

Belief and Imagination

Jack Frank Davis

Department of Philosophy

UCL

Submitted for the degree of Doctor of Philosophy

Declaration

I, Jack Frank Davis confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Two assumptions are often made about the nature of the cognitive attitudes that allow us to engage with fiction and in pretence: the *uniformity* and the *non-doxastic* assumptions. The uniformity assumption tells us that both of these activities involve the same cognitive attitudes. The non-doxastic assumption tells us that these cognitive attitudes are not beliefs, but belief-like states that we can call *belief-like imaginings*. I will challenge both of these assumptions in this thesis. In the case of the uniformity assumption, I will draw a distinction between *voluntary* and *involuntary* imaginative counterparts. I will argue that if a belief-like counterpart is involved in our engagement in pretence, it will be a voluntary counterpart, whereas an involuntary one will have to be associated with our engagement with fiction. Against the non-doxastic assumption, I will argue that we can explain our engagement with these activities by introducing beliefs with distinct contents. In the case of pretence, I will suggest that the relevant beliefs are of the form ‘[I believe] I PRETEND that “*p*”’. In the case of fiction, I will argue that the relevant beliefs are of the form ‘I believe *p* [in the fiction]’. This will lead to us challenging the non-doxastic assumption on the grounds that belief-like imaginings are *unnecessary* for explaining how we are able to engage with fiction and in pretence. I will also offer some arguments for why belief-like imaginings might be *insufficient* for explaining how we are able to engage with fiction and in pretence. In particular, I will argue that belief-like imaginings do not do enough to explain how we recognise when someone else is engaging in pretence, and that they struggle to make sense of why our representations related to fiction and pretence exhibit what Walton calls ‘clustering’.

Acknowledgements

To start off, I would like to thank my primary supervisor, Prof. Lucy O'Brien for her insightful comments and advice while I was preparing this thesis. I would also like to thank my secondary supervisor, Dr. Rory Madden, for pushing me to improve the clarity of my writing throughout my graduate studies. I am also grateful to Dr. Stacie Friend for reading through a late stage draft of chapter 5 and offering penetrating criticism of some of my arguments. My arguments will likely still fail to convince her that belief-like imaginings are not involved in our engagement with fiction, but I hope I have at least managed to respond to some of her concerns.

Having been at UCL for 6 years, there are so many people I want to thank in the department it is hard to know where to begin. As such, I will focus my gratitude on two groups of people. Firstly, I would like to thank the attendees of several work in progress sessions where I presented elements of this thesis. Secondly, I would like to thank my undergraduate students over the last few years for always being engaging in philosophical discussions and for improving my understanding of various philosophical issues. It didn't take long for me to realise you never truly learn philosophy until you start to teach it to others.

Outside of UCL, I have benefited from attending various conference, where I presented a version of chapter 4 and several papers on aesthetics. Sadly, these aesthetics papers, on issues like the morality of videogames and the appearance of fictional worlds, have not found their way into the final thesis. Nevertheless, I would like to thank the audiences at all the conferences I have attended over the last few years for their insightful questions, which helped me to clarify various issues that have made their way into this thesis in one way or another.

On a more personal note, I would like to thank my parents and my little brother for supporting me throughout the writing of this thesis. Their support, both financial and emotional, has been invaluable.

Finally, I would like to offer a special thanks to my partner, Justine Goy. Not only has her support made the lonely and isolating experience of writing a philosophy PhD

bearable, she also read through an entire draft of this thesis to spot my many typographic errors. Sadly, I suspect that I have introduced many new typographic errors while making final edits and last-minute additions, but the fault for these remaining errors is entirely my own.

Table of Contents

DECLARATION.....	2
ABSTRACT	3
ACKNOWLEDGEMENTS	4
INTRODUCTION.....	8
CHAPTER 1: IMAGINATIVE COUNTERPARTS	14
INTRODUCTION	14
1.1 VARIETIES OF IMAGINATION	17
1.2 VOLUNTARY COUNTERPARTS & SIMULATION THEORY.....	25
1.3 VOLUNTARY IMAGINING SEEING.....	31
1.4 INVOLUNTARY COUNTERPARTS & FICTIONAL TRUTH.....	33
1.5 INVOLUNTARY IMAGINING SEEING.....	42
1.6 EMOTIONAL COUNTERPARTS	46
CONCLUSION.....	55
CHAPTER 2: DISTINCT ATTITUDES & DISTINCT CONTENTS	56
INTRODUCTION	56
2.1 A COUNTERPART TO BELIEF.....	60
2.2 A SINGLE ATTITUDE	62
2.3 A DISTINCT ATTITUDE.....	65
2.4 DISTINCT CONTENTS.....	75
2.4.1 <i>Counterfactual Belief</i>	75
2.4.2 <i>Leslie & Meta-representations</i>	80
2.4.3 <i>Beliefs about Fiction</i>	82
2.5 DISTINCT ATTITUDES OR DISTINCT CONTENTS?	84
CONCLUSION.....	92
CHAPTER 3: THE IMAGINATION BOX & PRETENCE.....	93
INTRODUCTION	93
3.1 CHILDHOOD PRETENCE	96
3.2 FEATURES OF PRETENCE	99
3.3 THE IMAGINATION BOX.....	103
3.4 DIFFERENCES BETWEEN BELIEF & BELIEF-LIKE IMAGININGS	111
3.5 DEALING WITH ASYMMETRIES.....	118
3.6 CLUSTERING REVISITED	122
3.7 THE ROLE OF BOXOLOGY.....	124
CONCLUSION.....	127
CHAPTER 4: PRETENCE RECOGNITION & MOTIVATION	128
INTRODUCTION	128
4.1 MOTIVATION & RECOGNITION	130
4.2 BROAD & NARROW	132
4.3 MENTALISTIC PRETENCE	137
4.4 DEFENDING BEHAVIOURAL RECOGNITION & MOTIVATION.....	145

4.5 HOW MUCH CAN PRETEND EXPLAIN?.....	155
CONCLUSION.....	163
CHAPTER 5: IMAGINATION & FICTION	164
INTRODUCTION.....	164
5.1 BELIEF-LIKE IMAGININGS & FICTION	166
5.2 COGNITIVE ENGAGEMENT.....	175
5.2.1 <i>Sophistication</i>	176
5.2.2 <i>Ontology</i>	179
5.2.3 <i>Fictions Are Weird</i>	183
5.3 FICTION & NON-FICTION.....	185
5.4 DISCOURSE PROCESSING	190
5.5 IMMERSION & AFFECT.....	196
CONCLUSION.....	207
CONCLUDING THOUGHTS	209
BIBLIOGRAPHY	211

Introduction

A young girl slowly toddles toward a miniature table, plastic teapot in hand. Upon reaching the table, she sets the teapot down carefully, placing it alongside several plastic cups that rest atop plastic saucers. She then turns to the teddy bears and dolls sitting around the table and proceeds to ask each of them if they'd like some tea. She indicates that they've answered in the affirmative, picks the teapot back up, and then walks around the table making as if to pour tea into their cups. Having served them all some 'tea', she sits back down and starts to act as if she were drinking from her own cup, then proceeds to hold conversations with the various toys. While this tea party is under way, her mother is sitting on a lounging chair nearby, Paul Beatty's Booker Prize-winning satire *The Sellout* in hand. In this book, she reads about events that have never happened and never will happen, and about characters who have never existed and never will exist, while her daughter continues to make as if to drink tea from an empty cup.

This thesis investigates part of what's going on in the head of individuals engaged in these two sorts of activities. One common view of the matter is that these two activities – childhood pretence and engaging with works of fiction – involve similar sorts of mental states. Most crucially, both are often thought to involve the imagination. Matravers notes in relation to a game where a child, Eric, pretends that nearby tree stumps are bears:

It is difficult to see how we could describe Eric's situation as anything but him imagining of himself that he is confronted by a bear. (Matravers 2013, pp. 11-12)

Everett says about engaging with fictions:

It cannot reasonably be denied that our engagement with fiction involves some sort of imaginative exercise in which we imagine the world of the fiction, its protagonists, the events which befall them, and so on. (Everett 2013, p.6)

If this is right, a crucial part of our explanation of what's going on in the head of the young girl described above is that she *imagines* she is having a tea party. In the *Sellout* example, her mother is similarly *imagining* the various goings-on described in the novel.

This leaves a question about what we mean when we say both activities involve the imagination. Walton (1990), who has been influential in drawing connections between

these two sorts of activities, argues that they both involve *make-believe*. In recent years, this has evolved into the idea that both these activities involve *propositional imagination*. A recent consensus has formed that propositional imaginings are what we can call an 'imaginative counterpart' to belief: they are *belief-like* imaginings. If we accept this, then when we say that pretence and fiction involve imagination, this amounts to maintaining that they both involve belief-like imaginings. The child imagines in a belief-like way that she is at a tea party; that she is pouring tea for teddy; that teddy has drunk his tea and needs a top-up; and so on. The reader imagines in a belief-like way that the main character of *The Sellout* has reintroduced segregation to his hometown in California and is currently at the Supreme Court accused of having violated the 13th Amendment by owning slaves.

I will develop my own contribution to this debate by calling this consensus view into question in two respects. Firstly, I will question whether belief-like imaginings are *necessary* and *sufficient* for explaining these two human activities, thus disputing what we can call the *non-doxastic* assumption. Secondly, I will examine whether a single mental state can explain imaginative activities like supposition, fiction and pretence, and so will question what we can call the *uniformity* assumption. I will argue that the cognitive attitudes involved in pretence and fiction are beliefs with *distinct contents*. I am going to propose that we can explain pretence by arguing that children can form meta-representational beliefs like '[I believe that] I PRETEND that "I am at a tea party"' (Leslie 1987, 1994). In the case of fiction, I am going to argue that we can make sense of our engagement by introducing beliefs subject to a fictional operator, such as 'I believe that Harry Potter is a wizard [in the fiction]' (Neill 1993, Tullmann 2016).

Before setting out my argument structure, it will be helpful to note three things that I will not be arguing for in this thesis. Firstly, I am not going to claim that there is no imaginative counterpart to belief. I think there may well be a sort of counterpart to belief related to supposing that *p* or entertaining the proposition that *p*. Furthermore, this counterpart has also been said to play a role in human activities like modal reasoning and third person mindreading. Since we will not have the space to consider these sorts

of activities in any great detail, I cannot adjudicate whether we need to make this association. What I will hope to demonstrate, however, is that we should be sceptical of an approach where, having decided that belief-like imaginings exist, we then proceed to put these belief-like imaginings to work in explaining various human activities without careful consideration of the explanatory work they are supposed to be performing. I will advocate for deciding whether to introduce counterparts on a case-by-case basis, particularly bearing in mind worries from philosophers like Kind (2013) about whether propositional imaginings are being made to bear a weight that cannot be borne by a single mental state.

Secondly, my argument shouldn't be read as being equivalent to saying that the imagination is simply not involved in fiction and pretence. I am happy to allow that other kinds of imagining – such as forming mental images and imagining experiences – may be important for understanding pretence and fiction. As such, when I refer to engagement with fiction and in pretence in this thesis, I will mostly mean this in the sense of our cognitive engagement. I am not presuming that settling on the right cognitive attitude will fully explain how we engage in these two activities.

Finally, there is an open question about what my project entails for the way that we talk about the attitudes involved in our engagement with fiction and in pretence. One way of responding to my claims would be to hold that we should cease to speak of *imagining that* in the context of these activities. Another way of responding to my claims would be to see them as suggesting that we can still legitimately speak in this way, but that we shouldn't take the notion of *imagining that* as implying that we are talking about a distinct attitude. On this approach, we should *reinterpret* what we mean when we talk about *imagining that*, at least in some instances. Not much turns on this, save the fact that if one wishes to continue to say that reading a novel or engaging in pretence involves the propositional imagination, I don't have to commit myself to saying this is a false claim, I only have to deny that this entails these activities involve a distinct belief-like attitude.

Having made these preliminary remarks, we can now set out the structure of this thesis.

In chapter 1, I will argue that there are two notions of an imaginative counterpart that can be discerned in recent philosophical discussions of the imagination. The first notion is of *voluntary* states which we can enter into at will, and which bear some resemblance to other mental states. The second is that of *involuntary* states which also resemble other mental states, but are entered into automatically when we engage with something that we recognise to be imaginary or fictional. Philosophical discussions of belief-like imaginings often cut across this distinction and my central claim in this chapter will be that introducing one sort of counterpart does not necessarily justify the introduction of the other sort of counterpart. In explaining why this is so, I will propose four principles that we can use to justify the introduction of both sorts of counterpart. This will help to shape our subsequent discussions of whether a counterpart to belief should be introduced, and whether one is involved in our engagement with fiction and in pretence.

In chapter 2, I will set out some different ways that we can explain our ability to propositionally imagine: the ‘just belief’, distinct attitude, distinct content, and distinct attitude and distinct content views. I will accept that we may well need to introduce some sort of counterpart to belief that can be associated with notions like *supposing*, but I will also introduce some initial reasons for doubting that a single counterpart can do all the explanatory work when it comes to explaining how we engage with fiction and in pretence. In this way, we will introduce our challenge to the uniformity assumption. This will set-up a standoff between distinct content and distinct attitude approaches when trying to make sense of our engagement with fiction and in pretence, which we will attempt to resolve in the remainder of the thesis.

In chapter 3, I will criticise the best developed argument for why we should associate a belief-like attitude with our engagement in pretence, which is set out by Nichols and Stich (2000, 2003). This discussion will reveal that their theory has difficulty explaining why belief-like imaginings are sometimes processed differently as compared to beliefs. This is because they place great emphasis on the idea that belief-like imaginings and beliefs share a ‘single code’ – which entails that they should be processed by our cognitive mechanisms in ‘much the same way’. In light of this commitment, I will argue that their view struggles to explain *clustering*, the fact that our representations about fiction and

pretence are associated with only a single fiction or a single episode of pretence (unlike our beliefs, which aim to provide a uniform picture of the world). This will reveal an initial way in which a distinct attitude is *insufficient* for explaining how we are able to engage in pretence and so will introduce our challenge to the non-doxastic assumption in the context of pretence.

In chapter 4, we will consider how children recognise that someone else is pretending, and how they recognise what sort of actions count as appropriate pretend actions when they engage in pretence. Nichols and Stich's account of pretence is a *behavioural* account, which means that they argue that young children understand pretence as a mere form of behaviour. I will argue that behavioural theories have difficulty explaining how children recognise the *content* of episodes of pretence, and this will reveal another way in which introducing a distinct attitude is *insufficient* for explaining pretence. I will contend that we should instead adopt a distinct content *mentalistic* view, where we hold that children have some understanding of the mental states that motivate pretence behaviour and that engaging in pretence involves forming beliefs that you are pretending. In light of this, I will suggest that it is *unnecessary* to introduce a belief-like attitude to explain how we engage in pretence. We will thus conclude our challenge to the non-doxastic assumption when it comes to explaining our engagement in pretence.

In chapter 5, we will turn to consider whether we should associate our engagement with fiction with a distinct belief-like attitude. We will begin by noting some differences between what an account of our engagement in pretence needs to explain as compared to our account of how we engage with fiction. This will reveal that we should reject the uniformity assumption, since, if a belief-like counterpart is involved in our engagement with fiction, it will be an involuntary counterpart. We will then consider several arguments for why we might need to introduce this sort of involuntary distinct attitude into our account of how engage with fiction and will find them all wanting. As such, I will argue that we can make sense of our engagement with fiction by proposing that we simply form beliefs like 'I believe p [in the fiction]'. This will entail that introducing a distinct belief-like attitude is *unnecessary* for explaining our engagement with fiction. I will also note that clustering remains unexplained by distinct attitude views, which suggests that

there is at least one respect in which a distinct belief-like attitude is *insufficient* for making sense of our engagement with fiction. In light of this, we will reject the non-doxastic assumption in the context of our engagement with fiction.

Chapter 1: Imaginative Counterparts

Introduction

In this chapter, we will seek to clarify the notion of an imaginative counterpart by considering what an imaginative counterpart is supposed to be, and what is entailed by saying that one mental state is a counterpart to another. We will also consider what sorts of mental states have counterparts, and when we are justified in maintaining that a mental state has a counterpart. In so doing, I will offer a novel theory of how to classify imaginative counterparts, and will set out some principles for when we should introduce them.

In section 1.1, I will give an overview of some recent discussions about the imagination, and I will bring out the distinction between propositional and non-propositional imaginings. I will note here that my arguments in this thesis will not rule out there being a role for non-propositional imaginings to play when it comes to explaining our engagement with fiction and in pretence.

In section 1.2, we will begin our discussion of imaginative counterparts. Currie and Ravenscroft (2002, p. 11) explain the general idea of imaginative counterparts by telling us that ‘when a form of imagining ... is X-like, ... it has state X as its *counterpart*’ (emphasis in original). However, I will argue that this is overly simplistic, since there are at least two notions of imaginative counterparts at play in contemporary philosophical discussions of the imagination. The first are mental states that resemble, in some sense, another mental state, but are under the control of the will. Since these states are under the control of the will, we can call them *voluntary counterparts*. This sort of notion of a counterpart can be associated with *simulation theory*. The second sort of counterpart is less familiar, and is what we can call an *involuntary counterpart*. This is where a counterpart still resembles another mental state, but the counterpart is not under the control of the will. It is hard to find clear examples of this sort of counterpart, but they can be associated with Walton’s (1990) influential theory of fictional truth, and in particular his notion of ‘quasi emotions’ (Walton 1978). This distinction is important,

since while both simulation theorists and Walton are widely understood as introducing counterparts, it has not been widely remarked that the sorts of counterparts they introduce can be distinguished.

To bring out the nature of these two counterparts, we will set out how simulation theorists justify the introduction of voluntary counterparts. I will contend that we can justify introducing these sorts of counterparts by arguing that one mental state resembles another mental state in introspective, functional and neurological respects but is under the control of the will. To give an example of when we might be justified in introducing this sort of counterpart, we will then consider whether seeing has a voluntary counterpart in section 1.3, and I will argue that we can justify this proposal by focusing on introspective, functional and neurological similarities between visualising and seeing.

In section 1.4, we will consider the nature of involuntary counterparts. I will bring out the motivation for introducing this sort of counterpart by setting out Walton's theory of fictional truth and how he allows for real feelings and actions to constitute fictional feelings and actions. We will then consider when we are justified in introducing this sort of counterpart in section 1.5 by examining whether seeing has an involuntary counterpart. I will suggest that in order to justify the introduction of an involuntary counterpart, we have to focus on specific human activities and consider whether either a) there are constraints on a genuine mental state being involved in this activity that necessitate the introduction of a counterpart; or b) there are puzzles that arise when making sense of this activity that are best explained by introducing a counterpart. Bearing these two issues in mind, I will suggest that we can potentially defend the introduction of an involuntary counterpart to seeing.

Finally, in section 1.6 we will consider whether the emotions have either voluntary or involuntary counterparts. I will argue that they do not have voluntary counterparts, but may well have involuntary counterparts if we accept Walton's arguments for introducing some 'quasi emotions'.

On the basis of these discussions, I will defend four claims about imaginative counterparts:

1. There are (at least) two kinds of putative imaginative counterpart i) a state that is under the control of the will and resembles a genuine mental state in introspective, functional or neurological ways; and ii) a state that resembles another mental state, which we automatically enter into when we engage with something we recognise as being fictional or imaginary, and has fictional content
2. Noting a mental state appears to have content that is recognised by the agent as fictional or imaginary is insufficient for introducing a counterpart of type (ii).
3. Instead, to justify the introduction of a counterpart of type (ii) we have to focus on specific human activities to see whether either a) constraints on a genuine mental state being involved in these activities necessitate the introduction of a counterpart; or b) puzzles that arise when making sense of these activities are best explained by introducing a counterpart.
4. Justifying the introduction of a counterpart of type (i) does not justify the introduction of a counterpart of type (ii) (and vice versa).

These initial points are important, since I will argue in chapters 2 and 5 that philosophers describe belief-like imaginings in a way that cuts across my distinction between these two sorts of counterpart, with the counterpart involved in pretence resembling my first notion of a counterpart and the one involved in fiction resembling my second notion of a counterpart. In light of this, we must offer separate justifications for associating a counterpart to belief with these two activities.

To begin, it will be helpful to say something about the more general notion of ‘imagination’ and to try and bring out what propositional imaginings are supposed to be.

1.1 Varieties of Imagination

The imagination is often regarded as being central to our mental life, as well as one of more interesting and romantic parts of it. For example, laypeople and philosophers alike have suggested that the imagination plays a central role in the generation of artworks, and in ordinary language we revere those individuals who deploy their imagination to create works of art.¹ We see some evidence for the importance of the imagination in the eyes of the folk in the way we use the word ‘imaginative’ to praise individuals: preferring the imaginative scientist or politician to the unimaginative. We also spend a lot of time engaging with what they might call the ‘imaginary’ worlds of novels, films, plays, video-games, and so on.

Somewhat less romantically, philosophers have argued that the imagination plays a role in allowing us to determine whether something is possible (Chalmers 2002 argues for a link between imagination, conceivability and possibility, for example) and in our counterfactual and hypothetical reasoning (Williamson 2005, 2016). This places the imagination at the heart of a good deal of our philosophical reasoning, suggesting for example that it plays a key role in how we develop and engage with thought experiments – such as Mary in her black and white room – and with hypothetical arguments more generally. In the Mary case, it looks like we imagine what it would be like to step into a world of colour for the first time, and this imagining tells us whether she would learn anything new.² In light of these sorts of connections, Williamson (2016, p. 115) suggests that the original evolutionary purpose of the imagination might have been to help our ancestors decide on what course of action to take when faced with obstacles like a river they were unsure whether they could safely cross.

¹ A recent summary of this link can be found in Wiltsher and Meskin (2016). I am not claiming here that all art creation necessarily involves the imagination. The point here is merely that it is often thought that art and imagination are intimately connected.

² A good introduction to debates about how much stock we should place in the imagination when thinking about what is possible can be found in the introduction to Gendler and Hawthorne (2002).

In making these introductory remarks, however, we see how diverse the uses of the word ‘imagination’ and its cognates like ‘imagines’ are. This reveals a difficulty with trying to explain what the imagination is, namely that it is not clear what sort of account we should seek. O’Shaughnessy (2000, p. 339-340) suggests that taking there to be a single mental activity of imagining is to ‘assume too much’ since this amounts to suggesting that ‘there is some one thing that is the phenomenon of Imagining.’ For example, we could argue that to understand the imagination we need to understand a certain sort of mental state, activity, attitude, event or process of imagining.³ An explanation of what the imagination is would then take the form of an account of this *sui generis* mental state.

One proposal we might be tempted by is the idea that imagining involves forming mental images. We could argue that to imagine something is to form a mental image of that something, and that explaining the nature of the imagination boils down to explaining the nature of this ability to form mental images. There has been a longstanding tendency in philosophy to associate the imagination with the formation of mental imagery.⁴ In Aristotle’s *De Anima* there is a relatively short discussion of the imagination, where he can be read as associating imagination simpliciter with the formation of mental images, proposing that imagination is ‘that in virtue of which an image occurs in us’ (*De Anima* iii 3, 428aa1-2).

Another historical discussion of the imagination that associates it with mental imagery is found in Descartes’ *Meditations*. This focus is demonstrated by Descartes’ argument that we can understand the idea of a chiliagon (a 1000-sided shape) but cannot imagine one:

[i]f I want to think of a chiliagon, although I understand that it is a figure consisting of a thousand sides just as well as I understand the triangle to be a three-sided figure, I do not in the same way imagine the thousand sides or see them as if they were

³ This list is adapted from Kind (2016, p. 2).

⁴ White (1990, Ch. 1-7) offers a good historical analysis of this link, finding it in Aristotle, Descartes, Spinoza, Berkeley and Hume among others.

present before me ... But suppose I am dealing with a pentagon: I can of course understand the figure of a pentagon, just as I can the figure of a chiliagon, without the help of the imagination; but I can also imagine a pentagon, by applying the mind's eye to its five sides and the area contained within them. (Descartes 1642/1996 p. 50-51)

In this quotation, it looks like Descartes is thinking about imagination as essentially involving images, and claiming that we cannot form an image of a chiliagon. Part of the thought here might be that any image we form of a chiliagon cannot be distinct from, say, the image we would produce to represent a 999-sided shape or a 998-sided shape.⁵

If we were to endorse this sort of view, then all human activities that involve the imagination will therefore involve the formation of mental imagery. If we return to some of my initial examples, reading a novel would necessarily involve forming mental images of the characters and events described therein; figuring out whether you can cross a river would necessarily involve forming an image of your attempted (or successful) crossing; and deciding whether Mary learns something outside her black and white room would necessarily involve forming an image of a girl leaving a black and white room.

However, this does not look like the right way of capturing what we mean when we say these sorts of activities involve the imagination. One could read a novel without forming any mental images whatsoever, and if we think about visual fictions such as film or comics, engaging with these art forms doesn't appear to involve the formation of mental images, since the fictional events are already depicted for us on screen or on the page. Likewise, we might think that we can decide what conclusion to draw from the Mary case without forming any images, or that we can decide whether we can cross a troublesome river without them.

The difficulties that arise if we attempt to reduce the imagination to one sort of mental state become even more acute once we recognise the wide range of uses of the word 'imagination' and its cognates in ordinary language. Strawson notes that:

⁵ McGinn (2005, p. 129) points out we can also question whether we could distinguish an actual 1000-sided shape from a 999-sided shape via ordinary visual perception.

The uses, and applications, of the terms ‘image’, ‘imagine’, ‘imagination’, and so forth make up a very diverse and scattered family. Even this image of a family seems too definite. It would be a matter of more than difficulty to identify and list the family's members, let alone their relations of parenthood and cousinhood (Strawson 1970, p. 31)

The most extensive recent taxonomy of the different ways we use ‘imagination’ and its cognates comes from Stevenson (2002), who picks out twelve different conceptions of imagination in philosophy and ordinary language, ranging from ‘the ability to entertain mental images’; to ‘the ability to think of something not presently perceived, but spatio-temporally real’; to ‘the ability to appreciate things that are expressive or revelatory of the meaning of human life.’ In spite of the wide variety of conceptions he picks out, Stevenson tells us his list is by no means exhaustive. Regardless, it is unlikely that any one mental state can explain these initial twelve conceptions. For example, the sort of mental state involved in the production of mental imagery is likely not the same as the one involved when we appreciate things ‘that are expressive ... of the meaning of human life’.

In light of the challenges that attend to offering a general theory of imagination, recent philosophical discussions have tended to proceed by making some relevant distinctions between different aspects of the imagination rather than trying to offer a broader theory of imagination.⁶ One notable example of this tendency is found in Walton’s *Mimesis as Make-Believe*, perhaps the most influential work on the imagination in recent years. Walton concludes his starting discussion of the nature of imagination by asking:

What is it to imagine? We have examined a number of dimensions along which imaginings can vary; shouldn't we now spell out what they have in common? —Yes, if we can. But I can't. (Walton 1990, p. 19),

⁶ A potential exception is White who attempts to develop a theory of imagination that covers all uses of ‘imagination’ and its cognates in ordinary language and argues this should be an important goal of any successful theory of the imagination. As he puts the point: ‘I submit that any acceptable theory of imagination must account for all ... common uses [of imagination and its cognates].’ (White 1990, p. 85)

Walton (1990, pp. 13-19) instead proceeds by making some relevant distinctions for his purposes, such as between *occurrent* and *dispositional* imaginings and *solitary* and *social* imaginings. For the purposes of this thesis, we will proceed by marking a distinction between *propositional* and *non-propositional* imaginings.

