Associates 2002.
- Green, M.C., Strange, J.J. & Brock, T.C. *Narrative Impact: Social and Cognitive Foundations*. Mahwah, NJ: Lawrence Erlbaum Associates, 2002.
- Gregory, W.L., Cialdini, R B. & Carpenter, K.M. "Self-Relevant Scenarios as Mediators of Likelihood Estimates and Compliance: Does Imagining Make It So?" *Journal of Personality and Social Psychology* 43 (1982): 89-99.
- Gupta, Anil. "Definitions." *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), edited by Edward N. Zalta, URL = <http://plato.stanford.edu/archives/spr2009/entries/definitions/>.
- Harris, Paul L. *The Work of the Imagination*. Mississauga: Wiley-Blackwell, 2000.

- Harris, Paul L. *Mental Simulation*. Cambridge: Blackwell, 1995.
- Hass, R. G. "Effects of Source Characteristics on Cognitive Responses and Persuasion." In *Cognitive Responses in Persuasion* edited by R.E. Petty, T.M. Ostrom, & T.C. Brock. Hillsdale, NJ: Lawrence Erlbaum Associates, 1981.
- Heal, Jane. *Mind, Reason, and Imagination: Selected Essays in Philosophy of Mind and Language*. Cambridge: Cambridge University Press, 2003.
- Hendricks, Vincent & Symons, John. "Epistemic Logic." *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/entries/logic-epistemic/>.
- Hjort, Mette, and Sue Laver (eds.). *Emotion and the Arts*. Oxford: Oxford University Press, 1997.
- Hofweber, Thomas and Everett, Anthony (eds.) *Empty Names, Fiction, and the Puzzles of Non Existence*. Stanford: CSLI Publications, 2000.
- Howell, R. "Fictional Objects: How They Are and How They Aren't." *Poetics* 8 (1979): 129-77.
- Howell, R. "Fictional Realism and Its Discontents." In *Truth in Fiction* edited by F. Lihoreau. Munich: Ontos Verlag, 2010.
- Hylton, Peter. "Willard van Orman Quine." *The Stanford Encyclopedia of Philosophy* (Fall 2010 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2010/entries/quine/>.
- Hyslop, Alec. "Emotions and Fictional Characters." *Australasian Journal of Philosophy* 64 (1986): 289-297.
- Jacquette, D. *Meinongian Logic: the Semantics of Existence and Nonexistence*. Berlin and New York: de Gruyter, 1996.
- Jacquette, Dale. "Intentional Semantics and the Logic of Fiction." *The British Journal of Aesthetics* 29 (1989): 168-176.
- Johnson, M. K. "Individual and Cultural Reality Monitoring." *The Annals of the American Academy of Political and Social Science* 560 (1998): 179-193.
- Johnson, M.K., Hashtroudi, S. & Lindsay, D.S. "Source Monitoring." *Psychological Bulletin* 114 (1993): 3-28.
- Kalderon, Mark. *Moral Fictionalism*. Oxford: Clarendon Press, 2005.
- Kripke, Saul. *Naming and Necessity*. Oxford: Blackwell, 1980.
- Kripke, Saul A. *Reference and Existence: The John Locke Lectures*. Oxford: Oxford University Press, 2013.

- Kroon, F. "Disavowal Through Commitment' Theories of Negative Existentials." In *Empty Names, Fiction, and the Puzzles of Non-Existence* edited by Everett & Hofweber. Stanford: CSLI, 2000.
- Lamarque, P. & Olsen, S.H. *Truth, Fiction, and Literature*. Oxford: Clarendon Press, 1994.
- Lambert, Karel. "The Definition of E! in Free Logic." In *Abstracts: The International Congress for Logic, Methodology and Philosophy of Science*. Stanford: Stanford University Press. 1960.
- Langland-Hassan, Peter. "Pretense, Imagination, and Belief: the Single Attitude Theory." *Philosophical Studies* 159 (2012): 155-179.
- Leslie, Alan M. "Pretense and Representation: The Origins of 'Theory of Mind'." *Psychological review* 94 (1987): 412-426.
- Lewis, D. "Individuation by Acquaintance and by Stipulation." *Philosophical Review* 92 (1983): 3–32.
- Lewis, D. "Postscripts to 'Truth in Fiction?'" In *Philosophical Papers* vol. 1 edited by D. Lewis. Oxford: Oxford University Press, 1983.
- Lewis, David. *Philosophical Papers*, vol. 1. Oxford: Oxford University Press, 1983.
- Lewis, David. "Truth in Fiction." *American The Philosophical Quarterly*, 15 (1978): 37-46
- Liggins, David. "The Autism Objection to Pretence Theories." *The Philosophical Quarterly* 60 (2010): 764-782.
- Loftus, E. F. "When a Lie Becomes Memory's Truth: Memory Distortion after Exposure to Misinformation." *Current Directions in Psychological Science* 1 (1992): 121-123.
- Ludlow, Peter. "From Sherlock and Buffy to Klingon and Norrathian Platinum Pieces: Pretense, Contextualism, and the Myth of Fiction." *Philosophical Issues* 16 (2006): 162-183.
- Mandelker, Amy. *Framing Anna Karenina: Tolstoy, the Woman Question, and Novel*. Columbus: Ohio State UP, 1993.
- Marcus, Ruth Barcan. "Some Revisionary Proposals about Belief and Believing." *Philosophy and Phenomenological Research* 50 (1990): 133-153.
- Meinong, Alexius. "Über Gegenstandstheorie." In *Untersuchungen zur Gegenstandstheorie und Psychologie* Leipzig: Barth, 1904. Reprinted in *Gesammelte Abhandlungen (Gesamtausgabe bd. II)*, Graz: Akademische Druck- und Verlagsanstalt, 1971. Translation by I. Levi, D.B. Terrell, & R.M. Chisholm in *Realism and the Background of Phenomenology* edited by R. Chisholm. Glencoe: Free Press, 1960.