A good way to introduce this distinction is by bringing out why philosophers have moved away from the view that imagining essentially involves the formation of mental images.⁷ One reason to be sceptical of this approach is it looks like imaginings can take different *forms*, and not all of these forms are essential imagistic in nature. For example, suppose someone asks you to imagine being on a beach in Spain. One way to do this would be to form a mental image of yourself sitting on a beach in Spain, perhaps with a cocktail by your side and a book resting at your feet. But you could also imagine this without forming mental imagery, you could imagine *that you are on a beach in Spain*. In other words, you could imagine this *propositionally*, where this looks like it involves taking an attitude towards a particular propositional content. These sorts of imaginings are thus often called ‘propositional’ imaginings.⁸ That being said, this does not mean that you think of yourself as merely taking an attitude towards a proposition when you imagine in this way. Instead, you are doing something like representing a state of affairs in a manner that isn’t imagistic.

⁷ Kind (2001) is a partial exception to this recent move away from imagery, and argues imagery is central to the imagination.

⁸ In some recent discussions, these sorts of imaginings have also been called ‘attitude’ imaginings (Van Leeuwen 2013, 2014). I will be using propositional imaginings in this dissertation because I prefer the locution ‘propositionally imagines’ to ‘attitudinally imagines’ and because this is more commonly used by the philosophers we will be primarily discussing. The notion of attitude imaginings also leaves it ambiguous whether the content of the imagining in question is propositional or non-propositional. Grzankowski (2012) argues that we should pay more attention to ‘non-propositional’ attitudes, and if we make room for these sorts of attitudes in ordinary cognition it might also be possible for us to form non-propositional attitude imaginings.

This brings out the difference between propositional imaginings and one kind of non-propositional imaginings, namely *objectual* imaginings (Yablo 1993, p. 27). A propositional imagining has a *proposition* as its content and involves a that-clause, whereas an objectual imagining has an *object* as its content. To make this a bit more explicit, the claim is that propositional imagining involves *imagining that p*, where ‘*p*’ is a proposition, whereas objectual imagining involves imagining *O*, where ‘*O*’ is an object. We can *imagine that there is a tree outside the window* and we can *imagine a tree outside the window*.⁹

With this distinction in hand, we can respond to Descartes’ scepticism about our ability to imagine a 1000-sided shape. We can argue that whilst we might not be able to objectually imagine a chiliagon (in the sense of forming a mental image of one that is suitably distinguishable from a 999-sided shape) we can nonetheless propositionally imagine one. We can imagine that *there is a 1000-sided shape*.

That being said, it is debateable whether objectual imagining something always involves forming a visual image of that something. For example, it looks like we can imagine the smell of a rose or the sounds of a symphony. One might prefer to describe these as something like *property* imaginings, but, nonetheless, we might be able to form objectual imaginings with an auditory or olfactory character. That is to say, that we are not necessarily limited to solely being able to form objectual imaginings with a visual character (Currie and Ravenscroft 2002, pp. 11-12). Furthermore, one might wish to leave space in one’s account of objectual imaginings for the possibility of imagining abstract concepts,

⁹ There is a sense in which we can say a proposition is the object of a propositional imagining. Crane (2012, p. 419) for examples draws on Prior (1971) and notes one sense of something being an object of thought: ‘When we believe or judge, what we believe or judge is sometimes called the object of our thought; normally these things are called ‘propositions’ and states of thinking them are now called propositional ‘attitudes’’. Another way to mark the distinction that I am making here would be to say that propositional imaginings have a proposition of some sort as their object, whereas objectual imaginings have some sort of non-propositional object as their content.

such as justice and fairness, which will be hard to make sense of in terms of forming mental images.¹⁰

Although, we can distinguish between these two different kinds of imaginings, one might wonder if both stem from a common faculty. For example, McGinn suggests there is a common faculty that produces propositional imaginings and mental images. He begins by considering why we use imagination to refer to what look like two different kinds of mental activity:

Once this distinction is appreciated, it becomes a question why we use the word “imagination” in such an inclusive way: Is it ambiguous between sensory and [propositional] imagination? Does it just lump together unrelated mental operations? This is a reasonable question, but I think it has an answer—namely, the two types of imagination employ different elements but involve the *same faculty*. Sensory imagination employs sensory elements, much as perception does—though, as we have extensively seen, these elements must not be conflated. [Propositional] imagination employs conceptual elements, much as thinking does: these elements are not intrinsically modality-specific, and combine to form propositional contents. What is in common is the general faculty that works on these elements—the imagination. It is essentially a creative combinatorial faculty that differs from perception and from belief (as we shall see more fully in a moment). My point is just that the same faculty may operate on distinct types of elements; the identity consists in the same type of operation being performed by a structurally uniform faculty. (McGinn 2005, pp. 129-130)

I will not take a stand on this faculty claim in this thesis. That being said, it strikes me that it is unclear exactly what is being claimed when we say these different kinds of imagination are produced by a single faculty. For example, what does it mean for the same faculty to employ ‘different elements’? There may be a compelling way of clarifying this faculty approach, but whether or not there is a faculty of imagination will not have any direct implications for my arguments in this thesis.¹¹

¹⁰ Indeed, some call objectual imaginings ‘imagistic’ imaginings, a locution I’ve avoided to allow for this possibility. Kind (2016, footnote 7) independently makes a similar point about the difference between ‘objectual’ and ‘imagistic’.

¹¹ The idea that there is a faculty of imagination occasionally receives implicit support in recent discussions. For example, Nichols (2004, p. 129) opens a discussion of how imagination relates to belief by saying ‘The imagination has always been one of the darker

A second important category of non-propositional imaginings is called ‘active’ imaginings (Gendler 2011). The motivation for introducing this class of imaginings is summed up by Walton (1997/2015, p. 274) when he notes that in addition to the sorts of imaginings we have talked about so far, we can also imagine ‘[d]oing things, *experiencing* things, *feeling* in certain ways’ (emphasis in original). This also relates to the idea of sympathetic imagination introduced by Nagel (1974), which he uses to refer to our ability to imagine experiences.¹² On the basis of the arguments that will follow in this thesis, we will not rule out there being an important role to play for both objectual and active/sympathetic imaginings in our engagement with fiction, and perhaps also in pretence.

Having made these initial remarks about the imagination and having attempted to mark out what philosophers have in mind when they refer to propositional imaginings, we should now turn to consider the notion of an imaginative counterpart in more detail. In recent years, a consensus view has emerged that holds propositional imaginings are best thought of as counterparts to belief. We can call these sorts of imaginings *belief-like imaginings*. Before we assess the arguments for this view, we should first formulate some general principles for when we should introduce imaginative counterparts and we should try to get clearer on exactly what imaginative counterparts are supposed to be.

faculties of the human mind’ (emphasis mine). Weinberg and Meskin (2006b, p. 176) note ‘The imagination, after all, is a *faculty* that creatures like us contingently possess and deploy in distinctive ways.’ (emphasis mine)

¹² This raises a question about whether this form of imagination can subsume propositional and objectual imaginings. We could argue that all mental imagery formation is a case of imagining seeing and that the same should be said for imagining smelling, or imagining hearing, and that as such objectual imaginings reduce to imagining various different kinds of experiences. Likewise, we could argue that propositional imaginings reduce to imagining believing.

1.2 Voluntary Counterparts & Simulation Theory

The first notion of a counterpart we can find in philosophical discussions amounts to something like the following:

VC: A mental state that is under the control of the will and resembles another mental state in introspective, functional and neurological ways.

Perhaps the most detailed accounts of how these sorts of counterparts behave and how we can defend their introduction can be found in the writings of simulation theorists. It will thus be helpful to begin our discussion of voluntary counterparts by reflecting on what simulation theory tells us about the nature of voluntary counterparts.

Simulation theory is a theory that seeks to explain how we understand other minds¹³, and was first developed by Gordon (1986), Heal (1987) and Goldman (1989). The capacities that allow us to understand other minds are typically called our *mindreading* capacities. This refers to a wide range of human capacities. Ravenscroft (2017) suggests that the relevant capacities include:

1. Our capacity to *predict* people's future behaviour,
2. Our capacity to *ascribe* mental states to other people and to ourselves,
3. Our capacity to use these ascriptions to *explain* people's past behaviour.¹⁴

¹³ It can also be argued that mindreading is the process we use to read our own minds (Carruthers 2009a, 2009b).

¹⁴ Barlassina and Gordon (2017), in a recent review article of simulation theory, summarise our mindreading abilities in a similar way, though allowing for some further capacities, suggesting that 'The capacity for "mindreading" is ... the capacity to represent, reason about, and respond to others' mental states.' Ravenscroft (2017) is also aware that a theory of mindreading may well need to explain things like our ability to 'reason about, and respond to others' mental states', noting: 'In addition to attributing mental states and predicting and explaining behavior, there is a wide range of closely related activities. ... we not only seek to predict and explain people's behavior, we also seek to predict and explain their mental states. ... we speculate about, discuss, recall and evaluate both people's mental states and their behavior. We also speculate about, discuss, recall and evaluate people's

There have been various theories advanced to explain how we are able to ascribe mental states to others and predict and explain behaviour using these ascriptions. The two most prominent ways of explaining these abilities are the *theory-theory* and the aforementioned *simulation theory*.¹⁵

Theory-theorists argue that we can ascribe mental states to others, as well as predict and explain their behaviour, because we possess a *theory of mind*, which contains information about how desires and beliefs (and perhaps other mental states) interact to motivate behaviour. This theory of mind is usually said to be *tacit*, which means it is not immediately clear what sorts of propositions are supposed to make up the theory. Suggested candidates are often elaborations of truisms, such as ‘people will perform actions to get what they want, *ceteris paribus*’, or ‘people believe what they see *ceteris paribus*’ (Botterill and Carruthers 1999, p. 77-78).¹⁶

dispositions to behave in certain ways and to have certain mental states; that is, we consider their character traits.’ Ravenscroft goes on to point out these activities may be explained in terms of our mindreading abilities. However, he also notes that this is not guaranteed. For this reason, we will stick with the restrictive sense of mindreading here to avoid introducing unnecessary controversies about the scope of our mindreading abilities.

¹⁵ Other approaches include the ‘intentional’ approach of Dennett (1987) and the ‘interpretation’ approach of Davidson (1984). A good, albeit opinionated, summary of these views, along with their attendant difficulties, can be found in Goldman (2006a, Ch. 2-5).

¹⁶ It is widely agreed that the relevant theory of mind must be tacit rather than explicit, but there is an important dispute about whether this theory of mind is *learned* or *innate*. According to one view, the so called ‘Child Scientist’ view, children are like little scientists, running around testing hypotheses about how other people behave. Eventually, this engagement with the world allows them to form a theory about how other minds work. This view is defended by Gopnik and various collaborators, e.g. Gopnik and Wellman (1994), Gopnik and Meltzoff (1997). The innate view maintains that our theory of mind is innate, perhaps in a similar way to the way in which Chomsky (1965) takes our theory of grammar to be innate, or perhaps because we have an innate domain specific mindreading module. This view is defended by philosophers like Carruthers (1996) and psychologists like Leslie along with various collaborators (e.g. Leslie 1994, Leslie and Scholl 1999). The main point in favour of the innate view is that children develop their mindreading abilities along a set development timeline. It is often pointed out that one would expect discrepancies based on intelligence and other factors if children independently figure out how minds work. Against the innate view, usually there is a worry about the nature of

Simulation theorists argue that the theory-theory approach is wrongheaded and that mindreading is *process*-driven rather than *theory*-driven (Goldman 1989, p. 173). Defenders of simulation theory argue that when we reason about how others will behave, or ascribe mental states to them, we *simulate* their decision-making processes, which is accomplished at least in part by *simulating* their mental states. If this is right, then being able to mindread means being able to engage in this sort of simulation process, and does not depend on possessing any sort of tacit theory of mind.

To see why we can associate simulation theory with the idea of voluntary imaginative counterparts, it will be helpful to set out roughly how a simulation theorist will explain how I predict which sandwich a friend (we can call him Laurie) will buy when he goes into *Pret a Manger*. According to simulation theory to predict what sandwich Laurie will buy, I will make use of my own decision-making system, but will take this system *offline*. This will mean that I input ‘offline’ beliefs and desires into my decision-making system (often referred to as ‘simulated’ or ‘pretend’ beliefs and desires) that share the same content as Laurie’s beliefs and desires: for example, he might have a desire for a tuna sandwich and a belief he can get one at *Pret*.

This means that instead of performing a series of actions, such as going to buy a tuna sandwich, you merely predict the action Laurie will perform: you output something like a simulated decision to buy a tuna sandwich. Currie and Ravenscroft (2002, pp. 11-23) refer to these offline beliefs and desires as ‘belief-like’ and ‘desire-like’ imaginings and they are both supposed to be examples of voluntary counterparts. They are states that bear some relation to genuine beliefs and genuine desires, but nonetheless are not genuine beliefs and desires, and since we are able to engage our mindreading capacities at will,

innate modules or the very possibility of innate capacities. Chiefly then, this debate turns on whether one finds it more implausible that we possess an innate theory of mind or that, in spite of developing their theory of mind independently, children still end up developing their mindreading capacities on a consistent developmental timeline.

we can also entertain belief-like and desire-like imaginings at will according to simulation theorists.¹⁷

Since I will not be offering a theory of mindreading in this thesis, we do not need to assess whether this offers a plausible account of our mindreading capacities.¹⁸ Instead, we should look at how simulation theorists set out the nature of imaginative counterparts, since one can reject the tenets of simulation theory, whilst agreeing that we should introduce a notion of voluntary imaginative counterparts.

To explain our ability to generate these sorts of counterparts, Currie and Ravenscroft (2002, p. 9) introduce a capacity they call the ‘recreative’ imagination, a sort of imagination that allows us to ‘recreate’ ordinary mental states. This gives the question of what

¹⁷ There is another sort of simulated state that some simulation theorists have drawn our attention to. Simulated beliefs and desires are examples of what we can call *high-level* simulated states. These are simulated states that are typically under control of the will, with us being consciously aware of them. Goldman (2006a, Ch. 6) argues there can also be *low-level* simulated states that are not supposed to be subject to the will and which we may not be consciously aware of. Goldman argues these low-level states are involved in *motor simulation* and *emotion simulation*. Goldman’s argument for introducing the idea of low level simulation trades on the importance of mirror neurons. Mirror neurons are neurons that have been found to fire both when performing a certain action and also when watching someone else perform a certain action. In the emotion case, there is evidence similar mirroring occurs (a helpful summary of this sort of research can be found in Rizzolatti and Sinigaglia (2008)). Whether the sorts of unconscious processes signified by these mirror neurons should be called simulation processes, and whether their products should be classified as imaginative counterparts is unclear. Spaulding (2012), for example, is sceptical that mirror neuron processes should be thought of as simulation. The distinction here isn’t one we need to worry about since so far as I am aware no one has proposed putting belief-like imaginings into the low-level category.

¹⁸ For example, there is a worry about how we decide which pretend beliefs and desires to input into our decision-making system (Goldie 2003, pp. 334-335). There are also worries about the so-called ‘threat of collapse’, the worry that simulation theory collapses into theory-theory, since we need to theorise that other people are relevantly similar to us (Jackson 1999). Bearing in mind these sorts of issues, many simulation theorists now embrace *hybrid* theories which allow for some role for simulation and some role for theory (such as Heal 2003 and Goldman 2006a).

makes one sort of state a voluntary counterpart to another, or to phrase this in Currie and Ravenscroft's terminology, what makes one state a 'recreation' of another?

We can begin by looking at what they tell us about how to answer this question:

Imaginative projection involves the capacity to have, and in good measure to control the having of, states that are not perceptions or beliefs or decisions or experiences of movements of one's body, but which are in various ways like those states—like them in ways that enable the states possessed through imagination to *mimic* and, relative to certain purposes, to *substitute* for perceptions, beliefs, decisions, and experiences of movements. These are what we are calling states of recreative imagination. (emphasis mine) (Currie and Ravenscroft 2002, p. 11)

This brings out first of all the importance of the will when thinking about voluntary counterparts. As they put the point, it is important that we are able to 'control' our having of these counterparts. In light of this, they do not commit to every mental state having a counterpart. For example, they argue that the emotions do not have counterparts. I will say more about this issue in section 1.6 after we've introduced the notion of involuntary counterparts.

Secondly, it is notable that they place emphasis of the notions of mimicry and substitution. What exactly does it mean for one state to mimic another and thus serve as a substitute? There are two sorts of answer we can give to this question. We can focus on the features of the supposed counterpart state in question and whether it *resembles* another mental state; or we can focus on the mechanisms involved in producing the supposed counterpart state to see whether they are the same as those involved in producing a genuine instance of that mental state. This approach amounts to interpreting mimicry as entailing the re-use of certain cognitive mechanisms (Hurley 2005 is the most prominent defender of this sort of approach).

If we adopt the re-use approach, we will end up committed to connecting the notion of a voluntary counterpart to the specific claims of simulation theory, since we will associate the notion of a voluntary counterpart with a general method or mechanism for

generating imaginative counterparts.¹⁹ This is because endorsing a re-use based account of counterparts will presumably involve us accepting that we have a general ability to take our ordinary cognitive mechanisms ‘offline’. On the other hand, if we focus on resemblance, this provides us with a sketch of voluntary counterparts that we can make use of even if we reject the tenets of simulation theory. Indeed, the idea of a mental state that resembles another but is a counterpart which is under the control of the will is intuitive and we can accept that this sort of counterpart exists without committing ourselves to introducing any specific mechanism or method by which these counterparts are generated.²⁰

That said, we need to be a bit more precise about what ‘resemblance’ means here, since everything resembles everything else in some trivial respects. To make this claim more specific, we can argue that to establish that one state is a counterpart to another in the voluntary sense, we need to consider whether the state in question is under the control of the will and also whether it shares with the target state:

¹⁹ It’s worth noting that Goldman (2008) is sceptical of any approach that sets aside resemblance. For example, if I am looking at a coat hanger and at the same time my girlfriend is visualising a tree this would suggest her mental state is a simulation of mine, since it involves the re-use of her visual mechanism. This cannot be the right way of understanding what it would be for a state to be a counterpart to my seeing a coat hanger. Primarily, as Goldman would emphasise, this is because my girlfriend’s visualisation of a tree doesn’t resemble the coat hanger in front of me in any meaningful sense. I’m not entirely convinced by this objection, since we can tighten up the notion of re-use so that it must be *symmetrical* rather than simply *synchronous*. That being said, I expect Goldman would object that this is just to smuggle in a reference to resemblance by introducing the idea of a ‘symmetrical’ process.

²⁰ To give some examples of philosophers who appear to endorse voluntary counterparts without necessarily committing to simulation theory, Budd (1989, Ch. 5) discusses something like the idea of voluntary imaginative counterparts in his discussion of mental imagery, calculation in the head, and inner speech, and Wollheim (1973) discusses something like the notion of voluntary counterparts when considering imaginative identification.

1. Introspective Features (e.g. phenomenology and how we self-identify the state in question: does it seem similar yet distinct from a genuine state)
2. Functional Features (e.g. does the counterpart share some functional similarities with the target state?)
3. Neurological Features (e.g. are similar brain areas involved in the imaginative activity in question as well as the genuine mental state?)

To see how reflection on these features can justify the introduction of some voluntary counterparts, we will consider whether forming mental images can be thought of as a voluntary counterpart to seeing.

1.3 Voluntary Imagining Seeing

The suggestion that forming mental images – which we can call *visualising* – is some sort of counterpart to seeing is a fairly common one to find in philosophical discussions. For example, the relation between the two has been frequently remarked upon by philosophers such as McGinn (2005, p. 7) and Martin (2002, p. 403). If visualising is a counterpart to seeing, it will be a voluntary counterpart since our ability to form mental images is under the control of the will. While writing this sentence, I can choose to form a mental image of a dog, or a cat, or anything else I am able to represent imaginatively, no matter how bizarre. However, the fact that this is a voluntary counterpart is not meant to suggest we freely choose to form every mental image we visualise. Sometimes we can form a mental image of something without having consciously decided to, and indeed even if we actively don't want to. For example, a person who is feeling insecure about their romantic relationship with their partner might find themselves picturing them with another lover, even if they don't want to form these images. The importance of the will here is that mental images are at least *in principle* under the control of the will.

If the way I suggested we should characterise voluntary counterparts is correct, we should be able to justify maintaining that visualising is a voluntary counterpart to seeing by considering introspective, functional and neurological similarities between the two.

We noted above that it is a commonplace to accept that when someone forms a mental image of something this seems to have something in common with really seeing that something, and it is not implausible to suggest introspection reveals this similarity between seeing and visualising. That being said, seeing and visualising can also be introspectively recognised as distinct. If we were generally unable to make the distinction between our mental images and genuine perceptions of the world, this could lead us to perform various inappropriate actions. If merely forming a mental image of a tiger looking hungrily in our direction was often mistaken for seeing an actual tiger this would lead to odd behaviours, such as trying to run away from a tiger when previously you were happily sitting on a train heading towards London.²¹

To find some potential functional similarities between these two states, we can rely on armchair reflection and note that mental images can be helpful for guiding our behaviour in a similar way to actual seeing. If I am unsure whether to buy a new sofa for my living room, I can visualise placing the sofa in my room to see whether it will clash with the existing décor. I could also do this by actual placing the sofa in my room and looking at it. From a somewhat more scientific perspective, Goldman (2006a, p. 157) points out that a stock question in research on visualising is to ask people to count how many windows are in their living room by visualising their living room. This sort of task is performed just as well when we visualise our living room as compared to when we can actually see our living room, again implying a functional similarity between seeing and visualising.

As for, neurological similarities, Goldman (2006a, pp. 151-157) helpfully summarises studies which show various areas of the brain are involved in both visual perception and the formation of mental images. To give one of his examples, he notes studies from Kanwisher *et al.* (1997) and O'Craven and Kanwisher (2000) that show the fusiform gyrus is activated when we see faces and when we visualise them. He also notes a study

²¹ As with my point about the will, this is not to say we never make these sorts of mistakes. We might think this occurs in some forms of hallucination or delusion.

from Damasio, Tranel, and Damasio (1990) which shows that damage to lesions associated with the fusiform gyrus impairs both our ability to recognise faces and our ability to visualise faces. This suggests that there are important neurological similarities between at the very least seeing faces and visualising faces.

This sort of ‘imagining seeing’ is thus a relatively uncontroversial example of a voluntary imaginative counterpart. This sort of imagining seeing is under the control of the will; we have introspective evidence that there are similarities but also differences between mental images and actual perception; there is evidence of some functional similarities; and we have neurological evidence of some similarities.

We will shortly consider whether belief has this sort of counterpart in chapter 2. Before we do this, however, we should turn to introduce my second notion of a counterpart – involuntary counterparts.

1.4 Involuntary Counterparts & Fictional Truth

An involuntary counterpart can be characterised roughly along the following lines:

IVC: A mental state that resembles another mental state, which we automatically enter into when we engage with something we recognise as being fictional or imaginary, and has fictional content

A first question to address is why we might want to separate out this sort of notion of a counterpart. As I noted in the introduction, one reason is that this notion seems to arise from Walton’s (1973, 1990) theory of fictional truth – which we will discuss shortly – but it has not been frequently remarked that this means that what Walton has in mind when he appears to introduce imaginative counterparts is somewhat different to what simulation theorists have in mind. A second way to bring out the motivation for introducing this notion of a counterpart is to note three claims about imaginative counterparts and engaging with fiction that we might take to be true, but which seem to be somewhat in tension with one another:

1. Imaginative counterparts can in principle be wilfully entered into.

2. Engaging with fiction involves some imaginative counterpart states
3. Our responses to fiction are somewhat involuntary.

Some take this to suggest that in fact (2) is false, and that there are no imaginative counterparts involved in our engagement with fiction (Tullmann 2016, pp. 787-788). Others argue that this entails that some voluntary counterparts are involved in our engagement with fiction, but that these counterparts behave in an automatic way in this context (e.g. simulation theorists such as Currie and Ravenscroft 2002 and Goldman 2006b). I don't think either of these responses to noting the automatic nature of some our responses to fiction is correct. Though I am sympathetic with Tullmann's view, at least when it comes to the question of whether a counterpart to belief is involved in our engagement with fiction, I think we can at least in principle separate out voluntarily generated imaginative counterparts, such as those that might be involved in mindreading, from involuntary counterparts that we enter into automatically when we engage with something fictional. In other words, I don't think we should simply accept (1) as true: there might be some imaginative counterparts that are not even in principle under the control of the will.²²

As with voluntary counterparts, an involuntary counterpart resembles another mental state in some respects, but unlike with voluntary counterparts, we cannot rely on mere introspection when it comes to establishing whether a state has this sort of counterpart. This is because these sorts of counterparts can sometimes be phenomenologically identical to the state they are supposed to be a counterpart to.

²² Of course, we often voluntarily decide to engage with fictional or imaginary content. The point here is that some of our responses that follow this decision will not be voluntary: having decided to watch *Amour* I do not voluntarily start to cry or feel sad. We'll discuss this further in section 1.6 where we will discuss whether emotions have imaginative counterparts.

Furthermore, because they are not subject to the will, we have no obvious introspective features we can reflect on to tell them apart from the state they are a counterpart to.²³

Instead, the first step to establishing whether a state has an involuntary imaginative counterpart is to consider whether we sometimes appear to enter into the state in question automatically when engaging with something we recognise to be fictional or imaginary. Following Walton (1990, pp. 35-43), we can understand ‘fictional’ here in a broad sense as relating to what we call fiction in ordinary language, but also potentially some other forms of representational activity, such as pretence and representational art.

When we engage with things that we recognise as fictional, our mental states will often have *fictional content*.²⁴ Fictional content is the sort of content our mental attitudes – imaginative counterparts or otherwise – have when directed towards fictional characters, events, and so on, which more generally we can call *fictional entities*. As such, I am accepting that fictional entities can be the intentional object of at least some of our mental states (Crane 2013, Tullmann 2016).²⁵ This notion of fictional content allows us to clarify that when considering whether to introduce an involuntary counterpart, we shouldn’t think that we have some reason to introduce one whenever we enter into a state as a result of engaging with something we recognise as being fictional or imaginary.

²³ For example, Scruton (1974, p. 94) notes the importance of subjection to the will for distinguishing counterparts from genuine mental states: ‘There is a feature of both imagery and imagining which serves to distinguish them from many mental states. This is the feature of subjection to the will.’

²⁴ ‘Content’ can be understood here as relating to potential propositional contents of our mental states and attitudes, and also non-propositional contents if one thinks, for the example, that the contents of perception are non-propositional (for a good discussion of this issue see Crane 2009).

²⁵ A slight complication arises when the intentional object in question is a real individual or place, etc., but as described in fiction. In light of this possibility, we can suggest that a mental state has fictional content when either its intentional object is a fictional entity, or when its intentional object is a real entity, but where we are supposed to take that entity to have qualities only ascribed to it in fiction (e.g. when we are reading the Sherlock Holmes novels, the intentional object of some of our thoughts will be London, but these novels ascribe properties to London, such as that there is a 221B Baker Street, that it does not have in reality).

The state in question must also have fictional content. If we didn't make this addendum, then coming to believe that 'Harry Potter is a novel' upon reading the Harry Potter novels could be taken to imply that we might need to say this belief is in fact an involuntary counterpart to belief.

One might also wonder why I added the requirement that the content in question has to be known to be fictional by the agent. This is to avoid reaching counter-intuitive conclusions in cases where, for example, one mistakenly regards the Harry Potter novels as non-fiction, and comes to form beliefs about wizards based on reading them. There is a sense in which we might want to say that at least some of these beliefs will have fictional content, if one allows that we can have beliefs or other attitudes with fictional content even if we don't recognise it as being fictional content. The knowledge condition in this sketch of an involuntary counterpart helps to bring out that in this sort of case, you will take genuine attitudes towards this content. If you mistake *Harry Potter* for non-fiction, you will not form belief-like imaginings about wizards, you will just form mistaken or false beliefs about them.