- Meskin, Aaron, and Weinberg, J. "Emotions, Fiction, and Cognitive Architecture." *The British Journal of Aesthetics* 43 (2003): 18-34.
- Murphy, S.T. "The Impact of Factual Versus Fictional Media Portrayals on Cultural Stereotypes." *The Annals of the American Academy of Political and Social Science* 560 (1998): 165-178.
- Nell, V. *Lost in a Book: The Psychology of Reading for Pleasure*. New Haven, CT: Yale University Press, 1988.
- Nichols, Shaun. "Just the Imagination: Why Imagining Doesn't Behave Like Believing." *Mind and Language* 21 (2006): 459-474.
- Nichols, Shaun. "Imagining and Believing: The Promise of a Single Code." *The Journal of Aesthetics and Art Criticism* 62 (2004): 129-139.
- Nichols, Shaun. "Imagination and the Puzzles of Iteration." *Analysis* 62 (2002): 182-87.
- Nichols, Shaun, ed. *The Architecture of the Imagination: New Essays on Pretence, Possibility, and Fiction*. Oxford: Oxford University Press, 2006.
- Nichols, Shaun and Stich, Stephen P. *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford: Oxford University Press, 2003.
- Nisbett, R. E., & Wilson, T. D. "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84 (1977): 231-259.
- Oatley, K. "Why Fiction May be Twice as True as Fact: Fiction as Cognitive and Emotional Simulation." *Review of General Psychology* 3 (1999): 101-117.
- Orenstein, A. "Fiction, Propositional Attitudes, and Some Truths about Falsehood." *Dialectica* 57 (2003): 177-190.
- Parsons, Terence. "A Meinongian Analysis of Fictional Objects." *Grazer Philosophische Studien* 1 (1975): 73-86.
- Parsons, Terence. *Nonexistent Objects*. New Haven: Yale University Press, 1980.
- Paskow, Alan. *The Paradoxes of Art: A Phenomenological Investigation*. Cambridge: Cambridge University Press, 2004.
- Pelletier, F. & Zalta, E. "How to Say Goodbye to the Third Man." *Noûs* 34 (2000): 165-202.
- Pennington, N. & Hastie, R. "Explanation-Based Decision Making: Effects of Memory Structure on Judgement." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14 (1988): 521-533.