This leaves the question of how we can justify the introduction of this sort of counterpart. As noted in the introduction, I take it that merely noting that a mental state is entered into as a result of engaging with something fictional or imaginary – and has fictional content – is insufficient for justifying the introduction of this sort of counterpart. This is because it looks like some of our ordinary states can be entered into when we engage with something we recognise as being fictional and can have fictional intentional content (e.g. I can come to believe that 'Sherlock Holmes is a detective [in the fiction]' when I engage with the Sherlock Holmes novels). This is the second claim about imaginative counterparts that I will defend in this chapter

Instead, in order to justify the introduction of an involuntary counterpart, we need to consider whether a) *constraints* on the mental state in question necessitate the introduction of this sort of counterpart, or b) whether *puzzles* related to a particular human activity are best solved by introducing this sort of counterpart. This is my third claim about imaginative counterparts, and we will elaborate on this principle in sections 1.5 and 1.6.

To try to find some examples of this sort of counterpart and to justify these two principles, it will be helpful to introduce Walton's theory of *fictional truth*. His approach to understanding fictional truth leads to him introducing some imaginative counterparts that are involved when real world attitudes, feelings, and actions constitute fictional attitudes, feelings, and actions. This sort of idea captures part of my intuition behind the introduction of a class of involuntary counterparts, since this sort of transformation from the real to the fictional is often supposed to be automatic when we consume a work of fiction, but also doesn't always entail the introduction of an imaginative counterpart. His theory also allows us to draw some initial parallels between fiction and pretence.

We can introduce the idea of fictional truth by noting that we often make seemingly odd statements about fictional characters and other fictional entities that do not appear to be literally true, such as 'Sherlock Holmes is a detective'. The difficulty here is that 'Sherlock Holmes' looks to be an empty name: Sherlock Holmes does not exist, so there's no object this sentence can refer to. Since this sentence appears to fail to refer, it looks like 'Sherlock Holmes is a detective' is at best neither true nor false, and possibly just false (Friend 2007, p. 143).²⁶ We could reply here that Sherlock Holmes is, in fact, an abstract object of some sort, but this will not immediately solve our mysteries here since it's not clear how an abstract object can be a wizard. We will say a bit more about these issues related to referring to fictional characters and entities in chapter 5, since it can be argued that issues related to reference to the non-existent suggest we need to introduce belief-like imaginings into an account of how we engage with fiction.

²⁶ Which way we go here in part depends on one's theory of reference when it comes to the nature of empty names (this grouping includes fictional characters, but also failed scientific posits such as Vulcan and other things we apparently refer to that don't exist). On one view, if the name 'Harry Potter' fails to refer, then sentences using the name will simply lack truth value and so cannot be evaluated as true or false. This is, of course, a view associated with Russell (1905a, 1905b). It stands in contrast to another view that holds that, since the name fails to refer, sentences using it are straightforwardly false. This is associated with Frege's (1948) view that fictional names lack reference but possess sense.

However, although we might be happy to accept that a sentence like ‘Sherlock Holmes is a wizard’ is false (or perhaps simply lacks a truth value), there is something intuitively right about the claim that Sherlock Holmes is a detective. If I found myself in a debate with a friend over whether Holmes was a detective or a wizard, I would insist that he is most assuredly a detective and I would be puzzled over why my friend has ended up thinking he is a wizard.

Walton makes sense of this sort of issue by suggesting that statements like ‘Sherlock Holmes is a detective’ make claims about what is *fictionally true* (or *fictional* for short), rather than about what is literally true. To express the claim that a sentence is fictional he uses the notation **p** (Walton 1973), which amounts to something like ‘it is fictional that *p*’, ‘it is true in fiction that *p*’, ‘it is fictionally true that *p*’ and so on. That being said, this does not mean that Walton thinks we need to introduce a new ‘species’ of truth. Instead, Walton is arguing that when we make these sorts of claims, we are engaging in a game of *make-believe*, where we pretend various things are true.

If we embrace this make-believe view, then when we say that ‘Sherlock Holmes is a detective’ what we are actually saying is that it is fictional that Sherlock Holmes is a detective (**Sherlock Holmes is a detective**). This involves ‘pretending’ that the empty name ‘Sherlock Holmes’ has a referent and that sentences containing the name can be ‘true’. As such, for Walton, strictly speaking these sorts of utterances are only meaningful within the scope of a game of make-believe.

This introduces the question of what makes something fictionally true. Walton suggests that what is fictionally true is determined by what we are *prescribed to imagine* by a given fiction (Walton 1990, p. 39).²⁷ The Sherlock Holmes novels prescribe imagining that Sherlock Holmes is a detective, and so it is fictionally true that Sherlock Holmes is a detective in the fictional world of these novels. This thus helps to explain why we have

²⁷ Walton (1990, p. 21) is using ‘imagine’ as a placeholder here. As such, he should not be read as saying simply that what is fictionally true is what we are prescribed to imagine in a belief-like way: he has a richer notion of imagining in mind, which means we could also be prescribed to imagine *seeing* or *doing* things.

different intuitions about the truth of ‘Sherlock Holmes is a detective’ as compared to ‘Sherlock Holmes is a wizard’. The former is something we are prescribed to imagine and is thus fictionally true, whereas the latter we are not prescribed to imagine. It is not up to us as mere readers what is true in the world of Sherlock Holmes.²⁸

In light of this point, Walton holds that something can be made fictional in part by facts about the actual world and the rules of the game of make-believe we are playing. As such, merely imagining something to be the case is often insufficient for making that something fictional (Walton 1990, pp. 37-38).²⁹ To clarify why this is so, it will be helpful to consider how Walton applies his theory of fictional truth to childhood pretence.

Let’s return to the child in the introduction to this thesis, who was pretending to have a tea party. She may well make various statements during this episode of pretence that cannot be understood as expressing what is literally true. She might say, for example, that ‘there is a cup of tea in front of me’ or that ‘teddy drank the last cup of tea’. Once again, these claims appear to be literally false since there was no tea at her pretend party. Walton argues that these statements also make claims about what is *fictionally* the case, where the relevant fiction is the pretence episode instigated by the young girl.³⁰

In an episode of pretence like this, sometimes fictional truths are determined by things that don’t depend on the actions and knowledge of the pretenders. Walton illustrates this with the example of a game where children pretend every tree stump nearby is a bear. If there is a stump nearby that the children haven’t noticed, then it remains true that *there is a bear nearby*, even though none of the game’s participants are aware

²⁸ One could argue that we can imagine that *Sherlock Holmes is a wizard* with this in effect creating our own fictional world in which this is fictionally true. The point here is that this won’t affect what is true in the world of the novels.

²⁹ Only ‘often’ because one might think that in cases like daydreaming, one’s imaginings make things fictionally true merely by being imagined. For example, if I start daydreaming about going to a picnic and imagine eating a ham sandwich, this looks like it will make it fictionally true that I am eating a ham sandwich.

³⁰ This means Walton has a highly permissive notion of fiction, which is roughly equivalent to our ordinary language notion of a ‘representation’. Friend (2008, p. 154) argues we should think of Walton as having defined a special category of ‘Walt-Fictions’.

of this fictional truth (Walton, 1990, p. 37). In this game, participants are prescribed to imagine a bear whenever they see a stump, and therefore a bear's presence in the game depends not on whether the children playing the game imagine a bear is present, but merely on whether a stump is present. The children's knowledge of how many stumps are in their vicinity does not impinge on fictional truths about the number of bears nearby.³¹ In Walton's terminology, this is because these stumps are 'props' that prescribe imaginings. In a similar vein, he argues that novels, plays, etc., are also props that prescribe imaginings and this is partly why it is not up to us what is fictional true in them, since we do not determine what these props prescribe imagining.³²

This helps to reveal that the notion of 'pretend' in Walton's theory is something of a term of art. Walton is not proposing that we consciously make a decision to pretend that sentences like 'Sherlock Holmes is a detective' are true. The sort of game we engage in when we read the Sherlock Holmes novels is supposed to be an *automatic* response to reading them. One doesn't have to make an active decision to engage in a game of make-believe while reading a work of fiction, even if one may well voluntarily decide to engage with the fiction in question. To flag up this point, he marks a distinction between *deliberate* and *spontaneous* imaginings (Walton 1990, pp. 13-16), which captures part of what I have in mind when I refer to voluntary and involuntary counterparts.

We can further associate involuntary counterparts with some of Walton's ideas here by noting that in addition to statements of 'fact' about fictions, we can also make statements involving propositional attitudes when we talk about our engagement with fiction.

³¹ Though their awareness, or lack thereof, of nearby stumps will generate truths related to what it is fictional that they are aware of.

³² Whilst what one is prescribed to imagine in games of pretence such as this bear game is usually fairly clear, it is not always so obvious how to set out what works of fiction prescribe imagining. Indeed, some fictions purposefully choose to be ambiguous, or contain partially unexplained plot points. An interesting recent example of this would be the third season of Mark Frost and David Lynch's *Twin Peaks* where the show was ambiguous throughout as to what was happening and why. We can note in response that fictions allow for *indeterminacy*. An interesting discussion of this issue can be found in Woodward (2012).

For example, we could report our ‘beliefs’ and say something like ‘I *believe* Sherlock Holmes is a detective’, or we can report on our ‘emotions’, making claims like ‘I was *scared* when the alien (from *Alien*) came down from the vents’.

Walton argues that these utterances also express what is fictional, rather than what is true. To make sense of how this could be so, Walton (1990, pp. 58-61) proposes that when we engage with fictions we create a ‘game world’, which consists of the *work world* of the fiction that we are engaging with and our own responses to the fiction (which are thus only fictional in the context of our specific game of make-believe). The work world can be thought of as ‘composed of those fictional truths that are fictional in all authorised game worlds’ (Meskin and Robson 2012, p. 209). The notion of ‘authorised’ here serves to rule out saying, for example, that if I engage with the Sherlock Holmes novels and decide to take it that Holmes is not a detective, this means ‘Holmes is a detective’ is not fictionally true in the work world. In taking it to be fictionally true that Holmes is not a detective, I create an *unauthorised* game world (Walton 1990, p. 60).

To see how this works in practice, consider the movie *Alien*. Everyone who engages with this work is prescribed to imagine that the alien kills Parker and Lambert, so this is fictionally true in the work world of *Alien*. However, it is only in my specific engagement with *Alien* that I am prescribed to imagine that I believe that the Alien killed Lambert and Parker, since ‘I’ do not feature in anyone else’s fictional engagement with the movie. This claim is thus only fictionally true in the game world I create when I engage with *Alien*.

This introduces the question of what makes these sorts of attitude reports fictionally true. In these sorts of cases, Walton argues that our real-world attitudes, feelings, and actions constitute fictional attitudes, feelings, and actions (e.g., Walton 1990, 246-247). However, in some cases these real-world attitudes etc. will be genuine mental states, and in others, they will be mere counterparts. For example, if it is true that *I believe Sherlock Holmes is a detective* in the context of my engagement with the Sherlock Holmes novels, this could entail that this is made fictional because I genuinely believe *Sherlock Holmes is a detective* (I believe it is fictional that Sherlock Holmes is a detective) or it could mean that this is made fictional because in reality I have entered into some sort of

counterpart state to believing (I imagine in a belief-like way that Sherlock Holmes is a detective).³³

If step outside Walton's framework, this helps to reveal that sometimes our ordinary attitudes appear to have fictional intentional content, and in some of these cases we will want to say this is because we have in fact entered into an imaginative counterpart state. As such, we need a way of figuring out when it will be fruitful to suppose an involuntary counterpart is involved in our engagement with the fictional or imaginary. I maintain that to justify the introduction of an involuntary counterpart, we need to note constraints on the genuine state in question or that puzzles arise if we don't introduce an involuntary counterpart. To see how we might establish whether either of these conditions is in place, it will be helpful to discuss whether seeing has an involuntary counterpart.

1.5 Involuntary Imagining Seeing

Consider someone wandering around The Courtauld Gallery with a friend. They come across Manet's wonderful *A Bar at the Folies-Bergère* and their friend asks what they see. They reply, 'I can see various things: bottles of wine, a woman looking somewhat pensively away from my gaze and a trapeze artist' (among other things, of course, bearing in mind the richness of this painting). Walton (1973, p. 284) points out that these claims about what the gallery goer can 'see' do not appear to be literally true, in much the same way statements we make about fiction or pretence don't appear to be literally true. Our gallery goer is just looking at paint marks on a canvas, not wine bottles, a woman and a trapeze artist.

The obvious point to make here is that they do not literally see these things, they merely see a representation of them. This raises the question of what it means for a

³³ By *I believe Sherlock Holmes is a detective*, I intend only to refer to cases where I am engaging with the Sherlock Holmes fictions and where for Walton this would be an accurate report of what is fictionally true of my engagement. If a fiction just happens to declare 'Jack Davis believes that Sherlock Holmes is a detective', this, of course, does not entail that I am in a counterpart state to belief, nor that I have beliefs about Sherlock Holmes.

painting to represent or depict something. One suggestion might be that paintings resemble what they represent, but the notion of resemblance we would need to introduce here is hard to specify. For example, a two-dimensional configuration of paint on a canvas seems to bear little resemblance to a physical, three-dimensional woman standing at a three-dimensional bar.³⁴

Walton argues that when looking at paintings, we *imagine seeing* what they depict (1973, 2008). Our gallery goer is imagining seeing a woman standing at a bar when they look at *A Bar at the Folies-Bergère*. Walton makes the case for this by arguing that looking at paintings is another instance of a game of make-believe and so for him, paintings are classified as fictions.

He suggests that paintings prescribe us to imagine various things, and so when looking at a painting it can be true that *I see a woman standing at a bar* if this is what we are prescribed to imagine by the painting (Walton 1973, p. 300). This introduces a further sense in which seeing can have an imaginative counterpart. If this is right, we can argue that there is an involuntary counterpart to seeing involved in our looking at paintings.

If we put the idea here in Waltonian terms, this sort of imagining seeing somehow involves a real-world action (looking at a canvas) constituting a fictional action (seeing whatever is represented by the painting). Walton is proposing that it is made the case that our gallery goer *sees a woman standing at a bar* because seeing the canvas in question doubles in some sense as *seeing a woman standing at a bar* (Walton 1973, p. 304). In more neutral terms, we could argue that the content of our 'seeing' is fictional content, since we recognise we do not literally see what the painting depicts and paintings can depict people and things that do not exist and which we know to not exist.³⁵ Our

³⁴ Two more sophisticated attempts to develop a resemblance theory are offered by Peacocke (1987) and Budd (1993), which are in turn criticised by Wollheim (1998).

³⁵ There is an interesting question here about whether there can be non-fictional painting. For example, is a portrait of a monarch a work of non-fiction? Even if this is so, this doesn't pose a problem for Walton's view since he can bring out that seeing a portrait of someone is not the same as literally seeing someone.

question then, is whether we are justified in introducing an involuntary counterpart to seeing to make sense of our experience of looking at paintings.

We can note at the outset that Walton's account of depiction is controversial, in part because theorists have been sceptical of the idea that our experience of paintings involves the imagination (e.g. Saville 1986, Wollheim 1991) As such, instead of introducing an imaginative counterpart here, we might think that a special sort of seeing is involved when we look at paintings. Wollheim (1980, pp. 205-226) argues that we see things 'in' paintings and introduces the notion of *seeing-in*, which is perhaps the most well-known version of this sort of approach. This helps to bring out the importance of my second claim about imaginative counterparts: that merely noting that a mental state or attitude has been entered into as a result of engaging with something fictional or imaginary and has fictional content does not *per se* justify introducing an involuntary counterpart to that state. In the case of seeing, we can instead argue that we can in some sense have episodes of visual perception with fictional contents, albeit thanks to a special variety of seeing such as seeing-in.

However, Walton can also be read as trying to introduce this counterpart with reference to the two ways I suggest in my third principle about counterparts: he notes *constraints* on seeing, and argues *puzzles* arise if we suppose genuine seeing is involved in our viewing of paintings. The constraint on genuine seeing seems to be that looking at a canvas which depicts a given scene is not an instance of literally seeing that scene. The puzzle is why we nonetheless describe ourselves as seeing things when we observe a painting. In response, we could attempt to solve this puzzle without introducing a counterpart to seeing, and we could argue that whilst there may be some sorts of constraints on ordinary seeing, these same constraints are not present when it comes to a special kind of seeing, such as seeing-in.

For present purposes, we do not need to settle whether or not Walton is justified in introducing a notion of involuntary imagining seeing. This debate instead merely serves to clarify what considerations might be introduced when debating whether to introduce an involuntary counterpart. We can also make two broader points about imaginative counterparts in general based on this discussion.

Firstly, we can note some differences between Walton's notion of imagining seeing and the one associated with visualising to bring out why I am separating out voluntary and involuntary counterparts. In section 1.3, I allowed for there being instances when voluntary counterparts sometimes behave automatically, such as if you form a mental image you don't much want to form. The reason this doesn't make involuntary counterparts reduce to voluntary counterparts is because involuntary counterparts are not even *in principle* subject to the will. In the case of imagining seeing in Walton's sense, we cannot choose to imagine seeing something apart from what is depicted by the painting we are looking at. Furthermore, however we spell out the relevant sort of imagining seeing here, it cannot reduce to the formation of mental images since seeing what a painting depicts does not involve forming a mental image of its contents (Walton 1973, p. 286).

Secondly, Walton's account of imagining seeing demonstrates my fourth claim about counterparts, that accepting that a state has a voluntary counterpart does not mean we have to accept that the state also has an involuntary counterpart, and vice versa. For example, one could deny that forming mental images should be thought of as involving an imaginative counterpart, whilst accepting that one is involved in imagining seeing things in Walton's sense. Indeed, many philosophers would likely accept the opposite of this, allowing that mental images are a counterpart to seeing, whilst denying that we need to introduce a counterpart to seeing in Walton's sense. We can further justify and bring out the importance of this fourth claim by considering whether the emotions have imaginative counterparts.

1.6 Emotional Counterparts

The emotions are intimately connected to the imagination. Harris (2000, p. 59) discusses the fact children appear to be scared of things they merely imagine, such as monsters.³⁶ Furthermore, entirely voluntary imaginings can lead to emotional responses. For example, studies have suggested imagining things we have a phobia of can prompt fear responses (Lang *et al.* 1983) and if you imagine slowly drawing a knife across your hand this will likely result in a shudder of apprehension Murray (1995, p. 116).

Carruthers, in a book review of *Recreative Minds* offers a good summary of some of these connections:

[i]magination can certainly evoke real emotions — imagined insults can make you angry; imagined danger can make you afraid; the death of a character in a novel or film can make you sad; and so forth. ... imagined delicacies can make you hungry (wanting food), as imagined sex can make you sexy (wanting sexual relief). (Carruthers 2003)

In light of this connection between imagination and the emotions, many philosophers have been reluctant to introduce voluntary counterparts to the emotions. For example, as we noted previously, Currie and Ravenscroft (2002, Ch. 9) explicitly deny that we can ‘recreate’ emotions. Part of the reason for this reluctance seems to be that imagining feeling sad, for example, will amount to your being really sad about what you imagine, rather than entering into a mere *sadness-like* state.³⁷

³⁶ When these children are asked about the nature of these monsters, they reveal they understand the difference between real and imaginary monsters (Harris *et al.* 1991). Nonetheless, this doesn’t seem to make them any less scared.

³⁷ There are some exceptions to this position, for example Goldie (2005, p. 133) thinks that we need to introduce voluntary counterparts to the emotions along these lines to explain the idea of ‘imagining feeling’ suggesting: ‘[i]t seems possible for me, for example, to imagine something threatening, and to imagine feeling afraid of the threatening thing that I imagine, where the imagined fear is part of the *content* of what I imagine, and not a response *to* what I imagine.’

There is a more heated debate about whether at least some emotions might have involuntary counterparts. The quote from Carruthers illustrates that it is usually accepted that we can exhibit affective responses to the events depicted in works of fiction. Radford (1975) argues that the fact we exhibit these sorts of emotional responses shows us to be *irrational*, since we ought not to exhibit affective responses to merely fictional persons and events. We can call this the *rationality* worry.

Radford's initial paper has generated a great deal of discussion about a more general puzzle that is typically called the *paradox of fiction*, a puzzle that is supposed to arise because we accept something like the following three claims:

1. We exhibit affective responses to works of fiction.
2. We do not believe that fictions depict real events and happenings.
3. Our emotions are subject to a coordination condition (Gendler and Kovakovich 2006, p. 241). In order to exhibit an affective response towards something, we normally need to believe that the object of our response is real (e.g. it would be odd if I told my friend I'm upset because my older brother drank my last beer, when I don't have an older brother and I'm aware of this fact).

As such, this paradox introduces in addition to Radford's rationality worry a *causal* worry about how emotional responses to works of fiction are generated, bearing in mind that we do not believe the events depicted or described in them to have actually happened. To illustrate this purported paradox, we can consider the opening montage of Pixar's excellent *Up*. During this montage, we watch the development of the main character's relationship with his eventual wife, followed by the hardships that beset their married life: their inability to have a child; their inability to travel; the way their savings also get eaten up by unexpected events. Near the end of the montage, we witness his wife's untimely death just before they get to fulfil their life's dream to travel to South

America. At this point, whenever I watch the film I notice tears starting to form.³⁸ A natural description would be that I'm feeling sad, and that I'm feeling sad *because* of the events depicted by the film. At the same time, I'm well aware that the events depicted aren't real. I know this couple don't exist and never have existed (and never will exist³⁹). And yet, as if to fly in the face of the coordination condition, this knowledge does nothing to dry my tears nor does it serve to make me any less upset.⁴⁰

There have been too many responses to this paradox to try to discuss all of them. Solutions tend to either reject one (or more) of these three claims, or, in the case of Radford, to bite the bullet and argue that since all these claims are true, we are in fact behaving irrationally when we exhibit emotional responses towards fiction. For now, we need only consider how this paradox relates to the idea of involuntary counterparts, which can be seen by introducing Walton's solution to this paradox.

Walton (1978, 1990, pp. 195-204; 241-249) argues that we can make sense of the paradox of fiction by introducing the notion of *quasi emotions*, which are directed towards fictional characters and events. In so doing, Walton can be taken as offering a clarification of claim (1), that we have emotional responses to fiction, and also as introducing a set of imaginative counterparts to the emotions.

He motivates this proposal by introducing the example of a moviegoer named Charles, who is watching a horror film about a malevolent slime who is rampaging across

³⁸ A straw poll of some friends suggested this is not an uncommon response to this opening montage. Some mooted that this might be because it led to them reflecting on actual misfortunes and the vagaries of fate, but it looks like such further reflection isn't *necessary* to feel an emotional response here.

³⁹ This claim is somewhat contentious depending on one's understanding of possibility. An interesting discussion of this sort of issue arises in a debate between Kripke (1980, pp. 156-158) and Dummett (1993) about whether unicorns might have existed in the actual world. For what's it's worth, I would side with Kripke that one cannot, on the grounds that there are no conditions a real-world animal could satisfy that would make it a unicorn.

⁴⁰ Harris (2000 p. 73) notes that sometimes noting something is merely fictional or imaginary can help on this score. We might try to alleviate our apparent fear of a horror movie monster by reminding ourselves we're only watching a movie and the monster cannot really threaten our safety.

the earth. Upon the climax of one of its rampages, the slime turns a single beady eye to the camera, as if to fix his glare directly on Charles, and then makes as if to head towards him. In response, Charles shrieks and clutches his seat tightly. After the film, Charles, self-reports that he was ‘terrified’ of the slime. Walton’s (1990, p. 196) starting question is whether we should take this claim at face value.

We saw in our discussion of fictional truth that Walton maintains that many of our statements about fictions should be interpreted as claims about what is fictionally true. For example, if Charles said, ‘I saw the slime’ we should understand him as saying *I saw the slime* (it is fictionally true that I saw the slime). Walton suggests this is also true of Charles’s emotional report. When Charles says, ‘I felt terrified’ we should understand him as saying *I felt terrified*. Charles is expressing what is make-believe rather than what is true: it is only fictional that he was terrified. However, as we noted earlier, this does not have to mean a counterpart to fear has made this fictionally true, it could be he genuinely fears the slime and this is what makes it fictional that he fears the slime.

To make the case for Charles’ feelings amounting to mere quasi fear in reality, Walton notes that Charles’ responses to his ‘fear’ are somewhat odd. If he really feared impending slime attack, he would be motivated to flee the cinema, or to call the police, or to tell his loved ones to head for shelter, and so on. One explanation for his attenuated motivational response is that he doesn’t believe the slime actually threatens him or anyone else (save the denizens of the fictional world of the film). As such, Walton suggests that it is made fictionally true that Charles fears the slime because in reality he feels ‘quasi fear’: he has entered an involuntary counterpart state that resembles fear but lacks an appropriate connection to belief and motivation (Walton 1990, pp. 201-202).

That being said, this does not mean Walton is arguing we simply do not exhibit emotional responses when engaging with fiction. For example, Walton tells us that:

It goes without saying that we are genuinely moved by novels and films and plays, that we respond to works of fiction with real emotion. ... My ... claim is only that our genuine emotional responses to works of fiction do not involve literally, fearing, grieving for, admiring fictional characters. (Walton 1997/2015, p. 275)

As I read Walton, he is arguing that we do in fact exhibit some sorts of affective responses to fictions – we undergo certain bodily reactions and so on – but these responses constitute ‘mere’ quasi emotions because they arise from our engagement with fiction. Thinking about my tears in response to *Up*, we might say these are genuine tears but not manifestations of genuine sadness: they are representative of my quasi sadness.

That being said, it would be a mistake to regard quasi emotions as *mere* bodily reactions. For Walton, perhaps the most crucial issue is how to understand the object of our fiction-directed emotion: he is concerned with explaining why, for example, it looks like we describe ourselves as literally fearing fictional characters, bearing in mind it seems odd to fear something that cannot possibly threaten you. As such, Walton argues Charles doesn’t just fictionally fear *simpliciter*: he fictionally fears a specific fictional slime because he believes that make-believedly the slime in question is bearing down on him (Walton 1978/2015, p. 260). As such, we can argue that being in a state of quasi fear involves having certain involuntary physiological reactions and certain beliefs about what is fictional. Genuine fear involves those same physiological reactions combined with beliefs about what is actually the case. In Walton’s terminology, this means that real world quasi-fear feelings can make it fictionally true that, for example, you fear a slime. In more neutral terms, when our fear appears to have a fictional intentional object, it can be argued that at least sometimes we have in fact entered into a state of quasi fear.

The most important point here for present purposes is that quasi emotions are not supposed to be voluntary counterparts, since quasi emotions are not supposed to be subject to the will.⁴¹ If Walton is introducing a counterpart here, then quasi fear (and

⁴¹ That being said, Walton (1997) draws a link between his account and simulation theory, which we might think implies some sort of connection between quasi emotions and the voluntary counterparts of simulation theory. However, his point is that quasi emotions might be informative as part of a wider simulation project when it comes to determining how you’d behave in a given situation. His example is of spelunking. While simulating working your way into the depths of a cave, you might have emotional responses (which he suggests are quasi emotions) that tell you how you’d behave if you really had to go

other potential quasi emotions⁴²) are involuntary counterparts: they are automatic responses to engaging with something we recognise as being fictional.

With the distinction between involuntary and voluntary counterparts in hand, we can respond to some common worries raised with Walton's response to the paradox of fiction. Carroll for example, argues that:

One reason to be suspicious of the notion that art-horror is a pretend emotion rather than a genuine emotion is that if it were a pretend emotion, one would think that it could be engaged at will. I could elect to remain unmoved by *The Exorcist*; I could refuse to make believe I was horrified. But I don't think that that was really an option for those, like myself, who were overwhelmingly struck by it. (Carroll 1990, p. 74)

This amounts to arguing that because quasi fear is a counterpart, it should be a state that we can enter into at will and thus that we should be able to elect to remain impassive when watching a scary movie like *The Exorcist*, which doesn't necessarily follow. If we allow for involuntary counterparts, then it is possible for quasi fear to be a sort of state that is not under the control of the will.

That being said, the introduction of quasi emotions is controversial and we should look more closely at how we can justify their introduction. In line with my third principle, we need to ask whether there are a) constraints on some of our emotions which rule out their being involved in our engagement with fiction, or b) puzzles about our engagement with fiction that are best solved by introducing quasi emotions.