- Perloff, R.M. "The Third-Person Effect: A Critical Review and Synthesis." *Media Psychology* 1 (1999): 353-378.
- Petty, R.E. & Cacioppo, J.T. "Issue Involvement Can Increase or Decrease Persuasion by Enhancing Message-Relevant Cognitive Responses." *Journal of Personality and Social Psychology* 37 (1979): 1915-1926.
- Petty, R.E. & Cacioppo, J.T. "Involvement and Persuasion: Tradition vs Integration." *Psychological Bulletin* 107 (1990): 367-374.
- Petty, R.E. & Cacioppo, J.T. "The Elaboration Likelihood Model of Persuasion." In *Advances in Experimental Social Psychology* Volume 19 edited by L. Berkowitz. Orlando, FL: Academic Press, 1986.
- Petty, R.E. & Wegener, D.T. "Attitude Change: Multiple Roles for Persuasion Variables." In *The Handbook of Social Psychology* 4th ed, vol 1 edited by D.T. Gilbert, S.T. Fiske, & G. Lindzey. Boston: McGraw-Hill, 1998.
- Petty, R.E.; Haugtvedt, C.P. & Smith, S.M. "Elaboration as a Determinant of Attitude Strength: Creating Attitudes That are Persistent, Resistant, and Predictive of Behavior." In *Attitude Strength: Antecedents and Consequences* edited by R.E. Petty & J.A. Krosnick. Mahwah, NJ: Lawrence Erlbaum Associates, 1995.
- Petty, R.E.; Ostrom, T. & Brock, T.C. *Cognitive Responses in Persuasion*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1981.
- Phillips, J. "Two Theories of Fictional Discourse." *American Philosophical Quarterly* 37 (2000): 107-119.
- Pollard, D.E.B. "On 'Talk 'About' Characters." *The British Journal of Aesthetics* 16 (1976): 367-369.
- Predelli, S. "Modal Monsters and Talk About Fiction." *Journal of Philosophical Logic* 37 (2008): 277-97.
- Prentice, D.A. & Gerrig, R.J. "Exploring the Boundary Between Fiction and Reality." In *Dual-Process Theories in Social Psychology* edited by S. Chaiken & Y. Trope. New York: Guilford, 1999.
- Prentice, D.A.; Gerrig, R.J. & Bailis, D.S. "What Readers Bring to the Processing of Fictional Texts." *Psychonomic Bulletin & Review* 5 (1997): 416-420.
- Prentice, D.A. and Gerrig, R.J. "Exploring the Boundary Between Fiction and Reality." In *Dual-Process Theories in Social Psychology* edited by S. Chaiken & Y. Trope. New York: Guilford, 1999.
- Priest, G. *Towards Non-Being: The Logic and Metaphysics of Intentionality*. Oxford: Clarendon Press, 2005.

- Pylyshyn, Z. "Computation and Cognition: Issues in the Foundations of Cognitive Science." *Behavioural and Brain Sciences* 3 (1980): 111-132.
- Quine, W.V.O. *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1953.
- Quine, W.V.O. "Two Dogmas of Empiricism." In *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1953.
- Radford, Colin. "How Can We Be Moved by the Fate of Anna Karenina?" *Proceedings of the Aristotelian Society Supplemental* 49 (1975): 67-80.
- Radford, Colin. "Neuroscience and Anna; A Reply to Glenn Hartz." *Philosophy* 75 (2000): 437-440.
- Ravencroft, Ian and Currie, Gregory. *Recreative Minds: Imagination in Philosophy and Psychology*. Oxford: Oxford University Press, 2002.
- Reeves, B. & Nass, C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. New York: Cambridge University Press, 1996.
- Robinson, Jenefer. *Deeper Than Reason: Emotion and Its Role in Literature, Music, and Art*. Oxford: Clarendon Press, 2005.
- Rorty, R. "Is There a Problem about Fictional Discourse?" in *Consequences of Pragmatism*. Brighton: The Harvester Press, 1982.
- Rosen, Gideon. "Modal Fictionalism." *Mind* 99 (1990): 327-54.
- Rosselli, F.; Skelly, J.J. & Mackie, D.M. "Processing Rational and Emotional Messages: The Cognitive and Affective Mediation of Persuasion." *Journal of Experimental Social Psychology* 31 (1995): 163-190.
- Russell, B. "On Denoting." *Mind* 14 (1905): 479-493.
- Saatela, Simo. "Fiction, Make-Believe and Quasi Emotions." *The British Journal of Aesthetics* 34 (1994): 25-34.
- Sainsbury, R.M. *Reference Without Referents*. New York: Oxford University Press, 2005,
- Sainsbury, R.M. *Fiction and Fictionalism*. London: Routledge, 2010.
- Salmon, N. "Nonexistence." *Noûs* 32 (1998): 277-319.
- Sauchelli, Andrea. "Fictional Objects, Non-Existence, and the Principle of Characterization." *Philosophical Studies* 159 (2012): 139-146.