Let's begin by considering constraints. Walton allows that quasi fear can on introspection be phenomenologically identical to real fear. Why, then, do these bodily reactions not constitute full-blown fear? As we noted earlier, Walton argues that there is an intimate connection between our emotions, motivations, and beliefs. In the Charles example, he obviously does not believe that the slime actually threatens him. If he did

spelunking. He does not think this means these quasi emotions are under the control of the will.

⁴² Walton doesn't argue every emotion has a quasi-version, which means it would be misleading to imply he thinks that every emotion has a counterpart.

fear an imminent attack, presumably he'd do more than just sit in his chair yelping. However, some philosophers are suspicious of taking this to justify the introduction of quasi emotions. For example, one might think that a state being properly described as 'fear' depends not on the subject's beliefs or motivations, but merely on the physiological reactions occurring in their body, a position historically associated with James (1884, p. 190), who suggests 'we feel sorry because we cry, angry because we strike, afraid because we tremble, and not that we cry, strike, or tremble, because we are sorry, angry, or fearful'. If this sort of account of the emotions is correct, then quasi emotions may just be ordinary emotions by definition. I suspect that in response, Walton would focus our attention on the fact that our somewhat attenuated responses to fictional stimuli still imply that we don't genuinely come to fear the fictional.

However, even if one is sceptical about whether we need to introduce quasi emotions on the basis of constraints on feeling a genuine emotion, there remain many puzzling aspects to our emotional engagement with fiction that need to be explained aside from the paradox of fiction. For example, we need to explain why we sometimes choose to engage with fictions that we know will frighten us, or lead to other unpleasant emotional experiences.⁴³ A philosopher who argues that genuine emotions are involved in our engagement with fictions will still need to make sense of these sorts of puzzles, and Walton (1990, Ch. 7) uses his theory of quasi emotions to try to make sense of many of these issues. As such, a full discussion of whether we should introduce quasi emotions would need to reflect on whether we need to introduce them to make sense of at least some of the puzzles that arise when we reflect on our emotional engagement with fiction.

To give one example of how we might justify the introduction of quasi emotions in this sort of way, we can note that Friend (2003, pp. 41-45) defends the introduction of quasi emotions by focusing on a puzzle that arises when we reflect on fictions that concern real individuals. In the film *JFK*, Oliver Stone presents Jim Garrison – a district attorney

⁴³ Carroll (1990) calls this first problem the 'paradox of horror'. The more general problem is called the paradox of negative affect by Smuts (2009).

who prosecuted the only trial related to JFK's assassination – as an admirable individual trying to fight injustice and conspiracy. Friend notes that this means we might have positive emotional responses towards him and want him to succeed in his task *in the fiction*. However, at the same time, we can be aware that in real life Garrison ruined a man's life when prosecuting this case, and as such have negative emotional feelings towards him in reality. Friend argues that the best way to explain this is to maintain that we have genuine emotional responses towards the prosecutor in the actual world and quasi emotions towards the prosecutor *in the fiction*.⁴⁴

To make a broader point about involuntary counterparts, when determining whether to introduce an imaginative counterpart in the involuntary sense, we must consider whether introducing this sort of counterpart helps to make sense of puzzles related to human activities that seem to involve the imagination, or whether introducing a counterpart simply creates new problems and exacerbates existing ones. We cannot justify the introduction of an involuntary counterpart solely by noting what looks to be an example of a genuine mental state that has fictional content. This is perhaps why some take an instinctive dislike towards the idea of quasi emotions: pre-philosophically it seems acceptable to say that we can exhibit genuine emotional reactions on the basis of engaging with a representation we know to be fictional.

So far, we have seen that imaginative counterparts in both my first and second sense can be plausibly introduced for seeing, and one can be plausibly introduced in at least my second sense for emotion. However, we should not begin our discussion of whether belief has a counterpart by presuming that a principle like 'X is a mental state, therefore X has an imaginative counterpart' holds.

⁴⁴ One might worry here that the prosecutor in the fiction is not the same as the real prosecutor and so our emotions simply have different intentional objects. Friend (2003, p. 46-50) argues against this on the grounds that fictions often ask us to take things to be fictionally true of real individuals, places, and things. For example, it seems odd to say Sherlock Holmes novels are not really set in London but in some fictional place that resembles London.

In the case of voluntary counterparts, we have already seen that the emotions appear to give an example of a state that does not have this sort of counterpart, since we cannot enter into emotion-like states at will. In the case of involuntary counterparts, both the notion of quasi emotions and Walton's notion of imagining seeing are controversial, and bearing this in mind, one could argue that I have failed to pick out a distinctive category of mental states. Instead, it could be argued that sometimes voluntary counterparts behave in an automatic way and that we do not need to introduce a second category of counterpart.

I will return to this sort of concern in relation to belief-like imaginings at several junctures in this thesis. My main response is that if one wishes to broaden the functional role of a counterpart in this way, one will need to give positive reasons for doing so. We should be cautious of moving too quickly from establishing that a given state has a voluntary counterpart to arguing this counterpart is involved in a given human activity but behaves in an automatic manner. As we will see in chapter 2 and chapter 5, this means we have good reason to separate out the question of whether an imaginative counterpart to belief is involved in our engagement with fiction from the question of whether belief has a voluntary counterpart.

Furthermore, there are some mental states and attitudes that do not look like they have counterparts in either the voluntary or involuntary sense, such as *amusement*. Both Walton (1994, pp. 43-44) and Currie and Ravenscroft (2002, pp. 189-191) agree that you can be genuinely amused by the merely fictional, so it looks like amusement having fictional content is unproblematic and we won't need to introduce a counterpart to make sense of this.⁴⁵

⁴⁵ I address this point at greater length in my MPhil dissertation (Davis 2015, pp. 31-37). One might object that there is something puzzling about your amusement being *caused* by something merely fictional because fictional entities lack the requisite causal powers. In this way, we might think we need to introduce an involuntary counterpart to amusement to explain this causal gap. I think this is mistaken because this simply serves to increase puzzles in this area, and we can equally explain this causal gap by suggesting the literal cause of your amusement is a story or narrative rather than a fictional individual.

Conclusion

We have seen that deciding whether our ability to propositionally imagine involves a counterpart to belief relates to a wider question about what sorts of mental states have imaginative counterparts. In sections 1.2 and 1.4, we saw that there are two notions of imaginative counterparts at play in contemporary philosophical discussions of the imagination: one of a state that resembles another but is under the control of the will, and one that is a more automatic response to engaging with something fictional or imaginary.

In sections 1.3, 1.5 and 1.6 we attempted to clarify the nature of these counterparts by considering when we are justified in introducing them. We did this by considering whether seeing and the emotions have either voluntary or involuntary counterparts. On the basis of these discussions, I argued for four claims. Firstly, I suggested that we can justify introducing a voluntary counterpart by reflecting on whether a state is similar to another state in introspective, functional and neurological ways. Secondly, I argued that we cannot justify the introduction of an involuntary counterpart by merely noting that we can enter into a given mental state as a result of engaging with fiction. Thirdly, in order to justify introducing an involuntary counterpart, we need to consider whether there are constraints on the state in question, or puzzles that arise when we reflect on a particular human activity if we do not introduce this sort of counterpart. Finally, we saw that we can deny that a state has one sort of counterpart, but accept that it has the other sort of counterpart.

In the next chapter, we will see that some philosophical accounts of belief-like imaginings cut across my distinction between these two sorts of counterpart. I will argue that this gives us some reason for thinking that we should reject the uniformity assumption when it comes to explaining the attitude involved in our engagement with fiction and in pretence.

Chapter 2: Distinct Attitudes & Distinct Contents

Introduction

As we noted in the introduction, it is widely agreed that we have an ability to propositionally imagine things. In recent years, it has been popular to defend the idea that this ability involves taking a *distinct attitude* towards p . This attitude is said to be a counterpart to belief, and so when it is said this attitude is involved in fiction and pretence, this amounts to accepting the *non-doxastic* assumption. Defenders of this view also tend to adopt the *uniformity* assumption, and argue that this single distinct attitude is involved in a wide range of human activities, such as hypothetical reasoning, fiction, mindreading and pretence. However, this is not the only way we can explain our apparent ability to propositionally imagine things. Instead, what philosophers have called propositional imaginings could involve a variety of different mental attitudes in different contexts.

To develop the arguments of this chapter, it will first be helpful to briefly set out four different ways in which we might explain our ability to propositionally imagine things (though since we will question the uniformity assumption, they do not have to be taken as mutually exclusive).

1. The Single Attitude Approach

This sort of view holds that what we call propositional imaginings are in fact merely ordinary beliefs. This view is unpopular for two reasons. The first can be seen by returning to the non-doxastic assumption that I mentioned in the introduction. It seems intuitive to think that there is at least some sort of belief-like attitude that we can take towards propositions: I can believe that I am on a beach; but I can also *suppose* that I am on a beach. The second can be seen by reflecting on the uniformity assumption that I mentioned in the introduction. If we argue that engaging in pretence and with fiction just involves simple beliefs, then this suggests that children and consumers of fiction face

cognitive confusion (e.g. children might come to believe that they are dogs during pretence, consumers of fiction might come to believe that monsters are real).

2. The Distinct Attitude Approach

These theories tell us that the propositional imagination is underpinned by a *distinct attitude*. Propositionally imagining p is held to be an attitude distinct from believing that p , but is an attitude which shares some similarities with belief. It is therefore an imaginative counterpart to believing.

3. The Distinct Content Approach

This sort of view holds that at least some of the states that we call propositional imaginings are in fact beliefs with special *contents*. There are at least three versions of this approach in the existing literature on imagination, pretence and fiction. The first is developed by Langland-Hassan, who argues that propositional imaginings are *counterfactual* beliefs. The second is developed by Leslie, who suggests that beliefs of the form ‘I PRETEND that “I am a dog”’ are what explain childhood pretence. The third is the view that our cognitive attitude towards fiction is beliefs under the scope of a *fictional operator*. According to this view, if I indicate that I think it is fictionally true that ‘Harry Potter is a wizard’, this can be understood as me saying that ‘I believe Harry Potter is a wizard’ [in the fiction]. The most recent, and also sophisticated, defence of this view is offered by Tullmann (2016), and several philosophers have defended this approach in relation to the paradox of fiction, suggesting that these sorts of beliefs are what lead to us exhibiting emotional responses (e.g. Neill 1993, Davies 2009).

4. The Distinct Attitude and Distinct Content Approach

This sort of approach would argue that propositionally imagining p sometimes involves both a distinct attitude and distinct contents. For example, we could argue that we take a non-doxastic belief-like attitude towards p , but that p is nonetheless subject to some sort of fictional operator. The main motivation for adopting this sort of view would be if

the distinct attitude and distinct content views both fail to explain a human activity that seems to involve our ability to propositionally imagine. However, if one of these views alone can do the requisite explanatory work, the appeal of this sort of conjoint view will be diminished. As such, I will not say much more about this approach in this thesis. That being said, some have suggested that this is how Leslie's view of pretence should be understood (e.g. Currie 1998, p. 41), and I will argue against developing Leslie's view in this way in chapter 4.⁴⁶

These different approaches to understanding the propositional imagination relate to a distinction between *force* and *content* associated with Frege. This is roughly a distinction between the attitude we take towards a given content, and the content itself. The consensus view holds that we need to introduce the notion of an 'imaginative force' to explain what it is to propositionally imagine something. I am instead going to focus on the relevance of content, and argue that a distinct content view offers the best account of the attitude involved in fiction and pretence.

We will begin in section 2.1 by considering whether belief has a voluntary counterpart. I will argue that it plausibly does, and that this counterpart can be associated with

⁴⁶ There are at least two further accounts that we could discuss here. The first would be a *really distinct attitude* approach, where we argue that some propositional imaginings are not counterparts to belief or any other kind of mental state. The second would be to argue that imagining or believing *p* involves taking a *further attitude* towards *p*, such as judging that *p* is fictional or true. For example, the difference between imagination and belief could depend on whether we judge that a given representation concerns the real or fictional, or whether we apply a norm of truth to the representation (Shah and Velleman 2005). The first sort of approach is interesting but unhelpful in the context of fiction and pretence. The reason that philosophers have been drawn to the idea that the attitude involved in fiction and pretence is a counterpart to belief is because the cognitive attitude involved in them seems to resemble belief. The second is compellingly criticised by Sinhababu (2013, 2016), and I share his concerns. As he rightly brings out, any purported further attitude or norm involved in belief could equally be applied to our imaginings. For example, one could apply a norm of truth to one's imaginings but this wouldn't make them into beliefs (Sinhababu 2013, pp. 155-157).

notions such as *supposing*. I will then bring out some of the initial reasons for why we might want to associate this attitude with fiction and pretence.

In section 2.2, I will further bring out why philosophers have been drawn to the view that a counterpart to belief is involved in fiction and pretence, and thus to make the non-doxastic assumption. This will be done by setting out why a simple belief account of pretence and fiction is problematic. The crucial difficulty with this approach is that if we literally believed the contents of fictions or of our episodes of pretence, this would lead to chaos in our representational systems.

In section 2.3, I will set out two different theories that seek to offer detailed sketches of the nature of belief-like imaginings: simulation theory and the cognitive theory of Nichols and Stich. Both of these theories adopt the uniformity assumption, and having set out these theories, I will then bring out some initial reasons for why we might want to question this assumption. We will do this by reflecting on the different ways that belief-like imaginings appear to behave when we engage in hypothetical reasoning as compared to when we engage with fiction or in pretence.

In section 2.4, I will set out three distinct content approaches that seek to make sense of our engagement with fiction and in pretence: the counterfactual view of Langland-Hassan (2012), the meta-representational view of Leslie (1987, 1994), and the fictional operator view. For the most part, we will just set out these views without passing judgement, though I will take this opportunity to argue that Langland-Hassan's view will struggle to make sense of how we engage with fiction.

Finally, in section 2.5, I will consider whether we have any *prima facie* reasons to favour a distinct attitude view of fiction and pretence over a distinct content view and thus to embrace the non-doxastic assumption. We will do this by reflecting on whether introspective, functional or neurological issues suggest that belief is not the only cognitive attitude involved in our engagement with fiction and in pretence. I will argue that none of these considerations is compelling. This will set up a standoff between the distinct content and distinct attitude views of fiction and pretence, which we will attempt to resolve in the remaining chapters of this thesis.

2.1 A Counterpart to Belief

As we saw in the previous chapter, there are two notions of a counterpart at play in philosophical discussions of the imagination. The first is a mental state that is under the control of the will and bears some similarities to another mental state. The consensus view of the nature of propositional imaginings argues that they are a counterpart to belief in this sense. The consensus view maintains that the attitude underpinning the propositional imagination is under the control of the will and bears important similarities to belief, but has a distinct functional role and is a distinct attitude. Neurological and introspective similarities are also referenced to justify this claim, as one would expect based on my sketch of this sort of voluntary counterpart.

Broadly speaking I think this is right. That is, I think that there is some sort of imaginative counterpart to belief that we can enter into at will and so is a voluntary counterpart. There are various words we use in ordinary language to mark out this counterpart. For example, we can speak of entertaining the proposition that p or supposing that p and thinking that p (in the sense of ‘thinking about X ’, not taking something to be true).⁴⁷ These sorts of attitudes can be distinguished from belief introspectively and functionally. It would be counter-intuitive, for example, to suggest that when I’m sitting at my desk and decide to entertain or suppose that p , this is in some sense just to believe that p , even if this belief has a special content. This sort of state also seems to bear some functional similarities to belief. For example, supposing looks to play a role in allowing

⁴⁷ One might worry here that these are three different sorts of mental states. I would be willing to accept this and not much will turn on this point in the remainder of the thesis. My only claim here is that I’m happy to allow that these terms refer to an imaginative counterpart to belief. My scepticism relates to linking this sort of counterpart to pretence and fiction. One might also worry that these sorts of attitudes are in fact not counterparts to belief. Arcangeli (2014) argues that supposition should be thought of as a form of imagination that is an imaginative counterpart to *acceptance*. One could also think ‘entertaining’ is a more basic sort of attitude than imagining, perhaps associated with a Fregean notion of ‘grasping a thought’ (Frege 1956). My point here is only that I do not want to rule out that these sorts of attitudes might be counterparts to belief.

us to engage with hypothetical arguments, since our suppositions can interact with our standing beliefs. If our suppositions didn't bear some relation to our beliefs, this would raise serious questions about the reliability of our hypothetical reasoning. However, just because we can make sense of there being some sort of counterpart to belief involved in supposing and entertaining propositions, this does not mean that we must accept that this counterpart plays a crucial role in explaining how children can engage in pretence, or how we can engage with fiction.

To bring out why the consensus view argues for this claim, we can note that there is widespread agreement among philosophers that there is an intimate link between imagination, fiction and pretence. One reason for endorsing there being this sort of connection is that we often associate fiction and imagination when discussing fiction in a non-philosophical context, for example we might describe a work of fiction as stimulating the imagination in everyday conversation.⁴⁸ The same is also true of pretence, and indeed in London there is an Institute of Imagination which focuses on the importance of 'play' among other aspects of childhood associated with the imagination, such as making and creating. However, while this might point towards there being a close connection between some sort of imaginative capacities and our engagement with fiction and in pretence, this sort of common-sense intuition does not reveal why philosophers have argued that *belief-like* imaginings play an important role in our engagement with fiction and in pretence.

We can note two things to bring out why philosophers have been keen to stress the importance of belief-like imaginings for engaging with fiction and pretence. The first is that, as we noted in section 1.1, we do not need to form non-propositional imaginings, such as mental images, to engage with works of fiction or in pretence. While watching a

⁴⁸ In a recent lecture, the author Neil Gaiman makes this link by suggesting: 'When you watch TV or see a film, you are looking at things happening to other people. Prose fiction is something you build up from 26 letters and a handful of punctuation marks, and you, and you alone, using your *imagination*, create a world and people it and look out through other eyes. You get to feel things, visit places and worlds you would never otherwise know.' (emphasis mine) (<https://www.theguardian.com/books/2013/oct/15/neil-gaiman-future-libraries-reading-daydreaming>)

film, I do not need to form mental images of the on-screen events in order to engage with the narrative, and while reading a novel I do not necessarily have to form mental images of the described events.⁴⁹ In the case of pretence, if a child is pretending to be a dog, presumably they don't necessarily have to form a mental image of a dog or of themselves as a dog.

Secondly, as we saw in section 1.4, fictions contain (or imply) propositions that don't appear to invite straightforward belief. Upon reading the Sherlock Holmes novels, the reader is not supposed to believe that it is literally true that there exists an individual called 'Sherlock Holmes' who is a detective. Instead, we are supposed to understand this as being *fictionally true*. Several philosophers have followed Walton in embracing the idea that a proposition is fictionally true if we are mandated or prescribed to *imagine* it (e.g. Currie 1990). Since this sort of imagination takes a fictionally true *proposition* as its content, the sort of imagination associated with fictional truth is standardly argued to be belief-like. Since Walton in turns associates his theory with pretence, we can also make the same point to defend associating pretence and belief-like imaginings.

A further reason for why this consensus view has formed can be brought out by considering why a simple belief-based account of how we engage with fiction and in pretence will be difficult to make work.

2.2 A Single Attitude

The simple belief proposal will face an immediate difficulty when we think about the role that propositional imaginings are supposed to play in making sense of childhood pretence. One thing we have to explain when offering an account of pretence is how children are

⁴⁹ Of course, this is not to say that we never form mental images when engaging with fictions. The point here is that it seems coherent for a reader to read and understand a novel without forming any mental images. That said, a study by Brooks (1967), referenced by Matravers (2013, p. 72) suggests that readers have difficulty reading a text and visualising the events described in the text at the same time.

able to maintain what we can call ‘cognitive order’. We need to explain why forming representations related to their pretence doesn’t lead to chaos in their representational system. Leslie sets up this problem by asking:

How is it possible for a child to think about a banana as if it were a telephone, a lump of plastic as if it were alive, or an empty dish as if it contained soap? If a representational system is developing, how can its semantic relations tolerate distortion in these more or less arbitrary ways? Indeed, how is it possible that young children can disregard or distort reality in any way and to any degree at all? Why does pretending not undermine their representational system and bring it crashing down? (Leslie 1987, p. 412)

Consider a child pretending to be a dog, who is running around the house making barking noises. If their pretence is underpinned by the belief *that they are a dog*, it looks like this would lead to chaos in how they represent the world. Suddenly the child would either hold contradictory beliefs – that they were and were not a dog – or they would replace some existing beliefs about being a human with the belief that they were a dog.⁵⁰

As such, if we argue that propositional imaginings are beliefs and don’t give this claim any qualification, we will end up in a position where children straightforwardly believe that they are dogs, or that bananas are telephones, or that there is real tea in their cups, and so on. This is a counter-intuitive conclusion to find ourselves committed to. Indeed, empirical studies focused on children’s understanding of pretence demonstrate that from a young age, children are able to keep track of the distinction between the real and the imaginary, at least to some extent (Wellman and Estes, 1986, Estes, *et al.* 1989).

A similar problem also arises when thinking about fictions. If moviegoers watching a dystopian movie like *Children of Men* really believed the events depicted on screen were

⁵⁰ Children can engage in pretence from a young age, and one might find it implausible that young children can have beliefs about being human that could be replaced by beliefs about being a dog. In terms of their internal representations, the relevant beliefs here might be demonstrative. The child used to believe that they are *like that* (human) and now believe that they are *like this* (a dog).

happening, they'd presumably become panicked about the fact that the human race appeared doomed.⁵¹ Furthermore, the contents of some fictions straightforwardly conflict with our standing beliefs. Thankfully upon watching *Scott Pilgrim* we don't start believing that videogame user-interface elements can occasionally pop up in reality.⁵²

We can modify the proposal that the sorts of propositional imaginings associated with fiction and pretence are just beliefs by suggesting that they are *temporary* beliefs. For example, the child pretending that they are a dog might temporarily believe that they are a dog during their game of pretence, rather than this ending up as a stable belief about the world. This relates to a proposal about how to make sense of our engagement with fiction that tells us we 'suspend our disbelief', an idea associated with Coleridge (1817). However, this sort of illusion based account of how we engage with fiction has been thoroughly, and compelling, critiqued by Schaper (1978) and Carroll (1990, pp. 63-68).

The crucial explanatory failing of these sorts of approaches is that they still leave us facing a disconnection between belief and behaviour. If a child ever came to believe they were a dog, it wouldn't matter how temporary this belief was: they would surely do more than just wander around the house periodically making a 'woof' noise, and would likely engage in some dangerous behaviour suitable for dogs but not children, such as barking at dogs in the park or trying to chase them.

The worry about our motivational responses to fictions also remains. As we suggested in section 1.6, if we temporarily believed in the existence of horror movie monsters, then instead of yelping or hiding behind the sofa, we would likely respond to horror films very differently. For example, we might call the police or try to immediately leave the cinema

⁵¹ This film depicts a dystopian future where women are no longer able to become pregnant, leading to mass riots and widespread disorder. The protagonist is tasked with guiding the last pregnant women on Earth to a secret research facility where she will be protected from the outside world.

⁵² For example, at one stage in the film the protagonist – Scott Pilgrim – 'dies' upon being stabbed by his recently acquired arch-nemesis. Thankfully, he finds a '1-up' floating in space above him and upon grabbing it returns to life.

to flee from the monster's advance. Introducing the notion of a temporary belief, or the suspension of disbelief, will still leave us without a satisfactory explanation of the sorts of attenuated behaviours people exhibit when engaging in pretence and with works of fiction.

What we need here is an account of how children and adults can consistently separate out what they take to be true of a fiction or of an episode of pretence from what they straightforwardly believe to be true of the world. An account that reduces imaginings to beliefs *simpliciter* leaves it unclear how to make sense of this.⁵³

This means we are left with at least two sorts of approaches for responding to these difficulties. Firstly, we can take the consensus view and argue that a distinct attitude of belief-like imagination is involved in fiction and pretence. Secondly, we can argue that belief remains the relevant attitude, but that we need to give a different specification of the content of the relevant beliefs. We will begin by considering how the former view has been developed.

2.3 A Distinct Attitude

There are two related theories about the nature of belief-like imaginings that associate them with fiction and pretence and argues that they are a voluntary counterpart. We hinted at the first one in chapter 1 when we noted that there is a notion of 'simulated' belief introduced by simulation theorists like Currie and Ravenscroft and Goldman. To defend the introduction of this notion, they bring out some functional similarities between

⁵³ Another, more sophisticated version, of the just belief view has been put forward by Schellenberg (2013). In her view, propositional imagination and belief are on a *continuum*. In effect, this means that she argues for a widening of our conception of belief. Since my eventual view of fiction and pretence relies on beliefs with distinct contents and will not result in us reconfiguring our folk conception of belief, I will take it to be preferable to this revisionary account. Furthermore, the primary motivation for adopting this view is to explain certain aspects of imaginative immersion, and Liao and Doggett (2014) offer convincing arguments for why this proposal is unnecessary for explaining these issues.

belief and what look to be belief-like imaginings. For example, Currie and Ravenscroft note some ways in which belief-like imaginings resemble beliefs (e.g. in how they preserve inferential links) and also note some differences (e.g. in how they don't aim to reflect what's true). In relation to the issue of inferences, they tell us that:

What is sometimes called imagining-that is belief-like imagining. An important feature of beliefs is their occupation of characteristic inferential roles; believing something tends to lead to believing other things, depending on what else you already believe. Imagining that you are famous is belief-like partly because it mirrors the inferential role of the belief that you are. (Currie and Ravenscroft 2002, p. 12)

Goldman calls this sort of imagination 'S-imagination' (suppositional-imagination) and tell us that:

S-imagination is typically formulated with a 'that'-clause, 'X imagines that p', where p can refer, unrestrictedly, to any sort of state-of-affairs. To S-imagine that p is to entertain the hypothesis that p, to posit that p, to assume that p. Unlike some forms of imagination, S-imagination has no sensory aspect; it is purely conceptual. (Goldman 2006b, p. 42)

Goldman goes on helpfully to distinguish this sort of imagination from what he calls 'E-imagination' (enactment-imagination):

Enactment-imagination is a matter of creating or trying to create in one's own mind a selected mental state, or at least a rough facsimile of such a state, through the faculty of imagination. Prime examples of E-imagination include sensory forms of imagination, where one creates, through imagination, perception-like states. Acts of visual and auditory imagination, which involve the production of vision-like or hearing-like states, are familiar types of E-imagination. Another type of E-imagination is motor imagination, where one produces action-directed representational states, without intending to execute the selected action. The term 'imagery' is commonly applied to these cases; there is visual imagery, auditory imagery, and motor imagery. (Goldman 2006b, p. 42)

Goldman also suggests that this means we might be able E-imagining believing, which may in turn mean that S-imagination reduces to E-imagination since we could argue that:

E-imagination is the fundamental kind of imagination, and that S-imagination is simply one species of it. ... the species in which the mental state enacted is belief. Supposing that p is E-imagining believing that p. (Goldman 2006b, p. 44)

Goldman and Currie and Ravenscroft also both accept the uniformity assumption, and argue that these belief-like states allow us to engage in hypothetical reasoning and also play an important role in our engagement with fiction and in pretence.⁵⁴

The second theory that seeks to link belief-like imaginings to pretence and fiction was originally introduced by Nichols and Stich (2000, 2003), and christened the ‘cognitive’ theory of propositional imagination by Nichols (2004, p. 129), who tells us that ‘Stephen Stich and I developed a cognitive theory of the imagination’.⁵⁵ The theory has been elaborated by Nichols in various single-authored papers (2004, 2006) and has received some friendly emendations from Weinberg and Meskin (2006b). It has also been given several sympathetic amendments by philosophers such as Doggett and Egan (2007, 2012) (who argue that someone adopting this approach should introduce something like the simulationist notion of desire-like imaginings into an account of fiction and pretence). Defenders of this sort of cognitive view also tend to endorse the uniformity assumption, and argue that these belief-like imaginings are involved in hypothetical reasoning, fiction and pretence, among other activities like mindreading.

Nichols and Stich’s cognitive account of the propositional imagination is embedded in a commitment to *representationalism* and *functionalism* about the nature of the mind. In relation to representationalism, they hold that our beliefs and desires, (and perhaps other propositional attitudes we might think need to be accounted for in our cognitive architecture) are ‘representational states’. In relation to functionalism, they accept that

⁵⁴ Some further simulation defences of the uniformity assumption can be found in Currie (1995), which makes the case for associating this sort of simulated belief-like state with fiction and pretence, and Gordon and Baker (1994), which offers a simulation based theory of pretence.

⁵⁵ The reader might at this point wonder why the term ‘cognitive’ is used since this is usually associated with beliefs, rather than imagination, in areas like meta-ethics. As Nichols and Stich use the term, it can be thought of as indicating that the theory concerns the psycho-functional workings of the mind and the notion of a representational mental state.

for a mental state to be classified as a belief, as opposed to a desire, it must serve a particular functional role in our cognitive architecture. They illustrate this idea with the metaphor of ‘boxes’ that contain functionally distinct representations in the mind. They note that most philosophers would accept that the human mind contains a belief ‘box’ that contains our belief representations and a desire ‘box’ that contains our desire representations, where the representations in these two boxes can be distinguished from one another with reference to their differing functional roles.⁵⁶ As they put the combination of these ideas:

[t]o believe that Socrates was an Athenian is to have a representation token whose content is Socrates was an Athenian stored in one’s ‘Belief Box’, and to desire that it will be sunny tomorrow is to have a representation whose content is It will be sunny tomorrow stored in one’s ‘Desire Box’. (Nichols and Stich 2003, p. 15).

These boxes are then postulated to interact with other elements of our cognitive architecture, such as ‘inference mechanisms’ and ‘affect generation systems’, with each box interacting with different bits of mental architecture to reflect their differing functional roles.⁵⁷

Nichols and Stich’s argue that in order to account for our ability to propositionally imagine, we need to introduce a further box, which we can call the *imagination box*.⁵⁸

⁵⁶ A philosopher may reject this metaphor of boxes as explanatorily unhelpful, but the point here is most philosophers would accept that we should distinguish between beliefs and desires, and it is not an uncommon approach to make this distinction with reference to their functional roles.

⁵⁷ This talk of boxes is not supposed to imply that our beliefs or desires will turn out to have a single location in a neural map of the brain. Weinberg and Meskin (2006b, p. 179) make this explicit, telling us that ‘[o]ur talk of ‘boxes’ ... should not be assumed to have neurophysiological implications’ in the course of their development of Nichols and Stich’s theory. Boxes are merely supposed to separate out aspects of our mental life by functional role, without making any claims about what realises these functional roles at a neural level. Indeed, it is perhaps more plausible to propose that several cognitive mechanisms underpin what Nichols and Stich label the ‘affect generation systems’ and ‘decision making systems’.

⁵⁸ This box has been given a variety of names. Nichols and Stich (2000, 2003) call this box the ‘possible worlds’ box or PWB for short (due to its supposed involvement in thinking

The imagination box is distinct from the belief and desire boxes since the representations contained within the imagination box have a distinct functional role.

As they put the point:

[the imagination box] contains representation tokens. However, the functional role of these tokens – their pattern of interaction with other components of the mind – is quite different from the functional role of either beliefs or desires. Their job is ... to represent what the world would be like given some set of assumptions that we may neither believe to be true nor want to be true. (Nichols and Stich 2003, p. 28)

This amounts to something similar to the notion of belief-like imaginings introduced by simulation theorists, but contains no commitment to there being other sorts of imaginative counterparts, or to a broad notion of simulated mental states. This distinction between the two approaches is highlighted by Goldman where he notes (in relation to the issue of how imagination is involved in our engagement with fiction):

[t]here is a substantial difference between the view that S-imagination covers all important uses of imagination in the consumption of fiction and the view that E-imagination (even in its non-suppositional variants) is essential to the consumption and appreciation of fiction. (Goldman 2006b, p. 44)

One way to read this sentence, is that Goldman understands the cognitive view as arguing that the only kind of imagination relevant to human activities like engagement with fiction is belief-like imagination. He then argues that this is misguided since we need to also allow for a variety of counterpart states to play a role here, such as desire-like imaginings. On the other hand, desire-like imaginings are explicitly rejected in various articles and books by cognitive theorists, such as Nichols and Stich (2003) and Weinberg (2013).

The main difference between the two views can be seen by returning to the idea of boxes. The simulation theorist can be understood as saying that there are no special

about hypothetical and counterfactual situations), and Nichols (2004) calls it the ‘pretence’ box (due to its role in allowing us to engage in pretence). I am choosing to use the name ‘imagination box’ that was introduced by Doggett and Liao (2014) since this best captures the wide role this box is supposed to play in our mental life.

imagination boxes, instead we have a general ability to take our ordinary boxes ‘offline’. Nichols and Stich and their followers do not commit themselves to a general ability to take boxes offline, and so they just introduce belief-like imaginings without committing to there being any further imaginative counterparts. If the simulation theorist did this, they would end up introducing a wide variety of new cognitive boxes when instead they can propose a general process that allows us to take our existing boxes offline.

Nichols (2006, p. 459) highlights this consensus about the existence of a distinct attitude that underpins the propositional imagination, and suggests that this proposal is ‘perhaps the most productive idea about the imagination that anyone has ever had’. I will not challenge Nichols’s claim about the emergence of this consensus, and he is right to say that this idea has been productive. Philosophers have suggested that these sorts of belief-like imaginings play a key role in a wide range of human activities, such as mindreading, pretend play, engagement with fiction, counterfactual reasoning and modal reasoning, and a unified account of these activities would constitute major philosophical progress if it were widely accepted.

In light of the explanatory promise of this way of thinking about the propositional imagination, Weinberg and Meskin (2006b, p. 177) suggest that we should shift our methodology for thinking about these human activities away from a ‘paradox and analysis’ model towards a ‘phenomenon and explanation’ model, where we make sense of their curious features within the framework of Nichols and Stich’s approach to the imagination. They argue that we should do this since:

[w]hen philosophers confront a puzzling set of propositions, their traditional approach is to look for a way to reconfigure the concepts deployed in the propositions, casting about for formulations that look sufficiently natural and motivated from the perspectives of metaphysics and folk psychology, and which dissolve the apparent contradiction. We advocate instead that the puzzling propositions be treated merely as initial descriptions of a set of phenomena for which the philosopher must now seek a good explanation. (Weinberg and Meskin 2006, p. 177)

They demonstrate this by considering the paradox of fiction. We saw in chapter 1 that this arises from our acceptance of three claims:

1. We exhibit affective responses to works of fiction.

2. We do not believe that fictions depict real events and happenings.
3. Our emotions are subject to a coordination condition (Gendler and Kovakovich 2005, p. 241). In order to exhibit an affective response towards something, we normally need to believe it is real.

We noted earlier that most attempts at resolving this paradox proceed by clarifying one of these three claims (e.g. Walton's quasi emotion solution works by offering a clarification of the first claim). With the cognitive theory of imagination in place, Weinberg and Meskin (2003, 2006, pp. 183-184) argue that we should view this paradox as a phenomenon for the cognitive theory to account for, rather than a puzzle to be gerrymandered away by sharp philosophical distinctions about the nature of the emotions. On the cognitive view, since imaginings are belief-like, this means that they can activate our affective systems in a similar way to beliefs. They argue this will provide us with the philosophical tools needed to resolve the paradox since:

[t]he functional similarity but non-identity between believing and imagining helps explain ... the ambiguous nature of fiction-driven affect. To the extent that we focus on the fact that many ordinary emotional responses are caused by beliefs, but affective responses to fiction are caused by distinct cognitive states (i.e. imaginings), and moreover, that belief-caused affect may interact with our motivational and action-production systems in a way that imagination-caused affect may not—to the extent that we focus on such differences, we will tend to resist characterizing the latter as full-fledged emotions. But to the extent that we focus on the functional similarity between belief and imagination, as well as the phenomenological and biological similarities between fictive and non-fictive affect, we will be pulled towards assimilating fictionally driven affect to the general category of the emotions. (Weinberg and Meskin 2006b, p. 184)

This application of the cognitive theory highlights one of the key merits of Nichols and Stich's approach: in order to explain the different functional roles of beliefs and belief-like imaginings they set out, in detail, the supposed similarities between them, along with their differences. In chapter 3, we will consider the specific claims of Nichols and Stich's cognitive theory to see whether it offers a good explanation of how children are able to engage in pretence. For now, we will consider whether we should challenge the general idea that the propositional imagination is underpinned by a single distinct attitude of voluntary belief-like imagining.

At this stage in the discussion, it might look like this distinct attitude approach to the propositional imagination is remarkably anodyne. Surely it is just obvious that imagining that p involves taking a distinct attitude towards p , one might think, and our interest should be in setting out how this attitude behaves rather than in debating its existence. Indeed, I even said in section 2.1 that I am happy to allow that there is some sort of distinct attitude of entertaining, supposing or thinking that p that can be understood as being a voluntary counterpart to belief. In response to this sort of point, we should note that the reason for this view appearing uncontroversial is that most philosophers will want to endorse something like the following claim:

PA: Propositional imagining p involves taking some sort of propositional attitude towards p

PA is hard to deny as presented: *prima facie* it seems right that our ability to propositionally imagine things involves taking *some* sort of propositional attitude towards p . The consensus view, however, goes beyond PA in what it says about the psychological states that grant us our ability to propositionally imagine things. The consensus view also maintains that the attitude PA refers to is best thought of as a distinct cognitive attitude which is a counterpart to belief, and that this particular attitude plays a role in a wide range of human activities. As such, the theorists we've discussed so far would endorse a more specific claim about the propositional imagination:

BA: Propositionally imagining p involves taking a distinct belief-like attitude towards p

PA is a mere platitude since it amounts to saying that humans have the ability to propositionally imagine things, and that this ability involves some sort of propositional attitude. It is not quite so clear why we ought to assent to BA and accept that this ability is explained by a *single* attitude which is a counterpart to belief. It is not clear, in other words, why we should accept the uniformity assumption and the non-doxastic assumption.

We can introduce a worry about accepting the uniformity assumption by reflecting on some remarks made by Kind (2013) about why it might be problematic to place too

much explanatory weight on the shoulders of a single imaginative attitude. Kind notes that belief-like imaginings⁵⁹ have been suggested to perform very different roles in mindreading, engagement with fiction, pretend play and hypothetical/modal reasoning. She suggests that even if we allow for belief-like imaginings to have a broad functional role, it will be difficult to give a plausible sketch of the nature of this sort of attitude.

For example, Currie and Ravenscroft (2002, p. 35) identify belief-like imaginings as the attitude involved in supposing something to be the case, noting that ‘On our view ... supposition is belief-like imagining’.⁶⁰ However, ordinarily we don’t think mere supposing or entertaining is enough to generate emotional responses, yet we do seem to exhibit affective responses to works of fiction and games of pretend (e.g. we don’t shed tears over Mary’s predicament when she is locked in a black and white room, but we may well shed tears when watching *The Room*).⁶¹ In light of this, Kind suggests that:

[t]he mere act of imagining is not enough to produce affect; rather, we must be employing some particular kind of imagining when we emotionally engage with fiction. ... In short, we ... see that there is nothing about the imagination itself that allows it to play all the different explanatory roles that it has been assigned. (Kind 2013, p. 14)

⁵⁹ Kind refers to imaginings *simpliciter* but her examples concern the sorts of human activities philosophers like Nichols and Stich would associate with belief-like imaginings.

⁶⁰ Weinberg and Meskin (2006b, pp. 191-199) offer a somewhat different account of the relation between belief-like imagining and supposing. They note various differences between these states, and argue suppositions are best thought of as a *subset* of belief-like imaginings, rather than making a simple identification between the two along the lines of Currie and Ravenscroft. This raises the question of why they don’t think supposing is a distinct mental attitude from belief-like imaginings. Their argument on this score is complex, but it boils down to the idea that there is more in common between supposition and imaginings than between belief-like imaginings and belief, and the differences do not warrant saying there are two different attitudes involved here.

⁶¹ Currie and Ravenscroft (2002, pp. 33-38) respond by introducing desire-like imaginings, and argue that these imaginings are involved in our engagement with fiction but not hypothetical reasoning. A good response to these arguments can be found in Nichols (2004b) and Kind (2011)

We could bite the bullet here and offer a series of complex disjuncts to explain the functional role of belief-like imaginings, such as allowing that they sometimes cause affective responses and sometimes do not cause affective responses, and so on. However, the more disjuncts we introduce, the more we risk defending the introduction of a gerrymandered mental state, whose complex specification is only justified by a desire for a unified account of activities like hypothetical reasoning, pretend play and our engagement with fiction. If we accept this sort of worry, then positing multiple attitudes that underpin the propositional imagination will be more plausible than introducing only one.⁶²

The theoretical importance of these discrepancies in how belief-like imaginings seem to behave comes to the fore when thinking about our engagement with fictions. As we noted in the previous chapter, when we read a novel, or watch a film, and so on, the cognitive attitudes we take towards the content of the work will be, to some extent, automatic and spontaneous. One doesn't have to consciously make a decision to 'imagine' the contents of works of fiction. This opens up the possibility that if we introduce an imaginative counterpart to make sense of our engagement with fictions, it will be an involuntary counterpart. As such, in line with my fourth principle about counterparts, this suggests that we cannot take the fact we can associate a notion of belief-like imaginings with a notion of supposing as offering some justification for associating a belief-like attitude with our engagement with fiction.

There is thus good reason to challenge the uniformity assumption and question whether there is a single attitude that can explain all the human activities in which philosophers have argued propositional imaginings play a role. We will now turn to consider the potential merits of allowing for a role for beliefs with distinct contents in our

⁶² Alternatively, we might think that it was a mistake to make room for a special category of propositional imaginings in the first place. For example, Kind (2013 p. 30) notes that philosophers don't always regard the imagination as producing 'states of fundamentally different types' when they discuss the distinction: 'When philosophers distinguish propositional imaginings from objectual imaginings, they do not consider an imagining with propositional structure to be a substitute or pretend belief'.

engagement with fiction and pretence. As I said in the introduction, these views don't have to be understood as arguing that propositional imaginings are not involved in these two activities. Instead, they can be understood as offering an alternative explanation of what is going on internally when someone is engaging in what philosophers describe as propositional imagining.

2.4 Distinct Contents

2.4.1 Counterfactual Belief

The first distinct content approach we will discuss is developed by Langland-Hassan (2012) to explain how children are able to engage in pretence. He argues that what philosophers have called 'propositionally imagining', at least in the context of pretence, amounts to asking oneself questions about what would happen if something was the case, then calling upon generalisations to answer this question. On the basis of these generalisations, we then form counterfactual beliefs. In light of this, he argues that no special attitude is needed to explain pretence: we just need to allow that children have an ability to call upon stored generalisations and form counterfactual beliefs. As he puts the point:

My view, in a nutshell, is that imagining that p amounts to making judgments about what would likely happen if p , from retrieved beliefs in relevant generalizations; and, pretending that p is using such judgments to act in ways that would be appropriate if p . (Langland-Hassan 2012, p. 157)

So according to Langland-Hassan, in order to pretend that they are a dog, a child will begin by asking (internally) 'what would likely happen if I were a dog?'. They will then answer this question by calling upon generalisations related to dogs, such as that they bark and run around. Based on these generalisations, the child will form counterfactual beliefs such as 'if I were a dog then I would run around barking'. To count as pretending

that p , this child has to perform certain actions on the basis of these counterfactual beliefs.⁶³

He goes on to argue that forming these sorts of counterfactual beliefs won't involve entertaining the proposition that is supposed to be imagined by a pretender and so avoids the quarantining difficulties that we argued afflict simple belief accounts of pretence:

[p]retending (and imagining) that the green cup is empty does not here involve entertaining the proposition the green cup is empty. Nor does pretending (or imagining) that the other cup is full require entertaining the proposition the other cup is full. (Langland-Hassan 2012, p. 166)

This is because if, for example, someone comes to believe that 'If you had poured tea into both cups, they would both now be full', this won't involve entertaining a proposition such as 'both cups are full'⁶⁴, and so we can make sense of how the pretender keeps cognitive order since:

[a]t no time during the pretense does the child entertain a representation with a content that conflicts with—or “duplicates”—that of any of her beliefs. This means there is no proposition in need of quarantining. When the cups are initially “filled” during the pretense, the child does not need to infer (or believe) that the cups are full; rather, she needs to recognize that the experimenter is acting as if he is pouring tea, and to infer that if tea had been poured in the cups, they would now be full. (Langland-Hassan 2012, p. 166)

Empirical evidence seems to suggest that children are capable of counterfactual reasoning from a young age (Gopnik 2009, pp. 23-34) so there is no *prima facie* reason to rule this

⁶³ This somewhat oversimplifies what Langland-Hassan takes pretending to be. In a later paper (Langland-Hassan 2014) he offers a sophisticated attempt to define pretending, concluding 'To pretend that p is to act with the intention of making some w function, at that moment, in z -like ways, in the furtherance of a Pretense Episode.'

⁶⁴ One might object here that forming this sort of counterfactual still involves entertaining the proposition 'both cups are full' in some sense. We can mark a distinction here between a thin notion of entertaining, where believing that not p also involves entertaining p , and a thick notion of entertaining where p is represented *simpliciter*. Langland-Hassan's view is compatible with the former view, since entertaining p in this thin sense won't lead to us being motivated to behave in p like ways whenever we believe that not p .

proposal out. Indeed, as Langland-Hassan points out, Nichols and Stich's also argue that counterfactual beliefs are involved in pretence.⁶⁵ He thus argues that his theory makes do with only a subset of the mechanisms they use to explain pretence and that their imagination box is thus a *redundant posit* (Langland-Hassan, pp. 157-158).

One might wonder whether we should classify this theory as a distinct content view of what it is to propositionally imagine. On the one hand, this account relies on beliefs with conditional 'if ... then' contents, and there is a sense in which this is a sort of distinct content. However, some of our everyday beliefs will also have 'if ... then' counterfactual contents, since at least some of our everyday beliefs will be conditional. To justify classifying Langland-Hassan as offering a distinct content theory, one could argue he'd also need to introduce some sort of content that only relates to pretence. For our purposes, I think it is more helpful to classify Langland-Hassan as a distinct content view because these counterfactual contents, though not specific to pretence, are different from the simple sorts of beliefs we considered previously. Putting the view under this heading highlights that Langland-Hassan doesn't think that the beliefs involved in pretence have simple contents like 'I am a dog'.

Regardless of whether one agrees with me on this classificatory issue, his view remains opposed to the distinct attitude approach, and is explicitly offered as an alternative to Nichols and Stich's theory of imagination. My main worry with this approach is that it struggles to explain pretence recognition (how children recognise others are pretending) and pretence motivation (how children figure out what pretend actions to perform). To explain these two issues, I will argue in chapter 4 that we instead need to embrace an alternative distinct content view put forward by Leslie. That said, I will make use of some insights from Langland-Hassan's view to defend some of the worries critics of Leslie have raised about the possibility of believing at will.

⁶⁵ We will set out why in chapters 3. Roughly, because Nichols and Stich deny imaginings can directly motivate us, they think pretence also involves forming counterfactual beliefs like 'if I were a dog, then I would bark', which motivate pretend actions when combined with desires to act as-if this was true.

Before we move on to introduce Leslie's view of pretence, it is worth briefly noting that this counterfactual theory will struggle to work as a theory of how we engage with fiction. Although Langland-Hassan doesn't make the case for having offered an account of how we engage with fiction, he does suggest his view might be able to serve as a general account of what is involved in propositionally imagining something (Langland-Hassan 2016, footnote 9). Ferreira (2014, pp. 73-77) helps to bring out the difficulties of developing an account of how we engage with fiction that relies on counterfactuals by noting that the contents of fiction are often *rich* and can at least in theory be *sparse*. For an example of a rich fiction, consider this quote from *The Remains of The Day*:

Lord Darlington wasn't a bad man. He wasn't a bad man at all. And at least he had the privilege of being able to say at the end of his life that he made his own mistakes. His lordship was a courageous man. He chose a certain path in life, it proved to be a misguided one, but there, he chose it, he can say that at least. As for myself, I cannot even claim that. You see, I trusted. I trusted in his lordship's wisdom. All those years I served him, I trusted I was doing something worthwhile. I can't even say I made my own mistakes. Really – one has to ask oneself – what dignity is there in that?

It would be odd to suggest that we need to engage in some sort of counterfactual reasoning in order to engage with this section of text: all we need to do is read these English sentences and understand them. That being said, there are at least three different ways we can seek to apply Langland-Hassan's view to our engagement with fiction:

1. We form counterfactuals which take the actual sentences of a novel (or lines of dialogue in a movie etc.) as antecedents. (E.g. when we read this passage we begin by forming a belief like 'if Lord Darlington wasn't a bad man then q ').
2. We form counterfactuals about the fictional truths we are prescribed to imagine by fictions (e.g. when reading *Pride and Prejudice* we form beliefs like 'if there were a woman named Elizabeth Bennet then q ')
3. We form a single counterfactual that ranges over the entire reading of a text (e.g. when reading *The Remains of the Day* we form a belief like 'If the utterances that make up this story reflected a genuine historical account then q ')

(1) and (2) look like they are badly placed to explain how we engage with fictions. In this case of (1), there is no reason to engage in counterfactual reasoning to understand sentences in a work of fiction: we simply need to read and understand the sentences in question. (2) looks like something we *can* do when engaging with a work of fiction. For example, we might do this to try and figure out what will happen next in a story, or why a certain event happened. Once again, however, this doesn't seem to be a necessary element of engaging with fiction. (3) is somewhat more promising. The chief difficulty with this approach is that it would be hard to specify a single counterfactual that can explain how we engage with fictions. For example, the one I suggested in my phrasing of (3) presumes the fiction in question has a relatively truthful narrator. We would need further counterfactuals to account for unreliable or infelicitous narrators, such as Humbert Humbert in *Lolita*.

We might propose this means we need a disjunctive account of the relevant counterfactuals, such as we either believe 'if the utterances that make up this story reflected a genuine historical account then q ' or 'if I have reason to doubt that the utterances that make up this story reflect a genuine historical account then q '. This will still be counter-intuitive and we can bring out why by reflecting on what it is like to read *Lolita*. It gradually becomes clear as one reads *Lolita* that the narrator is unreliable: this is not explicitly announced at the start of the novel. This seems to imply that if this disjunctive counterfactual view is right, then the relevant counterfactual governing our engagement with a fiction will change as we come to realise that, for example Humbert is not to be trusted. This is a strange way to conceptualise what happens when one reads a novel with an unreliable narrator. It looks like we start to question the felicity of their utterances as we are reading, rather than making some sort of global change to a counterfactual we entertained to begin engaging with the work. Once again, it is more plausible to say we simply understand the sentences we are presented with, without them being subject to the scope of some sort of counterfactual.

As for sparse fictions, Ferreira rightly notes one can write a fiction which consists of a single sentence, such as 'There once was a man who went to the moon', and it doesn't look like engaging with a fiction made up of a single utterance like this one involves

reasoning counterfactually. One might ask oneself questions like ‘what would happen if there was a man who went to moon?’ but this isn’t necessary for engaging with this fiction. This would only be relevant if we were trying to reflect on possible entailments this utterance has, e.g. if we want to speculate how the man got to the moon, or why he went there: Is he an astronaut? Did he ride in a shuttle? Asking these sorts of questions and calling upon generalisations doesn't seem to be necessary for engaging with a sparse narrative fiction like this one, instead, we merely need to understand this single English sentence.

This also further demonstrates why it will be difficult to make a version of (3) work. When given a single proposition like this, what would the antecedent be for the counterfactual governing our engagement with the fiction? If it is merely something like ‘if this utterance is true then q ’, then this adds nothing to our understanding of the fiction. Why this is a strange proposal can be illustrated by comparing reading a fiction made up of a single sentence to reading the same sentence having found it written down somewhere (perhaps written on a bench). In the latter case, we will presumably just read the sentence and come to understand it. In the fiction case we might, perhaps, try to discern more meaning from the utterance, but this could also happen in an everyday context since we might wonder why someone has written down this esoteric remark. What is not obvious here is that there is any difference between how we read and understand a sentence like ‘There once was a man who went to the moon’ when it is presented as a fiction as compared to reading it in any other sort of context.

That being said, this does not rule out the possibility of accepting Langland-Hassan’s theory of pretence, since I’m willing to accept distinct explanations of how we engage in pretence and how we engage with fiction. The point here is only that his theory cannot serve as a general theory of every instance of what philosophers call propositional imaginings.

2.4.2 Leslie & Meta-representations

According to Leslie, the attitude that allows children to engage in pretence is belief with a proprietary structure. Firstly, these beliefs have contents which are what he calls

‘decoupled’: they are removed from ordinary belief processing.⁶⁶ Secondly, they involve the mental state concept PRETEND⁶⁷, a concept possessed by even very young children because it is supposed to be ‘innate’ in some sense (Leslie 1987, pp. 419-421). This is because Leslie views pretence as an early instance of theory of mind, which he also takes to be innate (Leslie 1987, pp. 421-423). As such, he is a theory-theorist about the nature of mindreading, and thinks we have an innate theory of mind module.

If Leslie is right, this means that the representations that allow us to engage in pretence have a complex structure of the form: Agent – Informational Relation – ‘*p*’. The informational relation involved is supposed to be captured by the concept PRETEND and the quotation marks around ‘*p*’ represent the fact this proposition is decoupled. So, for example, a child pretending to be a dog will have a belief that looks something like ‘I – PRETEND – “that I am a dog”’, where these quotation marks represent the idea this proposition lacks its ordinary causal consequences, thus making sense of the issue we noted in section 2.2 about how children are able to keep track of the distinction between pretence and reality.

However, we might think that we should interpret Leslie as offering a distinct content and distinct attitude view, according to which engaging in pretence involves a) pretending that *p* (*qua* imagining that *p*) and b) believing you are pretending that *p*. I will argue in chapter 4 that we do not have to adopt this sort of view to explain pretence by making use of some of the considerations Langland-Hassan uses to justify his counterfactual theory of imagination.

Leslie originally called these sorts of beliefs ‘meta-representations’, but now sometimes calls them ‘M-Representations’ to avoid the connotation that children have the

⁶⁶ Nichols and Stich (2000, p. 141) argue that this aspect of Leslie’s theory merely makes it a ‘notational variant’ of their own.

⁶⁷ Capitals are the standard way of expressing that children possess this concept, and I will respect this way of signifying this. This is a technical notion and is not supposed to perfectly capture our everyday notion of pretending.

concept of a representation (Leslie and Thaiss 1992).⁶⁸ Weisberg (2015, p. 5) helpfully explains that Leslie takes engaging in and recognising pretence to involve meta-representation because ‘it involves representing someone’s representation of a state of affairs’. I will reflect on some worries about the sophistication this attributes to young children in chapter 4.

The key explanatory virtue of Leslie’s view is that it helps to explain how children are able to recognise pretence behaviours performed by others around the same time they begin to engage in such behaviours themselves, which is something that both the distinct attitude view and Langland-Hassan’s counterfactual distinct content view have difficulty explaining, an issue which will be the focus of chapter 4. Both these views are ‘behavioural’ views of pretence and this means they argue that recognition of pretence in others will involve recognising others are engaging in certain sorts of behaviour. Leslie argues this is insufficient for explaining how children are able to recognise pretence. I will argue this also shows belief-like imaginings are insufficient for explaining how children are able to engage in pretence behaviour, and that once we suitably clarify Leslie’s view, they are also rendered unnecessary.