- Schank, R.C. & Abelson, R.P. "Knowledge and Memory: The Real Story." In *Advances in Social Cognition* vol. 8 edited by R.S. Wyer, Jr. Hillsdale, NJ: Lawrence Erlbaum Associates, 1995.
- Schiffer, S. "Language-Created Language-Independent Entities." *Philosophical Topics* 24 (1996): 149-166.
- Schneider, B. & von Solodkoff, T. "In Defence of Fictional Realism." *The Philosophical Quarterly* 59 (2009): 138-149.
- Schnieder, Benjamin & von Solodkoff, Tatjana. "In Defence of Fictional Realism." *The Philosophical Quarterly* 59 (2009): 138-149.
- Schock, Rolf. *Logics Without Existence Assumptions*. Stockholm: Almqvist & Wiksell, 1968.
- Schroeder, Timothy & Matheson, Carl. "Imagination and Emotion" In *The Architecture of the Imagination: New Essays on Pretense, Possibility and Fiction* edited by Shaun Nichols. New York: Oxford University Press, 2006.
- Searle, John R. "The Logical Status of Fictional Discourse." *New Literary History* 6 (1975): 319-332.
- Shrum, L.J.; Wyer, R.S. Jr. & O'Guinn, T.C. "The Effects of Television Consumption on Social Perceptions: The Use of Priming Procedures to Investigate Psychological Processes." *Journal of Consumer Research* 24 (1998): 447-458.
- Slater, M. D. "Processing Social Information in Messages: Social Group Familiarity, Fiction vs Nonfiction, and Subsequent Beliefs." *Communication Research* 17 (1990): 327-343.
- Slater, M. D. "Entertainment Education and the Persuasive Impact of Narratives." In *Narrative Impact: Social and Cognitive Foundations* edited by M.C. Green, J.J. Strange, & T.C. Brock. Mahwah, NJ: Lawrence Erlbaum Associates, 2002.
- Smiley, Timothy. "Sense Without Denotation." *Analysis* 20 (1960): 125-135.
- Smuts, Aaron. "Art and Negative Affect." *Philosophy Compass* 4 (2009): 39-55.
- Stavroula Glezacos. "Truth and Reference in Fiction." In *Routledge Companion to the Philosophy of Language* edited by Gillian Russell and Delia Graff Fara. New York: Routledge, 2012.
- Stern, Laurent. "Fictional Characters, Places, and Events." *Philosophy and Phenomenological Research* 26 (1965): 202-215.
- Strange, J.J. "How Fictional Tales Wage Real-World Beliefs: Models and Mechanisms of Narrative Influence." In *Narrative Impact: Social and Cognitive Foundations* edited by M.C. Green, J.J. Strange, & T.C. Brock. Mahwah, NJ: Lawrence Erlbaum Associates, 2002.
- Strawson, Peter Frederick. "On Referring." *Mind* 59 (1950): 320-344



- Thomasson, A.L. *Fiction and Metaphysics*. Cambridge: Cambridge University Press, 1999.
- Van Inwagen, P. "Creatures of Fiction." *American The Philosophical Quarterly* 14 (1977): 299-308.
- Van Inwagen, P. "Quantification and Fictional Discourse." In *Empty Names, Fiction, and the Puzzles of Non-Existence* edited by Thomas Hofweber and Anthony Everett. Stanford: CSLI Publications, 2000.
- Van Leeuwen, Neil. "The Motivational Role of Belief." *Philosophical Papers* 38 (2009): 219-246.
- Walton, K.L. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge MA: Harvard University Press, 1990.
- Walton, Kendall L. "Fearing Fictions." *The Journal of Philosophy* 75 (1978): 5-27.
- Weinberg, Jonathan M. "Configuring the Cognitive Imagination. Imagining and Pretending." In *Mental Simulation In New Waves in Aesthetics* edited by Kathleen Stock & Katherine Thomson-Jones. Houndmills, Basingstoke: Palgrave Macmillan, 2008.
- Wheeler, S.C. Green, M.C. & Brock, T.C. "Fictional Narratives Change Beliefs: Replications of Prentice, Gerrig, and Bailis (1997) with Mixed Corroboration." *Psychonomic Bulletin & Review* 6 (1999): 136-141.
- Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G.E.M. Anscombe. Oxford: Basil Blackwell, 1953
- Woods, John and Isenberg, Jillian. "Psychologizing the Semantics of Fiction." *Methodos* 10 (2010). Accessed April 7, 2010. doi: 10.4000/methodos.2387.
- Woods, John. "Fictions and Their Logic". In *Philosophy of Logic* edited by Dale Jacquette, vol. 5 of *The Handbook of the Philosophy of Science*, edited by Dov M. Gabbay, Paul Thagard and John Woods. Amsterdam: North-Holland, 2006.
- Wyer, R.S. Adaval, R. & Colcombe, S.J. "Narrative-Based Representations of Social Knowledge: Their Construction and use in Comprehension, Memory, and Judgment." In *Advances in Experimental Social Psychology* vol. 34 edited by M.P. Zanna. New York: Academic Press, 2002.
- Yanal, R.J. *Paradoxes of Emotion and Fiction*. Philadelphia: Pennsylvania State University Press, 1999.
- Zeimbekis, John. "Thought Experiments and Mental Simulations." In *Thought Experiments in Methodological and Historical Contexts* edited by Katerina Ierodiakonou & Sophie Roux. Leiden: Brill, 2011.
- Zwaan, R. A. "Effect of Genre Expectations on Text Comprehension." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20 (1994): 920-933.