2.4.3 Beliefs about Fiction

The final distinct content view we will discuss argues that the attitude we take towards works of fiction is belief subject to a *fictional operator*, which we can also call (somewhat ambiguously) *beliefs about fiction* (where this is elliptical for something like beliefs about what is fictional, or beliefs about what is true in the fiction, and so on). On this approach, in some cases propositionally imagining p amounts to believing that p is fictional, pretend, make-believe or imaginary. To use Walton’s notation, what philosophers call propositional imaginings are sometimes mental states that take on something like the form ‘I believe that $*p*$ ’ rather than ‘I believe that p ’. This view receives some support from

⁶⁸ This misunderstanding motivated some of the initial critics of his theory, such as Perner (1991) who emphasised that children lack a concept of REPRESENTATION and so cannot be said to form meta-representations.

philosophers of language, such as Kripke (2011, 2013) and Thomasson (1999), who argue that much of our talk about fiction is subject to this sort of implicit sentential operator. To express that someone has this sort of belief, I will sometimes adopt the form used by Tullmann (2016) and Tullmann and Buckwalter (forthcoming) and maintain the relevant beliefs here are of the form ‘I believe p [in the fiction]’, since this form helps to bring out that this operator may not always be salient in our conscious experience.

Matravers tentatively points out that defenders of the distinct attitude approach often don’t give us enough information as to why we should reject this sort of content-based account. In relation to the attitude we take towards events that occur in the film version of *Sense and Sensibility*, he notes:

“[t]he claim is that the audience imagine (or “make believe”) that Marianne is heartbroken over Willoughby. However, with respect to this argument, there is no need to propose such a mental state: all that is needed is that the audience believe it is true in *Sense and Sensibility* that Marianne is heartbroken over Willoughby. (Matravers 2010, p. 191)

Neill also speaks positively of this view in relation to *Emma*:

There is certainly nothing "putative," "insincere," "hypothetical" or "provisional" about my belief that Emma Woodhouse was handsome, clever and rich. I do, actually, believe that (it is fictionally the case that) Emma had all of these attributes. There is nothing fictional about beliefs of this sort; it is their *content* that concerns the fictional. Beliefs about what is fictionally the case, that is, are just that: beliefs. (emphasis mine) (Neill 1993, p. 3)

This sort of change of content view is sometimes criticised for failing to explain why we can become immersed in works of fiction, and in light of this why we exhibit affective responses to fictions. I will respond to this worry at the conclusion of chapter 5.

Before moving on to consider the merits of these sorts of distinct content accounts of pretence and fiction as compared to distinct attitude approaches, it will be helpful to consider whether we have some *prima facie* reasons for resisting a distinct content view of pretence and fiction that can be found by reflecting on how we can defend the introduction of a notion of belief-like imaginings associated with supposing. For the sake of argument, I will assume this same counterpart can play a role in all three activities, and

I will set aside my worries about uniformity and whether fiction involves an involuntary counterpart. We thus need to consider whether there are *introspective*, *functional* and *neurological* arguments for distinguishing belief-like imaginings and belief, and if these arguments do enough to associate a belief-like attitude with either fiction or pretence. These arguments I think can do enough to introduce some sort of counterpart to belief, but they do not give us reason to associate it with fiction or pretence. As such, we will need to consider more specifically what a theory of these activities has to account for to see whether we need to introduce belief-like imaginings.

2.5 Distinct Attitudes or Distinct Contents?

From the perspective of introspection, it can be pointed out that we are usually able to tell the difference between things we believe and things we might take ourselves to imagine propositionally in the sense that I am associating with a notion like supposing or entertaining.⁶⁹ When I decide to take a break from writing and imagine *that I am on a beach* I can recognise that I am only imagining this. However, philosophers rarely place much emphasis on this sort of argument when it comes to justifying the introduction of belief-like imaginings into an account of fiction and pretence. It is true that in both activities we recognise that we don't straightforwardly believe things. Having watched *The Sopranos*, I don't believe that I could actually run into Tony's crew if I were to visit New Jersey. However, this doesn't do enough to show that I have belief-like imaginings about Tony and his gang as opposed to beliefs about the fiction. Introspection tells me that I don't straightforwardly believe Tony and his crew exist, but not that I have belief-like imaginings. A similar point is also relevant to pretence. If I pretend that *I am a dog*,

⁶⁹ Possible exceptions here include delusions and self-deception since it is debatable whether the states involved in these cases are beliefs or imaginings. Currie and Ravenscroft (2002, Ch. 8) argues delusion involves mistaking an imagining for a belief. Egan (2008) suggests delusions involve somewhat belief-like and somewhat imagination-like states called 'bi-imaginings'.

introspection tells me that I do not straightforwardly believe that I am a dog; it does not tell me that I am imagining this proposition in a belief-like manner.

When we made use of this sort of self-recognition argument to defend the idea that the formation of mental imagery is a counterpart to seeing, it was noted that mental images have a somewhat different phenomenology to actual seeing. There is ordinarily an introspective difference between forming an image of an apple and actually seeing one. With belief and imagination, the case is less clear since it is not self-evident what the phenomenology of belief is, nor what the phenomenology of propositional imagining is supposed to be. For example, what is the phenomenology of believing that *water is H₂O*? It's not obvious this belief is associated with any distinct phenomenology, nor is it obvious it has a different phenomenology as compared to imagining that *water is H₂O*.⁷⁰

Bearing in mind these sorts of limits on introspection, philosophers tend to note *functional* similarities and differences between belief and belief-like imaginings in order to further justify their introduction and to associate them with fiction and pretence. A first question to ask is what is (or are) the functional role (or roles) of belief.⁷¹ Schwitzgebel (2016) helpfully summarises some common proposals for the functional roles of belief:

1. Reflection on propositions (e.g., *q* and if *q* then *p*) from which *p* straightforwardly follows, if one believes those propositions, typically causes the belief that *p*.
2. Directing perceptual attention to the perceptible properties of things, events, or states of affairs, in conditions favorable to accurate perception, typically causes the belief that those things, events, or states of affairs have those properties (e.g.,

⁷⁰ Van Leeuwen (2013, p. 792-793) notes this sort of worry and points out that in general phenomenological reflection is not a good way of distinguishing belief and what he calls 'attitude imagination'.

⁷¹ There are other approaches we could take here. We could, for example, say that believing involves having certain *dispositions* or to *represent* things in a certain way. Since the philosophers who introduce belief-like imaginings into accounts of fiction and pretence often defend functional views of belief, we will not consider alternative accounts of belief.

visually attending to a red shirt in good viewing conditions will typically cause the belief that the shirt is red).

3. Believing that performing action A would lead to event or state of affairs E, conjoined with a desire for E and no overriding contrary desire, will typically cause an intention to do A.
4. Believing that p , in conditions favoring sincere expression of that belief, will typically lead to an assertion of p .

It's worth briefly reflecting on each of these proposed functional roles to see if they give us reason to associate a notion of belief-like imaginings with fiction and pretence. The first shows a functional similarity between belief and belief-like imaginings. If I imagine that I am on a beach in Orlando, and I believe that Orlando is in Florida, then I'll imagine that I am in Florida (or at least, I will be disposed to do this upon reflection). Whatever attitude we take towards fiction and pretence also behaves in this way. If I take Sherlock Holmes to be a man *in the fiction*, and believe all men are mortal, I will take him to be mortal *in the fiction* (Currie and Ravenscroft 2002, p. 14).

The second proposed functional role reveals what looks like a difference between the two attitudes. Belief-like imaginings are not sensitive to perception in the same way beliefs are. If I decide to imagine that *I am on a beach*, this may well be on a mere whim, rather than on any sort of perceptual observation of the external world. There is also no obvious thing I could observe to make me cease to imagine this. However, this distinction becomes less apparent when we think about the sort of attitude involved in pretence and fiction. In the game of bears described by Walton that we discussed in section 1.4, children do indeed come to think *that there are three bears nearby* on the basis of what they perceive: the location of stumps. This is also true when thinking about fiction. What we take to be true of fictions depends on what we see on the page or see on screen. The reason I think that it is fictionally true that Harry Potter is a wizard is not because of a mere whim but because of the sentences I read in the Harry Potter novels.

(4) is also somewhat ambiguous when thinking about pretence and fiction. The claim here is along the lines of you will sometimes assert p if you believe p . We saw in section 1.4 that people do assert things like 'there are pies in front of me' when playing games

of pretence or ‘Harry Potter is a wizard’ when engaging with a fiction. Again, it’s not obvious this gives us a reason to introduce belief-like imaginings into our account of fiction and pretence.

So far, all these mooted functional differences and similarities suggest that there may well be some sort of counterpart to belief, perhaps as I argued earlier one associated with notions like *supposing* or *entertaining*, but they do not show that this attitude must be associated with fiction and pretence. Beliefs with distinct contents can equally explain these functional differences. In the case of (1), my inferences when engaging with fiction might be made up of a collection of ordinary beliefs (such as that men are mortal) and beliefs about the fiction (such as that Holmes is a man). In the case of (2), we can argue that sometimes what we perceive should lead to us forming beliefs about what is true in a fiction or in an episode of pretence. Finally, (3) can be made sense of by suggesting that when people talk about fictions and pretence they are asserting what they believe to be fictional or pretend.

(3), the point about belief being connected to motivation, tends to have the most weight placed on it in contemporary discussions of belief and imagination when it comes to justifying why a counterpart to belief must be involved in our engagement with fiction and pretence. Velleman sums up a version of this approach to understanding belief as:

[a]ll that’s necessary for an attitude to qualify as a belief is that it disposes the subject to behave in certain ways that would promote the satisfaction of his desires if its content were true. (Velleman 2000, p. 255)

This reflects a difference between belief-like imaginings and belief. Consider, for example, my imagining *that I am on a beach in Orlando*. On its own, merely imagining this proposition won’t motivate me to do anything. It may well do if it stimulates further beliefs and desires – I might realise I desire to be on a beach in Orlando and thus come

to believe I should book a flight, for example – but this imagining cannot motivate me *qua* imagining.⁷²

Unlike the first few functional roles we've discussed, this also illustrates a difference in the case of fiction. Sitting in a theatre watching *Romeo and Juliet*, you realise Juliet will shortly cast a dagger through her heart, and yet you do nothing to try to save her. You merely sit impassively in your chair, perhaps shedding tears at her plight. It could be argued that if the distinct content approach were right, we would have beliefs about Juliet in this sort of scenario and we ought to be motivated to perform actions on the basis of these beliefs. Since we are not motivated to try and interfere with fictions, beliefs, even those with distinct contents, cannot be involved in our engagement with fiction.⁷³

However, it's apparent from how Schwitzgebel phrases (3) that the relation between belief and motivation is complex. In order for a belief to motivate us, we must also believe a given action we can perform will help to bring about a state of affairs that we desire to bring about. The importance of this condition can be seen by supposing that you believe that a monster is attacking [in the fiction]. In this sort of case, you would also believe you were merely engaging with a work of fiction. As a result of this further belief, you will lack what we can call 'conditional' beliefs that would be necessary to motivate actions (Matravers 1991, pp. 34-35). In the Juliet case, we'd have to believe there were some actions we could perform to interact with the fiction and save her from her fate. Walton (1990, pp. 192-195) brings out that usually this sort of direct participation with fictions is ruled out. If we run on stage and attempt to snatch the knife out of Juliet's hand, this will not constitute saving Juliet *in the fiction* as opposed to rather annoyingly interrupting the performance by taking the knife out the actress's hands.

⁷² Everett (2013, p. 11) also highlights this disconnection in relation to imagining that a book you need is in the library. Merely imagining this will not make you go to try and check it out.

⁷³ The case for pretence will be more complex, since children do indeed perform pretend actions. We will see in chapter 3 that Nichols and Stich explain this by suggesting beliefs and desires are what actually generate pretence behaviour when combined with belief-like imaginings.

Interactive fictions like videogames offer an interesting contrast to illustrate this point. Take the recent role-playing game *Persona IV*. Like many videogames, the player is able to direct their on-screen avatar in various ways. For example, you can decide which non-player characters to spend your free time with, and eventually, which one to form a romantic relationship with.⁷⁴ Now, when playing this sort of game, you are to some extent able to change the course of the fiction. This means that you possess the relevant sorts of conditional beliefs about being able to affect events and happenings in the fictional world. If you want your avatar to date a certain character you can make this happen *in the fiction*.

If this account of why we don't interfere with ordinary non-interactive fictions is right, we can resist the introduction of belief-like imaginings into an account of how we engage with works of fiction. Our lack of motivation stems from a recognition of the limits of how we can interact with fictions, rather than the distinct functional role of some sort of belief-like attitude. In turn, this helps to explain why we can be motivated to perform pretend behaviours when engaging in pretence: this is also a case where we recognise that we are able to interact with a fictional world. As such, issues related to motivation do not force us to link belief-like imaginings and fiction.⁷⁵

⁷⁴ Alternatively, one can choose to not form a relationship with any character, or to form one with multiple characters. This latter course of action comes to haunt you on (in-game) Valentine's Day when your multiple lovers turn up at school with gifts for you, each delivering a line of dialogue about how you've broken their heart. These sorts of examples raise an interesting question about whether interactive fictions allow for some affective responses we cannot exhibit towards non-interactive fictions, such as guilt.

⁷⁵ One could also question whether we really do lack motivation when engaging with non-interactive fictions. Buckwalter and Tullmann (forthcoming, p. 12) point out that one might strive to only buy certain kinds of diamonds after witnessing the cruelty of the trade in *Blood Diamond*. Fiction is of course powerful in these sorts of ways, hence the controversy over certain political and social works. One might object that this is not sufficiently being motivated by the fiction *qua* fiction. To this Buckwalter and Tullmann (forthcoming, p. 13-15) speculatively point out that we do still exhibit affective responses, which is some sort of motivational response, albeit one that is somewhat attenuated.

In light of all this, solely focusing on the functional differences between belief and belief-like imaginings might help to further justify the introduction of some sort of notion of belief-like imaginings associated with supposing and entertaining, but does not give us reason to put this attitude to work in an account of fiction or pretence. This leaves us with neurological similarities. Schroeder and Matheson (2006) purport to offer some neurological evidence for both the existence of this distinct cognitive attitude and for associating it with how we engage with fictions. They do this by tracing causal pathways from the propositional imagination to our affective systems:

[t]he view has emerged that these acts [imaginative acts, such as imagining the contents of a fiction] have their power to move us through their activation of special cognitive attitudes, akin to beliefs in structure and in some of their effects, but distinguished from beliefs in others. This view, positing what we will call a ‘distinct cognitive attitude’... is ultimately an empirical thesis in certain important respects, ... We aim to add the blessings of neuroscience to the view, and thereby to put a final seal of approval on it. (Schroeder and Matheson 2006, p. 19)

They begin by discussing how sensory impressions and quasi sensory impressions are generated by our sense organs. The contemporary neuroscientific view of this issue is that our sense organs produce neural signals that generate patterns in the brain called *unimodal* sensory representations. The ‘uni’ here captures the fact that these are representations from a single sense modality. Unimodal representations from various modalities combine to form *multimodal* sensory representations, which are representations of things in our environment we can experience via multiple senses. This applies to everything from the iMac I’m typing on, to the trains opposite my flat, to any other ordinary object in my environment.

Schroeder and Matheson go on to point out that multimodal representations also play a role in generating affective responses. Studies suggest that they are able to interact with various areas of the brain associated with affective responses, such as the orbito-frontal cortex, the affective division of the striatum and the amygdala (LeDoux 1996).

Their next move is to point out that this same pattern of interaction occurs regardless of whether a multimodal representation is of something real, or something imaginary/fic-

tional. In particular, they point out that the same interactions between multimodal representations and areas associated with affective responses occur when engaging with works of fiction and when we engage with actual objects in our environment. This also holds true for imaginary representations that we generate at will.

They argue that this sort of data demonstrates that belief-like imaginings and beliefs interact with our affective system in much the same way (Schroeder and Matheson 2006, p. 29-30). However, it is equally possible that what these studies show is that our affective systems respond to beliefs with fictional or imaginary contents in much the same way as beliefs about the real world.

I presume they take this interpretation to be ruled out because they accept the arguments for belief and belief-like imaginings having distinct functional roles. For example, they note (Schroeder and Matheson 2006, p. 29) that imaginings don't appear to motivate us. As such, they may well take this as sufficient for justifying that there is a counterpart to belief involved in our engagement with fiction. However, as we noted above, reflecting on interactive fictions gives us reason to question this strict demarcation between beliefs and the cognitive attitude involved in our engagement with fiction. As such, since we have denied that functional arguments can demonstrate that a distinct attitude is associated with fiction, this neuroscientific data doesn't have to compel us into accepting a distinct attitude view of our engagement with fiction.⁷⁶

Bearing all this in mind, purely considering introspective, functional and neurological issues does not do enough to show we should associate belief-like imaginings with either fiction or pretence. As such, we will need to look at these activities in more detail to see whether the best explanation of them is one that involves belief-like imaginings.

⁷⁶ Buckwalter and Tullmann (forthcoming, pp. 16-20) independently offer a similar response to this neurological argument.

Conclusion

In this chapter, we saw in section 2.1 that there are good reasons to accept that there is some sort of voluntary counterpart to belief, which we can associate with notions like ‘supposing’. However, in sections 2.3 and 2.5, I argued that there is no *prima facie* reason to introduce this counterpart into an account of how we engage in pretence and with fiction, or to reject a distinct content view of how we engage in these activities. We noted that this is particularly true in the case of engaging with fiction, where it looks like the relevant sort of counterpart to belief philosophers have in mind may well be an involuntary one.

As such, the most productive approach to seeing if we should associate a belief-like counterpart with pretence and fiction will be to investigate what an account of how we engage in these activities needs to explain, and whether these desiderata can be met by beliefs with distinct contents, or if we need to allow a role for belief-like imaginings. We will begin by considering pretence. I will argue that ultimately pretence is better explained by a distinct content approach and that the distinct attitude view cannot explain all the issues related to how we engage in pretence. We will thus see that a distinct attitude is both *insufficient* and also *unnecessary* for explaining how we engage in pretence.

Chapter 3: The Imagination Box & Pretence

Introduction

We saw in the previous chapter that we cannot rely on generic arguments for taking belief to have a counterpart in order to justify associating belief-like imaginings with pretence. In this chapter, we will do two things. Firstly, we will introduce the explanatory demands that any satisfactory theory of pretence needs to account for. Secondly, we will criticise Nichols and Stich's arguments for the claim that introducing a counterpart to belief is the best way of meeting these explanatory demands. Our focus will be on Nichols and Stich because they offer the best-developed account of why we should associate a belief-like attitude with pretence.

I will focus on some difficulties that arise from their claim that belief-like imaginings and beliefs share a *single code*, a concern also discussed by Stock (2011a). My argument will differ from hers by emphasising the importance of *clustering* when it comes to explaining how we engage in pretence and with fiction. I will argue that clustering might be better explained by introducing a distinct content view, which will reveal a way in which we can challenge the non-doxastic assumption in the context of pretence by questioning the sufficiency of introducing belief-like imaginings.

The single code hypothesis tells us that beliefs and belief-like imaginings have the same logical and representational structure, and are therefore processed by our cognitive mechanisms in a similar way.⁷⁷ If this is right, this means that a belief like *the cat is on the mat* has the same logical and representational structure as the belief-like imagining that *the cat is on the mat*. This doesn't necessarily entail that these representations are

⁷⁷ An interesting question here is whether the notion of code is meant to have some sort of neurological implications. Nichols and Stich note that we don't know exactly what the code of belief is, or whether there is only one code underpinning belief or multiple codes, but it looks like we have to think there is some sort of neurological link here, otherwise it is hard to get a grip on the notion of a 'code'.

linguistic or quasi linguistic in form; the suggestion is merely that some sort of internal code determines how these representations are processed by our cognitive systems (Nichols and Stich 2003, p. 15).⁷⁸

If this proposal is right, functional similarities between belief and belief-like imaginings are explained by maintaining that code determines processing. This raises a question about the role of their ‘boxes’ when it comes to cognitive processing. Does the fact that a given representation is found in our belief box, rather than our imagination box, have implications for how it is processed? Or, alternatively, do only differences in code affect processing? It might be thought obvious that the contents of different boxes are processed in different ways, at least on some occasions. This, after all, is part of what it means to say that two states have different functional roles. However, whether Nichols and Stich can say this about belief-like imaginings and beliefs is unclear. They appear to suggest that the reason that there are similarities between belief and imagination is because they are in the same code. It is harder to see what explains the differences between belief-like imaginings and beliefs if we accept their theory. I will consider the difficulties their view faces when it comes to accounting for affective responses and clustering.

We could reject this specific commitment to a single code and endorse something like a simulationist view of belief-like imaginings. However, the appeal of the single code based approach is that it is supposed to give a principled way of explaining various features of pretence. If we jettison this hypothesis, we lose some of the reasons for preferring a distinct attitude theory of pretence to a distinct content theory of pretence.

To develop this argument, in section 3.1 we will begin by setting out why childhood pretence has been regarded as puzzling by philosophers and psychologists. I will also introduce some specific examples of pretence that are offered by Nichols and Stich, since these will be helpful for framing what a theory of pretence needs to explain.

⁷⁸ As such, this also shouldn’t be read as necessarily introducing a Fodorian language of thought. For example, it might be the relevant code is ‘map-like’ (Camp 2007).

In section 3.2, I will offer my own interpretation of what a theory of pretence needs to explain, which will somewhat simplify Nichols and Stich's account by placing the relevant features under three headings: *set-up*, *elaboration* and *output*.

In section 3.3, I will explain why Nichols and Stich think that these three features are best explained by introducing a distinct belief-like attitude. I will also note that their explanations of the motivation of appropriate pretence behaviours and the recognition of pretence in others are somewhat unconvincing, and we will return to consider these issues in more detail in chapter 4.

In section 3.4, I will question how we are able to explain the differences between Nichols and Stich's distinct belief-like attitude and our genuine beliefs, bearing in mind that they think they both share a 'single code'. I will highlight two particular issues. Firstly, we will discuss the fact that our affective responses to what we appear to imagine in a belief-like way often vary from our ordinary responses. Secondly, we will discuss clustering, the way that our representations related to fiction and pretence form clusters and thus are not viewed as being fictionally true *simpliciter*.

In section 3.5, I will consider how Nichols and Stich might make sense of these two asymmetries by considering the functional role response and the desire response. We will see that neither of these responses is entirely convincing. I will then suggest in section 3.6 that the best way to make sense of affective asymmetries is by allowing for clustering to play a role in shaping our responses to fiction and pretence. However, I will also argue that it is hard to accommodate clustering if we accept the single code hypothesis. This will reveal an initial way in which a distinct belief-like attitude might be insufficient for explaining our engagement with fiction and in pretence and so will give us an initial reason to challenge the non-doxastic assumption.

Finally, in section 3.7 I will reflect on whether we can defend Nichols and Stich's theory against my worries by considering whether I have misunderstood the aims of this sort of 'boxological' theory. I will argue that these difficulties with explaining differences between beliefs and belief-like imaginings are indeed enough to make us reject Nichols and Stich's idea of a 'single code'. However, I will then suggest that a distinct attitude view which dispenses with this commitment will be less explanatorily satisfying. This will

set the stage for considering a distinct content approach to making sense of our engagement in pretence in chapter 4.

3.1 Childhood Pretence

Trying to explain how young children are able to engage in pretence offers a fascinating challenge for philosophers and psychologists. Pretence behaviour is so ubiquitous in children that it is easy to overlook how remarkable it is. From around the age of 15 months (Bosco *et al.* 2006), children are able to engage in behaviour where they pretend the world is different to how it really is, even though they still lack a sophisticated understanding of the world around them. My girlfriend's young niece developed a habit of climbing on top of various objects and making a 'woof' noise just before turning a year and a half old, whilst still only being capable of saying a few basic words.

Since the late 1980s, there has been a rapid advance of empirical and philosophical work that seeks to explain why children are able to engage in pretence from such a young age. To list a few examples, Baron-Cohen (1985) has done work on the relation between childhood pretence and autism; Leslie's (1987, 1994) has investigated the link between pretence and theory of mind; Currie (1995) and Gordon and Baker (1994) have attempted to link simulation theory and pretence; and Lillard (1993) and Harris and Kavanaugh (1993) have investigated children's understanding of the nature of pretending. Much of the psychological work on pretence is summarised by Harris (2000), and more recently in review articles by Weisberg (2015) and Gendler and Liao (2011), with the latter bringing together both recent philosophical and recent psychological work on pretence in order to suggest further avenues for investigation.

Before we begin our discussion of what mental attitudes might help to explain how children are able to engage in pretence, it will be helpful to introduce some examples of pretence. Nichols and Stich helpfully set out three specific examples of pretence from the psychological literature and also describe a series of their own experiments in which they requested adults engage in pretence. From these cases, they then bring out what any theory of pretence must explain.

I will set out all of their examples of childhood pretence, though for the sake of brevity I will only describe two of their experiments involving adult pretenders, since these two cover what Nichols and Stich aim to bring out about the nature of pretence from their adult examples.

1. The Banana Phone

In this first example, introduced by Leslie (1987, p. 416), a child engages in a game where they pick up a banana and act as if it were a phone, e.g. by holding it up to their ear and speaking into it. Nichols and Stich note two ways a child might engage in this sort of pretend scenario. Firstly, a child might spontaneously pick up a banana and start talking into it, perhaps saying ‘Hello grandma! Are you coming over later?’. This is an example of what we can call *solitary pretence*. Secondly, the child’s mother might pick up a banana and say ‘Hello? Yes, he’s right here, just a minute!’ and then pass the banana to her child saying, ‘It’s for you!’. Frequently, children will go along with this sort of pretence by ‘taking the banana and saying ‘Hello’ into it’ (Nichols and Stich 2003, p. 19). This is what we can call *group pretence*.

Setting up these two versions of the banana phone example helps show that we need to do more than just explain how children initiate games of pretence. We must also make sense of how they *recognise* when games of pretence have been initiated by others.

2. The Tea Party

The tea party example is perhaps the most famous example of pretence in the literature on pretence, and was first introduced by Leslie (1994, p. 222). In Leslie’s original experiment, children were asked to pretend that they were at a tea party and to ‘fill up’ two cups in front of them with ‘tea’. The experimenter then upturned one of the ‘filled’ cups. The children taking part in the experiment were then asked which of the two cups was ‘empty’ and which was ‘full’. Leslie found that two-year-olds were reliably able to pick out the previously upturned cup as the ‘empty’ one, even though in reality both cups were empty throughout the tea party.

An account of pretence needs to say something about both how children keep track of what's true in their pretence and about how these sorts of inferences are made.

3. Monsters and Parents

The third example Nichols and Stich introduce comes from CHILDES, a database created to study children's language use. In the example they select from this database (Nichols and Stich 2003, pp. 20-21), two children are pretending a toy car is a bus, then one child says, 'pretend there's a monster coming ok?' and the other replies, 'No let's don't pretend that... Cause it's too scary that's why.' They then proceed to play another game, where they pretend to be mother and father, and 'cook' a pair of shoes for food.

This seemingly simple example has numerous implications for a theory of pretence. Firstly, it shows that children can refuse to engage in pretence and that seemingly ordinary emotions (like fear) can play a role in this refusal. Secondly, children can spontaneously make use of props that resemble what they are pretending them to be (e.g. the toy car standing in for a bus) but also ones that don't (e.g. the shoes as food). Finally, much like the banana case, this example shows that deciding what to pretend is not always a solitary activity. In this case, the two children decide amongst themselves what they should pretend to be the case.

4. Dead Cat

Nichols and Stich's final example of pretence in children comes from Gould (1972, p. 212). In this example, a child on a climbing frame says 'I'm a pussycat. Meow. Meow.' then climbs down from the frame, lies flat on the ground, and says 'I'm dead. I'm a dead pussycat... I got shot.'

Once again, this example raises several interesting questions. Firstly, we must ask why the child made a 'meowing' noise whilst pretending to be a cat. How did they know this was an appropriate action to perform? Secondly, we can ask why the child continued talking if he was supposed to be pretending to be a dead cat. Finally, the child suggests he is dead because he 'got shot'. This wasn't entailed by the pretence beforehand. Why did the child say this is how his cat-self died?

5. Adult Pretence

Nichols and Stich (2003, pp. 20-24) give several examples of pretend play in adults based on their own experiments. In their experiments, they gave university students a series of premises telling them what to pretend. Some of these premises related to solitary scenarios (e.g. pretend this banana is a phone) and other premises related to group scenarios (e.g. pretend you are at a fancy restaurant and that one of you is the server and one is the diner). They draw special attention to two episodes of pretence that occurred during their experiments, both of which occurred during the fancy restaurant group scenario.

In the first pretence episode, the person pretending to be a diner refuses the server's offer of a wine list, and tells them that they're allergic to pepper and thus can't eat something with pepper in the sauce. They order a house salad instead. In the second pretence episode, the waiter performs several unusual actions, such as pretending to grind peppercorns with the heel of his foot (since he has no grinder available) and bringing out a sword for the diner to eat his food with. Perhaps most unusually of all, the server is inattentive with this pretend sword, and at one point accidentally cuts off the head of the diner's imaginary companion.

3.2 Features of Pretence

Drawing on these examples, Nichols and Stich suggest that there are five things an explanation of pretence must make sense of. I'll summarise them under three headings: *set-up*, *elaboration* and *output*.

1. Set Up

In all these examples, the pretence episode began with what Nichols and Stich (2003, p. 24) call 'an initial premiss or set of premisses'. In the tea party example, the initial premise is that the child and the experimenter are going to have a tea party; in the fancy restaurant case, the premise is that one of the participants is a diner at a fancy restaurant and one is the server at a fancy restaurant; in the banana phone example, the premise is something like 'this [banana] is a telephone'.

These examples also bring out the fact that these premises (or sets of premises) can be generated in two ways. Firstly, the pretender can spontaneously produce the premise (as in the banana phone example). Secondly, the pretender can go along with someone else's premise (as in the fancy restaurant and tea party examples). A theory of pretence needs to explain both how the initial premise of a game of pretence can be spontaneously generated, and how a pretender is able to figure out what the premise is when they engage in games started by others. Under the 'set-up' heading we thus have two things that need to be explained by any theory of pretence: *generation* and *recognition*.⁷⁹

2. Elaboration

In all the examples we discussed, the pretence episode seemed to proceed in-line with elaborations made by the pretender. These elaborations were both *inferential* and *non-inferential*. In relation to inferential elaborations, Nichols and Stich note that:

[f]rom the initial premiss along with her [the pretender's] own current perceptions, her background knowledge, her memory of what has already happened in the episode ... the pretender is able to draw inferences about what is going on in the pretence. (Nichols and Stich 2003 p.25).

The tea party scenario is the paradigm example of this in the pretence literature. In order to correctly select the 'empty' cup, the child must infer that the upturned cup is empty *in the pretence*, even though both cups in the experiment will be literally empty throughout, since no actual tea has been poured. These sorts of inferential elaborations were also

⁷⁹ One could also argue that we should explain refusal here, as occurred in the CHILDES case where the child refused to pretend because 'it was too scary'. My reason for not including this as a separate category is because contra Nichols and Stich, I don't think we need to say much about this feature of pretence. This strikes me as similar to asking why someone turned off a horror movie: it could be they had an unpleasant startled sensation; they might have found the plot boring; they might have found the representations of women in the film problematic; they might have just been hungry or have needed to go to the bathroom. I am sceptical of there being a simple factor we can point to that will neatly explain refusal beyond noting that at least some of our ordinary attitudes are involved in pretence, and that these will sometimes make us reject a pretence premise.

made in the fancy restaurant case, since some of the participants calculated how much ‘change’ they were owed based on simple mathematical reasoning. Variations of the tea party experiment have revealed various examples of children drawing inferences related to their pretend games. Harris (2000 pp. 17-19) recounts an experiment where a ‘full’ cup of tea was emptied over a teddy bear’s head and children were asked what would happen to the bear. Children phrased their responses in various ways, but agreed the bear would be covered in ‘tea’. (e.g. some said the bear would be ‘teay’)

These examples also bring out that many of the elaborations that occur during episodes of pretence are non-inferential. In a banana phone scenario, the child might ‘talk’ to a specific individual on the ‘phone’, such a grandma, but it isn’t logically entailed by the premise that *the banana is a telephone* which specific individual they are calling. In the example from CHILDES, when the children decided to put shoes in the oven as a prop for food, this also wasn’t entailed by the pretence premises. The two examples of adult pretence also involved non-inferential elaboration, since many of the actions that occurred during these episodes (e.g. rejecting the wine list, requesting a pepper-free dish, chopping a guest’s head off) were not entailed by the initial premises.

As Nichols and Stich (2003, p. 25) note, some of these non-inferential elaborations reflect real world preferences of pretenders. For example, one participant said she didn’t like to ask for the wine list in real life as she doesn’t like being asked for identification to prove her age. However, this participant also said she was allergic to pepper in the pretence, but she later clarified that she wasn’t allergic to pepper in real life: she merely pretended that this was true.⁸⁰

A successful theory of pretence must explain how these sorts of non-inferential elaborations are generated along with inferential elaborations.

⁸⁰ The server who managed to chop the head off one of the guests thankfully admitted he would not be so careless with a sword in real-life!

3. Output

One of the more obvious features of pretence brought out by these examples is that pretenders do in fact perform various pretend actions during their episodes of pretence. A child who pretends that a banana is a telephone will actually pick up the banana and talk into it. Nichols and Stich (2003, p. 26) report that one server in the fancy restaurant case pretended to scribble down the order of their diner, while another actually wrote it down on a piece of paper.

Children also seem to respond *emotionally* to their games of pretend, such as the boy in the monster case who refused to pretend a monster was approaching because it was ‘too scary’. Harris (2000, p. 58) discusses an example from Taylor (1998, p. 212) of a girl who started crying, and when asked why explained that her (pretend) horse wasn’t able to come outside with her. She was apparently unable (or perhaps unwilling) to rectify the situation by choosing to pretend the horse was present.

These sorts of phenomena leave us with several questions. In relation to action, we have to ask what motivates pretenders to actually perform pretend actions. What combination of mental states motivates pretence behaviour? We also have to ask how a pretender knows what the appropriate pretend action to perform is: how does a pretender know that in order to pretend to be a cat they should ‘meow’ rather than ‘bark? Finally, since some of the behaviours described only dimly resembled what the pretender was pretending to be the case, we need to ask what counts as an appropriate pretend action in the first place. We saw that declaring ‘I’m dead’ appears to be an appropriate way of pretending to be a dead cat, but a dead cat would be silent (and in any case, even an alive cat wouldn’t speak in English or any other language).

In the case of affective responses, we need an account of how things we merely pretend to be the case can generate emotional responses, and an account of why these responses appear to roughly track the responses we would exhibit to similar real-life scenarios.

Although children respond emotionally to some of the things they pretend to be the case, in other respects the effects of pretence seem to have no effect on the subsequent mental states of the pretender and are instead *quarantined* from their ordinary attitudes

towards the world. Most obviously, Nichols and Stich (2003, p. 27) note that pretenders do not believe the pretended events really happened when they conclude their game of pretence. When a child pretends a banana is a telephone and uses it to pretend to speak to grandma, she doesn't believe she actually spoke to grandma at any point before, after, or during the pretence.⁸¹ Any theory of pretence thus needs to explain *motivation, affective responses* and *quarantining*.

Having noted these features of pretence, Nichols and Stich move quickly to establish that any satisfactory explanation of these features will involve the introduction of a distinct belief-like attitude. The first step in their argument is to establish that a distinct cognitive attitude is required to explain pretence, before moving on to argue why this attitude must be a counterpart that shares a single code with belief. I will now reconstruct their argument so that we can introduce some worries about their suggestion that belief and belief-like imaginings share a single code.

3.3 The Imagination Box

We saw in section 2.3 that Nichols and Stich symbolise the idea that pretence involves a distinct attitude, which is a voluntary counterpart to belief, by saying that children possess an 'imagination box' where pretend representations are stored. This allows them to explain why these representations are quarantined from our ordinary beliefs: pretence representations are located in this imagination box; whilst our beliefs reside in our belief box. Langland-Hassan draws out this merit of their account, and other distinct attitude accounts, by noting that:

⁸¹ This being said, Nichols and Stich (2003, p. 27) note that there are various limits to this sort of cognitive quarantining. For one thing, we are often very good at remembering the events we pretended to occur, and forming second order beliefs about what we have pretended. Gendler (2006) discusses cases where this quarantining seems to break down, and we treat things we appear to merely imagine like beliefs.

[t]he existence of a distinctive imaginative attitude helps explain how one can believe that p while imagining (or pretending) that not p , or imagine (or pretend) that p while believing that p . For, provided that imagining is, like desire, its own cognitive attitude, there is no obvious epistemic difficulty—no threat of “inferential chaos”—presented by the fact that one imagines that p while believing that not- p , just as there are no such problems inherent in believing that p while desiring that not- p . (Langland-Hassan 2012, p. 161)

It is worth noting that a distinct content view of pretence can also explain quarantining. We can argue that pretence representations are quarantined because they have different contents to our ordinary beliefs, without needing to suppose that they reside in a separate ‘box’. In a later article, Nichols (2004, p.130) suggests that considerations related to *synchronous processing* rule out a distinct content based account of how we are able to engage in pretence.

Synchronous processing refers to fact that it looks like we can believe that p and also pretend that p (*qua* belief-like imagining) at the same time. Consider the tea party example. The child who was pretending to have a tea party pretended that one cup was empty and that the other was full, whilst also believing that both cups were empty. This makes it look as if the child was able to process representations with identical contents at the same time: they pretended that *the cup is empty* and also believed that *the cup is empty*. As such, Nichols and Stich argue the difference between our real and pretence representations cannot merely be a difference in content. Nichols puts the point as follows:

[a] pretense representation and a belief can have exactly the same content. So, pretense representations are quarantined from beliefs, and yet the distinction is not driven by differences in content. The natural cognitivist proposal, then, is that pretense representations differ from belief representation by their function. (Nichols 2004, p. 130)

I am not convinced by this argument. All the tea party example reveals is that *on the surface* it looks like we can have pretence representations with the same content as what we believe. For example, Leslie – the person who first introduced this empty cup example – would argue that the child believes that the cup is empty, and during pretence also believes that ‘I PRETEND that “the cup is empty”’. This special sort of content allows

Leslie to explain synchronous processing without introducing a distinct attitude. To put the point more generally, we can suggest that when a child takes it to be true that ‘the cup is empty’ in their game of pretence, this doesn’t have to mean they are straightforwardly entertaining a representation with the content ‘the cup is empty’.⁸² For now though, we can accept this point about synchronous processing for the sake of argument.

Accepting these arguments about quarantining and synchronous processing only gives us reason to hold that *some sort* of distinct attitude plays a role in pretence, not that this attitude is necessarily a counterpart to belief. These arguments also don’t inform us whether we should think of this counterpart as being voluntary or involuntary.

To elaborate on why they think the relevant attitude is a voluntary counterpart to belief, we can bring out that Nichols and Stich’s ‘imagination box’ will need to make sense of how initial pretence premises are *generated*. We must account for how the child pretending a banana is a telephone is able to generate the premise that *this [banana] is a telephone* and how the child in the tea party example was able to follow the instruction to pretend that they were having a tea party.

Nichols and Stich (2003, p.29) leave the mechanics of this somewhat vague, but Weinberg and Meskin (2006b, p. 182) fill in the gap here by postulating an ‘inputter’ that places representations into the imagination box either spontaneously or in response to being asked to imagine something. This notion of an inputter is supposed to represent one of the key functional differences between beliefs and belief-like imaginings, since the inputter cannot similarly input representations into the belief box at will. Ordinarily, I thankfully cannot choose to believe that a banana is a telephone.

This helps to illustrate why, if there is a counterpart to belief involved in pretence, it will be an example of a voluntary counterpart. Nichols and Stich have in mind here a mental state that we can in principle enter into at will. This is not to say, of course, that

⁸² Of course, they will have a belief with this content in the tea party example. My point here is that when the child says the cup is empty after seeing a cup being upturned, this doesn’t necessarily entail they now have two representations with the simple content ‘the cup is empty’.

is always entirely up to the pretender what is true in an episode of pretence: the point here is that power over what is true in a given episode of pretence is at least in principle possessed by the pretender. Going back to Walton's game where stumps represent bears, it is made true by the presence of stumps that there are bears nearby, and this fictional truth does not depend on the pretender's knowledge. However, even in this sort of rule-governed case, the child could declare that they want to pretend otherwise (as we saw in the CHILDES example where what the children are pretending constantly changes) and that therefore the stumps are no longer bears.

However, noting that we can produce pretence representations at will does not show that the distinct attitude involved in pretence is belief-like; it only shows that this attitude is not belief. A first argument for why this attitude is belief-like can be found by considering how this imagination box can make sense of pretence *elaboration*. Various inferences and elaborations are made in a typical game of pretence on the basis of the initial premise, and we must explain how these are made. Nichols and Stich (2003, p. 29) explain inferential elaboration by proposing that an 'inference mechanism, the very same one that is used in the formation of real beliefs, can work on representations in the [imagination box] in much the same way that it can work on representations in the Belief Box.'

In other words, the pretence scenario gets filled in after we generate the initial premise because our ordinary inference systems infer from the pretence premise what we would infer if we believed the premise were true. For example, if the pretence premise is 'I am a dog' we might infer this entails 'I should bark'. This also helps to justify the claim that belief-like imaginings share a single code with belief, since both attitudes are said to interact with our inference mechanisms in 'much the same way'. They elaborate on this point by suggesting that:

[r]epresentations in the [imagination box] have the *same logical form* as representations in the Belief Box, and ... their representational properties are *determined in the same way*. (Nichols and Stich 2003, p. 32, emphasis in original)

The idea that beliefs and belief-like imaginings have 'the same logical form' and that 'their representational properties are determined in the same way' is what the single code

hypothesis amounts to. This has the consequence that cognitive ‘mechanisms will process [pretence] representations in *much the same way* that they process beliefs.’ (Nichols and Stich 2003, p. 33, emphasis mine).

However, merely introducing inference systems that operate on the initial pretence premise doesn’t do enough to explain the sort of richness that characterises episodes of pretend play. For example, from the single premise ‘this [banana] is a telephone’ we cannot directly draw many inferences, and almost certainly we cannot draw enough inferences to guide an entire episode of pretence. As such, this does not fully explain inferential elaboration during pretence.

Nichols and Stich make sense of this by proposing that the contents of the belief box are placed in the imagination box during pretence, and that representations in both boxes are governed by an *updater*. In relation to this first proposal, they suggest that during an episode of pretence the imagination box is filled not only with the initial pretence premise (or premises) but also with every representation from the belief box (Nichols and Stich 2003, p. 29). Since these imported representations are in the same code as the pretence premise, this allows Nichols and Stich to explain why we are able to make inferences during episodes of pretence that go beyond the initial premise: we can draw inferences based on how our pretence premise relates to our ordinary beliefs about the world.

That being said, Nichols and Stich note that there is initial difficulty with this suggestion, namely that it will lead to there being contradictory representations in the imagination box (Nichols and Stich 2003, p. 30). For example, during the tea party example, the imagination box will contain the representation that *one cup is full and one is empty* along with the conflicting representation that *both cups are empty*, since this representation reflects our actual belief.

To make sense of this, Nichols and Stich (2003, p. 30) point out that our ordinary beliefs seem to be governed by some sort of ‘updater’. For example, if you believe that the Earth is flat, and I show you a photograph of a round Earth, you will (hopefully!) automatically update your belief box by getting rid of your representation that *the Earth is flat* and replacing it with a representation like *the Earth is round*. This mechanism’s

inner workings are unknown to us, but it seems to work to revise our beliefs on the basis of our having new visual experiences, or hearing testimony we take to be reliable, and so on.

They argue that this updater also operates on the contents of the [imagination box], and that it does so in much the same way as it operates on contents of the belief box (Nichols and Stich 2003, pp. 31-32). This further illustrates why they argue the representations found in the belief and imagination boxes share a single code.

Nichols and Stich (2003, pp. 31-32) suggest that there are two ways in which this copying of contents from the belief to imagination box could happen. Firstly, the updater might act as a sort of filter, stopping beliefs being placed in the imagination box if they conflict with the initial pretence premise or premises. Secondly, it could remove beliefs that are found to conflict with the initial premise or premises. On both pictures, the end result will be the same: we do not have outright contradictory representations in our imagination box during pretence.⁸³

Even having introduced the idea of an updater, there are still some aspects of elaboration that remain unexplained, since, as we already noted, many elaborations in pretence episodes are non-inferential. For example, we still have no explanation as to why the child in the dead cat example says he ‘got shot’ or why one of the adult pretenders cut someone’s head off with a sword.

In order to account for such non-inferential elaborations, they introduce another mechanism into our cognitive architecture. They call this component the ‘Script Elaborator’, which fills in the details of an episode of pretence that ‘cannot be inferred from the pretence premiss, the (updater-filtered) contents of the Belief Box and the pretender’s knowledge of what happened earlier in the pretence.’ (Nichols and Stich 2003, p. 35). This piece of cognitive architecture only interacts with the imagination box, and so is meant to illustrate a functional dissimilarity between the belief and imagination boxes.

⁸³ At least we won’t have them because of a background belief being present in the imagination box. It might still be the pretence scenario asks us to take contradictory representations to be true.

At this point, Nichols and Stich have offered us an account of the second feature of pretence, elaboration, since this is explained by the inference mechanisms, updater, script elaborator, and more generally, the notion of a single code. We also have a partial explanation of set-up, since we saw the *generation* of a premise is made possible by the inputter. The inputter can also explain how some forms of *recognition* occur, namely ones where the child is told explicitly what the pretence premise is. At this point, however, we haven't been given an account of how children figure out what another individual is pretending when the premise isn't made obvious. We also have a partial explanation of the third feature, output, since Nichols and Stich are able to explain quarantining by drawing attention to the fact that pretence representations involve a distinct cognitive attitude. However, we still have no explanation of how affective responses to what we pretend are generated, nor of how pretend actions are motivated.

In relation to motivation, Nichols and Stich propose that the motivation for performing pretend actions stems from the combination of three kinds of mental states: imaginings, beliefs and desires. Children perform pretend actions because they imagine that they are something, or that they are doing something; they form desires to act similarly to how the thing they are imagining would act, or to how they would act if they were doing what they imagine doing; and they form beliefs about how the thing they are imagining would behave if it were real, or how they would behave if they were doing what they imagine themselves to be doing (Nichols and Stich 2003, pp. 37-38). This means something like the following combination of mental states is supposed to motivate pretence behaviour:

1. Propositional imaginings (that I am phoning grandma, that I am having a tea party, that I am a dead cat, etc.)
2. Beliefs about how things behave (that you speak into phones, that tea parties involve serving tea, that dead cats don't move, etc.)
3. Desires to behave in a way that would be appropriate if your imaginings reflected what was actually the case (a desire to behave in a way that would be appropriate if *I am phoning grandma*, a desire to behave in a way that would be appropriate

if *I am having a tea party*, a desire to behave in a way that would be appropriate if *I am a dead cat*, etc.)

These combinations lead to pretend actions like talking into a banana, pouring out pretend tea and lying on the floor like a dead cat. The desire element helps to explain the general motivation for engaging in any sort of pretence behaviour, and the belief component is supposed to explain why appropriate pretence actions are performed. This means belief-like imaginings themselves do not directly motivate behaviour, allowing Nichols and Stich to argue that imaginings do not have a direct connection to our decision making and motivational systems. This is said to reveal a further functional difference between belief and imagination.

There are numerous ambiguities in this sketch of the motivation of pretend actions. For example, there is a question about how closely beliefs about how things behave track appropriate pretend actions (e.g. cats don't really make the meowing noise humans do when pretending to be cats, and as we noted earlier dead cats don't say 'I'm dead' or that they 'got shot'). For now, we will leave these worries to one side, and will return to issues regarding motivation in Chapter 4, where they will serve to illustrate one of the benefits of adopting Leslie's distinct content view when it comes to explaining how we are able to engage in pretence.

As for recognition when the pretence premise isn't made obvious, they argue that this involves recognising someone is behaving in a way that would be appropriate if such-and-such were the case (Nichols and Stich 2003, p. 53). However, as with their account of motivation it is ambiguous how this will work in practice, and I will set out why this will struggle to account for pretence recognition in chapter 4.

In relation to why we exhibit affective responses to what we pretend, Nichols and Stich propose that in light of beliefs and belief-like imaginings sharing a single code, they are processed by our affect generation systems in a similar way, and so belief-like imaginings can lead to affective responses. The single code hypothesis thus allows us to make sense of affect in the same sort of way it allowed us to explain inferential elaboration.

With these developments and elaborations of their theory in place, it looks like the introduction of a distinct propositional attitude combined with the hypothesis that it

shares a single code with belief, allows us to explain the three features of pretence that Nichols and Stich draw our attention to. However, a serious worry with this approach arises if we reflect on precisely what is entailed by two mental states sharing a single code. To see why this notion is potentially problematic, we should begin by noting some differences between ordinary beliefs and the sorts of representations involved in pretence.

3.4 Differences Between Belief & Belief-Like Imaginings

If we accept Nichols and Stich's account of our cognitive architecture for the sake of argument, difference between beliefs and belief-like imaginings can be placed under three headings⁸⁴:

1. Input Asymmetries

This is where a mechanism only takes representations produced by either the belief box or imagination box as input. For example, Nichols and Stich argue that our decision-making and motivational systems only take beliefs as input, since propositional imaginings cannot motivate us directly.

2. Output Asymmetries

This is where a mechanism only outputs representations into either the belief box or imagination box. For example, it looks like the inputter can only output belief-like imaginings, since beliefs are not under the control of the will.

⁸⁴ I have adapted these headings from Weinberg (2013, pp. 188-190) who also suggests we need a fourth heading here, 'phenomenological' asymmetry. This is because he notes that from a young age we seem to be able to tell the difference between what we imagine and what we believe. As I argued in chapter 2, I think we can equally explain this difference when thinking about pretence and fiction by saying we can tell the difference between what we believe to be fictional and what we straightforwardly believe.

3. Processing Asymmetries

This is where a mechanism processes representations in the belief box and imagination box differently, or at least processes them differently in some instances. For example, the updater seems to resist getting rid of our initial pretence premises when we are engaging in pretence.

Now, on the face of it, one might think that these three sorts of asymmetries are readily explained if we accept Nichols and Stich's cognitive theory of imagination. Since they argue that propositional imaginings are a counterpart to belief, we should expect differences in input, output, and processing: it is these differences that demonstrate imaginings have a distinct functional role as compared to belief, after all.

We can call this the 'functional role' explanation of why these asymmetries are present. This sort of response will entail holding that it is just a peculiarity of the differing functional roles of beliefs and belief-like imaginings that leads to differences in processing, input and output, rather than something we can systematically explain. For example, maybe some mechanisms *just do* produce only imaginings, and maybe others *just do* process imaginings and beliefs differently sometimes. In other words, we defend the theory by suggesting functional roles are somewhat arbitrary and that the single code hypothesis shouldn't be taken as being all that explains how our cognitive mechanisms process imaginings. This sort of response is hinted at by Nichols in relation to differences in input and output when he tells us that:

[i]f pretence representations and beliefs are in the same code, *then mechanisms that take input from* the [imagination] box and from the belief box will treat parallel representations much the same way. (emphasis mine) (Nichols 2006, p. 461)

For example, we have noted that pretence episodes are ordinarily initiated when a pretence premise (or set of premises) is placed into the imagination box by the inputter. This reflects the fact that imaginings are subject to the will while beliefs are not – the inputter cannot place a representation into our belief box at will. For example, I can choose to imagine that David Cameron is a lizard, but I cannot choose to believe this. The functional role response doesn't give us a deep explanation of why this is so, but we

still at least have some sort of explanation of why this asymmetry in input is present. A similar suggestion can also help to make sense of differences in output: belief-like imaginings cannot motivate actions directly because there just isn't a pathway from the imagination box to our various decision-making mechanisms.

However, endorsing this sort of response highlights a worry about the explanatory merits of 'boxological' accounts of the mind. There is a risk that if we endorse this sort of approach we end up labelling distinctions between mental states and attitudes without actually explaining them.

Regardless of whether one finds this a satisfactory explanation of differences in input and output, this functional role response is less helpful as an explanation of processing differences. Indeed, Nichols goes on to note that:

[i]f a mechanism takes pretense representations as input, the single code hypothesis maintains that if that mechanism is activated by the occurrent belief that *p*, it will also be activated by the occurrent pretense representation that *p*. More generally, for any mechanism that takes input from both the pretense box and the belief box, the pretense representation *p* will be processed much the same way as the belief representation *p*. (Nichols 2006, p. 461)

Accounting for differences in processing is challenging for Nichols and Stich thanks to their emphasis on the relationship between processing and code. It looks like they maintain that differences in processing ought to be explained by differences in code, since it is sharing a single code that leads to beliefs and belief-like imaginings being processed in a similar way. This leaves it an open question whether Nichols and Stich think that differences in functional role between two representations can affect processing. If not, it is unclear why there are processing asymmetries between beliefs and belief-like imaginings.

We can bring out the worry here by making use of a reconstruction of part of Nichols and Stich's theory by Stock (2011a pp. 272-274):

1. A mechanism is of type *M* iff it is a member of the set of mechanisms which realise/contribute to the realisation of *D* in the entities it processes
2. In certain contexts, imagining and believing that *p* each realise *D*.

Therefore:

3. In those contexts, imagining and belief that p each are processed by a mechanism of type M.

In this reconstruction, Stock gives an account of how we justify the introduction of a piece of cognitive architecture and settle on its nature in premise (1), where she formalises the idea that mechanisms ought to be defined by what they realise (e.g. decision-making systems realise decisions, affect-generating systems realise affective responses). She then goes on to formalise the idea that this means beliefs and imaginings interact with at least some of the same mechanisms in (3). Stock thinks this view becomes problematic when we note that Nichols and Stich appear to endorse a further commitment, namely that:

4. No mechanism of type M makes any discrimination between imagining and belief as input.

This claim is problematic since there can be processing differences between beliefs and imaginings. To set out the scope of this worry, we should begin by noting two specific processing differences between belief and what Nichols and Stich are calling propositional imaginings (Everett 2013, pp. 10-13) offers a more detailed list of further purported differences). Setting out these two asymmetries will involve introducing some considerations that arise from thinking about how Nichols and Stich's theory relates to fiction and philosophical thought experiments, since Nichols (2006) introduces one of these asymmetries in relation to worries about fiction and hypothetical reasoning.

1. Affective Processing

We have already noted that sometimes our emotional responses to things we pretend to be the case are much the same as the responses we would exhibit to similar beliefs. This is easiest to bring out if we introduce some examples related to fiction rather than to

pretence. One might pity (or perhaps quasi pity) Anna Karenina in much the same way that one would pity a real woman who suffered her fate.⁸⁵

However, these similarities seem to break down in other cases. This can happen in two ways (Nichols 2006, pp. 464-465). Firstly, there are cases of *absent* affect or judgement. This sort of case is best demonstrated by considering philosophical thought experiments. Take Mary, the neuroscientist locked in a black and white room her whole life. When we engage with this thought experiment, we don't find ourselves distracted from questions about qualia because we feel sad that Mary has lived in this drab room her whole life, nor do we feel a sense of moral approbation towards her carers for apparently having locked her up in such bizarre and cruel way.

Nichols (2006) explicitly associates this sort of hypothetical reasoning and the use of thought experiments with the imagination box, but if one is not convinced by this link (perhaps because one wishes to distinguish mere supposing from belief-like imagining) we can also find these sorts of examples in works of narrative fiction, such as videogames. Many children's videogames contain elements that might be expected to lead to some sorts of negative emotional responses but due to their cutesy presentation, do not. For example, in the *Kirby* video game series Kirby, the titular avatar (who is more or less just a cute, constantly smiling, pink blob), sucks up enemies so that he can steal their powers and use them to solve puzzles.⁸⁶ I've yet to hear of any children (or parents) who have responded with horror or disgust upon seeing Kirby suck up an enemy on-screen. To give another example, so far as I can tell, there has also not been any moral panic

⁸⁵ A similar point can be made about our moral processing. One might judge that Lady Macbeth behaves immorally in much the same way that one would judge that a real woman who behaves like her is behaving immorally. This is part of what guides an interesting question about whether there are things we cannot imagine, or perhaps resist imagining (Gendler 2000) or perhaps which we cannot take to be fictionally true (Walton 2006).

⁸⁶ For example, you might suck up an enemy who is carrying a hammer in order to gain the ability to swing a hammer around. This sometimes allows you to break certain blocks in the game in order to find hidden treasures.

instigated by Mario jumping on – and thus killing – turtle-like creatures in his various videogame adventures.

However, if we learned of a real person who grew up in Mary's circumstances, we would probably respond emotionally to their plight. As for our videogame examples, we are (at least) sometimes disgusted and horrified by the thought of eating other sentient beings alive, and would likely not approve of a plumber who went around killing turtle-like creatures by jumping on top of them. These examples thus bring out that our emotional responses to what we imagine do not always behave in 'much the same way' as our responses to what we believe.

Secondly, there are cases of what Nichols (2006, p. 464) calls 'discrepant' affect. He gives the example of the film *Dr. Strangelove*, but his worry relates to many other black comedies. In *Dr. Strangelove*, the fate of the world is threatened by all out nuclear war and this leads to several humorous scenes in the American War Room, such as one where various generals discuss military options with the President of the United States. However, while we might laugh at this scene when engaging with the fiction, we certainly would not do so if we learnt about such happenings in the real world. If these imaginings about *Dr. Strangelove* are in the same code as our beliefs, why do our affect generation systems appear to process them differently? Videogame examples are also helpful for illustrating this sort of discrepant affect. While playing first-person shooter games, one might take great pleasure in shooting various enemies, while in real-life being horrified by these sorts of actions.

2. Clustering

An important point made by Walton (2015, pp. 18-19) about our representations related to works of fiction and episodes of pretence is that they exhibit 'clustering'. For example, belief-like imaginings about *Romeo and Juliet* concern only that particular fictional world, and representations about tea parties concern only that particular pretend tea

party.⁸⁷ This contrasts with belief, since our beliefs all seem to be interconnected in a single ‘cluster’ and aim to reflect the truth of the actual world. If I believe that if p then q , and believe that p , I should refrain from believing q in any situation unless I come to believe that p is false.

With fiction and pretence, this sort of inference only holds if p and q belong to the same cluster. If I believe that *if all men are mortal then Jeremy Corbyn is mortal*, I should only cease to believe that *Jeremy Corbyn is mortal* if I realise either he is not a man or that some men are indeed immortal. If, on the other hand, I engage with a work of fiction where some men are portrayed as immortal, I can happily take this to be true in the world of that fiction, whilst recognising that it won’t be true in other fictional worlds or in the real world. This clustering of our representations related to fiction and pretence thus allows us to explain why we can sometimes take on conflicting representations related to works of fiction. This is also true in relation to specific episodes of pretence: if I pretend that ‘that [banana] is a telephone’ I won’t take this to be true in every game of pretence.

If Nichols and Stich are right that these representations are belief-like imaginings, then why can they belong to clusters when beliefs cannot? For example, in the world of *Midnight’s Children*, we take it as fictionally true that people can read minds, while in the world of *Disgrace* we take it as fictionally true that such supernatural feats are impossible. But if an imagining about *Midnight’s Children* shares the same code as an imagining about *Disgrace*, how does this clustering come into play? It will be problematic for Nichols and Stich to hold that imaginings can be marked as belonging to a certain cluster while beliefs cannot be so marked, since this will risk constituting a systematic difference at the level of content between beliefs and imaginings.⁸⁸

⁸⁷ As noted earlier, Walton’s notion of imagining does not amount to mere belief-like imaginings. However, Nichols (2004) takes belief-like imaginings to be the relevant sort of imaginings here and so presumably wouldn’t object to my characterising our engagement with fiction in this way.

⁸⁸ One suggestion here could be that imaginings are only ever to be found in our imagination box on a temporary basis, and that different clusters thus switch in and out in-line with

To see how we might make sense of these processing difference, we will consider two responses in what follows, the functional role response and the desire response. I will argue that neither of these responses offers a convincing explanation of these asymmetries.

3.5 Dealing with Asymmetries

1. Functional Role Revisited

We might be able to make the initial functional role response work as an account of processing differences by drawing an analogy between our cognitive systems and a vending machine. Picture a machine that gives out either cans of Coke or cans of Pepsi. Let's say that this machine represents our affect-generation systems. To simplify matters, we'll say that Coke represents 'positive affect' and Pepsi for 'negative affect'. This machine accepts various coins, but only outputs one of these cans if you insert a single pound coin.

There are two slots on the machine. Put a pound coin in one slot and you get a Coke. Put it in the other and you get a Pepsi. Now, there are two people who use this (rather odd) vending machine every day. These two people can be said to represent the *belief box* and the *imagination box*. They both have wallets filled with various coins, and every day they both put a single pound coin into the machine. We can say that their respective pound coins share a *single code*. Ordinarily, they both put their pound coin into the Coke slot. However, on one occasion, for no particular reason, one of our Coke-loving individuals just happens to put his pound into the Pepsi slot and is provided with a Pepsi.

our imaginative project. This would introduce an odd puzzle about what happens when we read a novel whilst the TV stays on in the background playing a film we're half watching. It seems odd to suggest that switching between the TV and your book involves rapidly filling and emptying your imagination box as we pay more attention to one or the other fiction.

In this example, there is no principled explanation we can offer for why one Coke lover has a Pepsi this time round: it just so happens that they put their pound coin into a different slot. We could propose that imaginings and beliefs also just are processed differently on occasion. There is no elaborate reason we can give for why one happened to interact with the affective systems differently to normal when engaging with *Dr. Strangelove*, it just so happens that a slightly different interaction to normal has taken place even though the relevant representations remain in the same ‘code’. The problem with this suggestion, however, is that it doesn’t seem to be merely random that this difference in processing occurs: it looks like there is some sort of reason why we laugh at approaching nuclear war while watching *Dr. Strangelove* when ordinarily we would be horrified by this prospect. Allowing that representations sometimes just happen to be processed differently does little to help explain the differences in processing we set out above.

2. Desire

A second proposal comes from Nichols (2006, pp. 469-472), who suggests that we can find a better explanation of differences in processing by looking at the desires associated with our beliefs and imaginings. In relation to discrepant and absent affect cases, he argues that our desires are behaving differently to the way in which they would behave if we had beliefs rather than imaginings.

For example, in cases of absent affect like the Mary thought experiment, we can suggest that in order to feel sorry for Mary you would need to have a desire for her not to have been raised in such restrictive conditions. With this desire absent, the affect is also absent. Nichols (2006, p. 471) explains this point with the following two examples:

1. Someone walks into your office and says: Imagine that everyone outside of this room is dead, what would Utilitarianism say about the importance of our interests?
2. Someone walks into your office and says: Everyone outside of this room is dead; what does Utilitarianism say about the importance of our interests?

He suggests, plausibly, that your affective response to (1) will be very different to (2). He argues that this is because of the different desires associated with these two examples:

In both (1) and (2) we are encouraged to have the desire to figure out the entailments of Utilitarianism. In case (2), our desires about the real world would swamp any desire we have to work out the Utilitarian calculus. Hence, our inferences and recollections would be guided by these prevailing desires. In case (1), by contrast, we plausibly have no particularly pressing desires about the inhabitants of the imaginary scenario. As a result we are not compelled to draw the inferences and recollections that would follow in case (2). Rather, our desire to answer the question about Utilitarianism can be pursued without the intrusion of salient desires and concerns about the inhabitants of the imaginary situation. (Nichols 2006, p. 471)

His thought here seems to be that we don't always form desires related to the wellbeing of the inhabitants of fictional or imaginary worlds. On the other hand, when we engage with the real world, we cannot simply 'detach' our standing desires to preserve the lives of our loved ones and so on. As he puts the point:

Imaginings can be constrained, filtered and directed in all sorts of ways that are not available to beliefs. For our desires about the imaginary scenario will depend on the context, the intent of the author, the tone of the work, the point of the thought experiment, and so on. Our desires about the real world are much less flexible. (Nichols 2006, p. 472)

For discrepant affect cases, like *Dr. Strangelove*, he suggests that a different desire is present as compared to if the scenario occurred in real life, perhaps a desire to be amused, or some internal desire about the contents of the fiction:

When it comes to black comedy, we typically do not have such powerful desires for the preservation of human life in the imaginary scenario. Hence, we are not compelled to draw out disturbing inferences like billions of innocent people will die horrifically painful deaths. Rather, genre considerations make us want to focus instead on Slim Pickens' exuberant missile ride. (Nichols 2006, p. 472)

If this is right, then the discrepant affect occurs because the desire involved is different to the one that would combine with a similar belief (which would presumably be a straightforward one about avoiding the horrors of nuclear war). However, it's not entirely clear whether this proposal preserves the single code hypothesis. If we embrace this

response, we now face the question of why our desires sometimes interact differently with our imaginings as compared to our beliefs. For example, Nichols argues that:

The explanation for the asymmetries is not that the affective mechanism itself responds differently to imagining that *p* and believing that *p*. Rather, the asymmetries arise because the affective mechanism is sent quite different input depending on whether one imagines that *p* or believes that *p*. (Nichols 2006, p. 472)

But, he tells us little about *why* this difference in input is present; we are merely told that it is present thanks to it being possible for imaginings to be ‘sparse’. In particular, we have no explanation of why sometimes our desires interact with our imaginings in the same way that they interact with our beliefs, whilst in other cases they do not. For example, in their discussion of belief-like imaginings and motivation, Weinberg and Meskin suggest that:

[t]he fact that our imaginatively driven responses do not result in the full gamut of behavioural responses can be explained by ... pointing out the obvious fact that many behaviours require relevant motivational input ... While a belief that one is being threatened by a tiger will typically interact with a (standing) desire not to be harmed and result in flight behaviour, imagining that one is threatened by a tiger does not interact in the same way with that standing desire. (Weinberg and Meskin 2006b, p. 184)

This claim about motivational input and affective responses seems plausible, but this point does not explain why imaginings have a complex relationship with desire and why this should be seen as compatible with the idea that beliefs and belief-like imaginings share a single code. This quote seems to imply a systematic difference in how imaginings and beliefs relate to desire, and it is unclear why this difference would be present if these representations share the same code. Stock shares a similar worry, noting that:

[t]he single code theory ... looked committed to arguing that it was precisely a representation’s code which determined which inference and affective mechanisms it interacted with, and to what extent. Yet here Nichols seems to suggest that two identically code representations can ... give rise to different affect, depending on what desire are concomitant ... Insofar as the original view held that imagining and belief that *P* produce similar output, Nichol’s claim that different desires can accompany a belief and an imagining that *p* ... requires further explanation. (Stock 2011a, p. 278)

What the desire response helps to capture, however, is the intuition that how we respond to a given representation related to a fiction or episode of pretence will depend on how it relates to other representations. This is perhaps why Nichols noted that imaginings can be ‘filtered’ and ‘constrained’ in various ways. I will draw on this point to offer my own response to these processing differences. This will not allow us to salvage the single code hypothesis, but will tell us something important about how the representations involved in pretence and fiction behave.

3.6 Clustering Revisited

We have already noted that the representations involved in pretence and fiction are subject to clustering. One important consequence of this is that how we respond to a given representation associated with an episode pretence or a work of fiction will depend on which other representations are associated with it in a given cluster.

Suppose, for example, a pretender has formed a representation like *a bear is nearby*. If one is playing a game where they’re pretending to be a bear, this representation won’t lead to a fear (or quasi fear) response, in fact it might make them happy because they realise their fellow pretend-bears are nearby. If, on the other hand, one is playing a game like the one in Walton’s bears example, then this representation may well lead to a fear response since it will be associated with representations related to being human, needing to hide from bears, and so on.

The reason Nichols and Stich’s distinct attitude theory struggles to explain differences in processing is because they focus for the most part on single representations associated with fiction and pretence and the features these representations do and do not share with ordinary beliefs. To explain processing asymmetries, we need to pay attention to the entire cluster of representations related to engaging with a work of fiction or engaging in an episode of pretence.

However, Nichols and Stich’s imagination box combined with the single code hypothesis is ill placed to explain why clustering occurs. In the case of the imagination box, the idea that we place our imaginings in a ‘box’ doesn’t account for the fact that we will have to associate representations in the box with various different fictions (Walton 2015,

pp. 18-19). This is because simply putting a representation related to fiction into this sort of mental workplace doesn't explain how we come to associate it with a particular work of fiction or episode of pretence. It is not enough to just take some sort of belief-like attitude towards something we take to be true in a work of fiction or episode of pretence; we also have to recognise that we are engaging with a *specific* work of fiction or in a *specific* episode of pretence.⁸⁹ This reflects one of the difficulties that arises if we try to make a single belief-like attitude do all the explanatory work in explaining hypothetical reasoning, pretence and our engagement with fiction. Placing a proposition into this box might be sufficient for coming to suppose that p , but it is not so immediately clear how it will explain coming to take p to be true in fiction.

Furthermore, Nichols and Stich's commitment to the single code hypothesis makes it difficult to find a way to accommodate clustering in their distinct attitude theory. For example, one-way Nichols and Stich could attempt to explain why clustering occurs would be to argue that in fact, this doesn't reflect an asymmetry with belief. Beliefs are in some sense clustered because they are implicitly subject to an operator like *in the real world*, whereas imaginings can be subject to a variety of these sorts of operators, such as *in War and Peace* or *in Romeo and Juliet*. This sort of response will put pressure on the single code hypothesis, however, unless Nichols and Stich can explain why this doesn't amount to there being differences in content between beliefs and imaginings.

As such, this reveals an initial way in which we can question the non-doxastic assumption, since introducing a distinct attitude is *insufficient* for explaining this aspect of our engagement in pretence and with fiction, since merely introducing a belief-like attitude will not explain clustering. I will reflect a bit more about the implications of this in chapter 5.

⁸⁹ Nichols and Stich could argue in response to this worry that there is something distinctive about fictional contents which leads to them being processed in a different way to other representations in the imagination box, but they have not given us a mechanistic explanation of what might lead to this sort of difference in processing.

3.7 The Role of Boxology

At this point, a certain sort of response can be anticipated in defence of Nichols and Stich, which is that the boxological account of the mind has been misunderstood and that I have been too demanding in my expectations of its explanatory power. Weinberg, responding to Stock, argues that:

[o]ne need not think of mechanisms as the sort of thing that have one characteristic function that applies the exact same way in all cases, without any exceptions, contextual variations and so on ... These theories [boxological theories] are marking out large and stable trends in the functioning of the cognitive systems that they are characterising, and the worth of such theories is found in such terms, and not in anything like metaphysical precision. (Weinberg 2013, p. 187)

Weinberg (2013, p. 188) would charge me with trying ‘to apply [my] standard philosophical tools to theories whose home is more in the sciences.’ I shouldn’t be expecting a robust explanation of every processing difference or difference in input/output between belief and belief-like imaginings since Nichols and Stich are merely aiming to identify ‘large and stable’ trends related to how the propositional imagination functions. This looks somewhat similar to my emphasis on the importance of clustering, since Weinberg specifically mentions the importance of ‘contextual variations’.

However, this sort of move calls the explanatory value of Nichols and Stich’s cognitive theory of imagination into question. Understood in this way, the theory has limited explanatory and predictive power, since it only predicts things we already know happen (e.g. that we sometimes respond emotionally to things we imagine and sometimes do not) and it does little to explain why these things happen. If all this complex philosophical machinery only serves to explain and predict things we can already explain and predict with ordinary folk psychology and philosophical reflection, one wonders why we ought to accept their cognitive architecture and why we should try to make sense of mysterious notions like that of a ‘single code’.⁹⁰

⁹⁰ Stock (2011a) shares a similar concern.

Furthermore, this defence misrepresents the nature of the challenge which is introduced by Nichols and Stich maintaining that belief-like imaginings and beliefs share a single code. If they merely argued that imaginings were functionally distinct from beliefs but with some important differences, this defence would ring true (albeit leaving us with a lingering worry about the explanatory value of the theory). It is Nichols and Stich's focus on the notion of a single code that allows these worries about processing to come to the fore, since they seem to imply that there should be no exceptions to how imaginings are processed as compared to beliefs. Indeed, issues related to processing do seem to be viewed by Nichols as a problem for his theory, at least in relation to differences in processing that relate to the same mechanism, since he says that this is 'what the single code hypothesis says won't happen' (Nichols 2006, p. 465).⁹¹

Since this theory risks being rendered unilluminating, we are left with three ways we could proceed in trying to explain how children are able to engage in pretence. Firstly, we could try to develop a theory of pretence that relies on introducing a belief-like state, but which either doesn't depend on the idea of a single code or substantially weakens this notion. We could then investigate if this more minimal theory can still explain the three features of pretence we noted. However, this project has in some sense already been undertaken by simulation theorists like Currie (1995) and Gordon and Baker (1994), and

⁹¹ This also relates to a worry raised by Goldman (2006, pp. 282-283) about whether belief and desire share the same code since they interact with many of the same cognitive mechanisms. Nichols denies this in personal communication with Goldman (2006, Ch. 11 footnote 1). Goldman finds his reply puzzling noting: 'After reading a previous draft of this material, Nichols (personal communication) commented that he doesn't think that desires and beliefs use the same code. That's because he takes code talk as a metaphor for the "computational features" of a representation. This raises many delicate issues, and I confess that I lose my grip on Nichols's code talk at this juncture. The view needs to be spelled out in more detail, including a spelling out of the entire boxology architecture.' I share Goldman's concern here. There is something seductive about the idea that beliefs and imaginings share a single code, but once we submit the notion of code to closer inspection it becomes unclear precisely what this claim amounts to, particularly if Weinberg is right that the notion of code doesn't commit us to thinking isomorphic representations will always be processed in a similar way.

their theories of pretence have been convincingly critiqued by Nichols and Stich (2003, pp. 39-47) for being overly vague. Indeed, this is part of what motivated them to introduce their heavy philosophical machinery and the notion of a single code.

The notion of a single code does a good deal of heavy lifting in explaining set-up, elaboration and output, so if we adopt a distinct attitude view along the lines of Nichols and Stich's but weaken or abandon the notion of a single code, we risk endorsing a view that merely stipulates how to explain these three things. Take for example the question of why we respond emotionally to things we merely pretend to be the case. The single code hypothesis let us say this is because beliefs and belief-like imaginings are processed in much the same way by our affective systems. If we now maintain that this won't always be the case, this means that all the theory tells us is that sometimes we will respond emotionally to what we pretend in much the same way as we respond to what we believe, and sometimes we won't. Furthermore, if make this move, it will remain unclear how we should make sense of clustering without endorsing something like a distinct content and distinct attitude view of pretence. If we reject the single code hypothesis but continue to argue that belief-like imaginings have the same contents of our beliefs, this will struggle to account for how our representations related to fiction and pretence come to be associated with particular fictions and episodes of pretence.

Secondly, we could consider whether we can explain pretence by introducing a *really distinct* attitude which isn't a counterpart to another mental state. Bearing in mind the similarities between the representations involved in pretence and our ordinary beliefs we have noted, this will be a difficult approach to make work, so I will not attempt it in this thesis.

The final way to proceed in our investigation is to look at whether we can explain these three features of pretence by focusing on *content* rather than attitude. This is what we will do in the next chapter where I argue that this sort of approach is the only one that can offer a compelling account of pretence recognition and motivation.

Conclusion

In this chapter, we saw in sections 3.1 and 3.2 that we can propose that any satisfactory theory of pretence needs to explain set-up, elaboration and output. In section 3.3, we discussed Nichols and Stich's arguments for why the best way of making sense of these three features is by introducing a distinct belief-like attitude. We began the discussion of their theory by noting that their arguments based on quarantining and synchronous processing do not do enough to show that it is necessary to introduce a belief-like attitude into an account of how we engage in pretence. Nonetheless, accepting that their 'imagination box' is involved in pretence for the sake of argument, we saw that this posit does manage to explain these three features of pretence, albeit with some lingering worries about how their accounts of motivation and recognition will work in practice.

However, we then saw in section 3.4 that their notion of a 'single code' makes it difficult for them to explain processing differences between belief-like imaginings and beliefs. I suggested in section 3.6 that this is particularly true in relation to clustering, which I argued reveals an initial way in which we can challenge the non-doxastic assumption, since belief-like imaginings appear to be insufficient for explaining why our representations related to fiction and pretence exhibit clustering.

Having considered these worries with Nichols and Stich's framework, we should now turn to consider whether it is more plausible to embrace a distinct content theory of pretence. We will do this by reflecting on recognition and motivation.

Chapter 4: Pretence Recognition & Motivation

Introduction

We will now discuss some further challenges for Nichols and Stich's distinct attitude view of pretence: that it struggles to account for pretence *recognition*, and in a related way the *motivation* of appropriate pretend actions. To introduce this next stage of our discussion of pretence, it will be helpful to make a distinction between different kinds of theories of pretence aside from the one I made earlier between distinct content views and distinct attitude views. Broadly speaking, theories of pretence can be placed under two headings. Firstly, we have behavioural theories of pretence, of which Nichols and Stich's is an example. These sorts of theories propose that engaging in pretence involves children engaging in certain sorts of behaviour. Secondly, we have mentalistic theories of pretence. These sorts of theories maintain that in order to be able to pretend, children have to recognise something about the mental states that motivate pretence behaviour. The most prominent defender of this approach is Leslie (1987, 1994), who we saw in chapter 2 argues that pretenders possess the mental state concept PRETEND, and this is what allows them to engage in and recognise pretence.

This distinction between behavioural and mentalistic cuts across the one we have made between distinct attitude and distinct content views of pretence. Langland-Hassan offers a behavioural, distinct content view to explain pretence, whereas Leslie offers a mentalistic distinct content view. A common worry is that mentalistic theories of pretence attribute too much cognitive sophistication to young children because they suggest that children can form meta-representations while they still fail standard false belief tests. As such, in this chapter I will attempt to show why an account that attributes what might seem to be a surprising degree of sophistication to young children is necessary for explaining pretence. It is worth highlighting here that behavioural theories do not tend to dispute that adults and older children might have a mentalistic understanding of pretence. The issue is whether a behavioural account is sufficient for explaining younger children's understanding of pretence.

In section 4.1, we will set out Nichols and Stich's approach to motivation and recognition in more detail. In section 4.2, we will then consider an argument from Friedman and Leslie which suggests that Nichols and Stich's behavioural approach cannot explain recognition, before setting out why it looks like a version of Leslie's distinct content mentalistic view that focuses on intentions to pretend can explain these two issues in section 4.3. This will reveal that if we allow for meta-representational beliefs to be involved in pretence, one could argue that we also need to introduce a distinct attitude to explain how pretence premises can be generated at will. I will draw on some helpful considerations from Langland-Hassan to argue that this proposal can be avoided. This will show that it is unnecessary to introduce belief-like imaginings into our theory of pretence.

Having done this, in section 4.4 we will then discuss Langland-Hassan and Stich and Tarzia's recent attempts to respond to these worries by refining the behavioural view. These behavioural accounts both suggest that recognising that someone is pretending involves recognising that they are making some X saliently Y-like. However, I will argue that it is not clear how this is supposed to explain how children recognise the *content* of pretence episodes, or how they figure out what pretend actions they should perform. I will suggest that this is because an account of how children recognises pretence contents must allow that children have some sort of mentalistic understanding of the mental states that motivate pretend behaviours. As such, we will see another way in which introducing a distinct attitude is insufficient for explaining pretence.

In section 4.5, we will consider a worry raised by Stich and Tarzia about whether the mentalistic view is also ill-placed to explain how children grasp the content of an episode of pretence, and how they figure out what pretend actions to perform. I will respond to this challenge by reflecting on the fact that pretending involves *communication*. I will suggest an important role for intentions to pretend in the recognition of pretence, and I will argue that we can explain how children figure out which pretend actions to perform by introducing desires to make things *fictional*. As such, I will conclude that the mentalistic theory should be our preferred account of how children engage in

pretence, and that having embraced this view we can reject the non-doxastic assumption when it comes to explaining how we engage in pretence.

4.1 Motivation & Recognition

One of the most striking demonstrations of how adept children are at understanding pretence is that they can recognise when others are engaging in pretend play at the same time as they begin to engage in pretence behaviour themselves. This has been described as these two abilities being *yoked* together in development: there is no development stage where children can engage in pretence yet fail to recognise that others are pretending (Leslie 2002, pp. 105-108). Indeed, many experiments on children that try to shed light on pretence rely on children being able to recognise that the experimenter is pretending. For example, in Leslie's tea party experiment, the child had to recognise that the experimenter is having a tea party.

In the case of motivation, it is easy to overlook the challenges posed by the need to recognise what pretend actions are appropriate ones to perform. For example, how does a young 15-month or 18-month-year-old child realise that to pretend a banana is a phone, one lifts the banana in question up towards one's ear?

Recent behavioural theories of pretence, such as Langland-Hassan's, have tended to explain recognition and the generation of appropriate pretend actions in terms of children being able to recognise that they, or another individual, are playing a certain sort of *game*. Stich and Tarzia offer a helpful summary of this sort of approach:

[o]bservers can understand what someone playing the pretense game is doing by noting that the person playing the game is creating a state of affairs that is similar, in salient ways, to what is going on in an appropriate depiction of an imaginary world. (Stich and Tarzia 2015, p. 7)

Mentalistic accounts, such as the one offered by Leslie, maintain that children need to possess the mental state concept PRETEND in order to be able to recognise pretence and to generate appropriate behaviours. Leslie agrees with Nichols and Stich that recognising pretence will involve coming to recognise a pretence premise, but he argues that

the only way to explain how children can recognise pretence premises is by supposing that they possess the concept PRETEND, which can allow them to recover the way the pretender is representing their action (Friedman and Leslie 2007, p. 108).

To see the motivation for this argument, it will be helpful to compare desiring and pretending. There is a stage where children can desire things, for example milk, whilst lacking a concept of desire and while being unable to recognise that others desire things.⁹² With pretending, there is no stage where one can pretend without being able to recognise that others are pretending. Leslie (1987) explains this by proposing that PRETEND is an innate mental state concept and that engaging in pretence is an early example of our mindreading capacities.⁹³ This is also supposed to explain how appropriate pretend actions are motivated. Children can recognise that they are pretending that such-and-such, where this means recognising that they are in a certain kind of mental state, and this allows them to generate appropriate behaviours.

Some philosophers and psychologists are uncomfortable with Leslie's theory because it appears to attribute a relatively high degree of conceptual sophistication to young pretenders. Some experiments seem to suggest that children have a rather confused understanding of pretence (Lillard, 1993) so we should question whether children really do possess the mental state concept PRETEND.⁹⁴ Leslie responds to these sorts of worries

⁹² Some studies suggest an ability to recognise desires arises relatively early in development, and before an ability to recognise beliefs (Repacholi and Gopnik 1997). This doesn't alter our point here, since this ability doesn't arise early enough to predate children being able to have desires.

⁹³ One could of course accept the contention that PRETEND is a mental state concept, but deny that it is an innate concept.

⁹⁴ These experiments concern a Troll named Moe who is hopping around like a kangaroo, but does not know what a kangaroo is. Children are asked whether he is pretending to be a kangaroo. At a young age children answer 'yes' and this has been taken to show at least at younger ages children do not possess a mentalistic concept of pretence. German and Leslie (2001) respond that all this demonstrates is that children don't recognise that one requires knowledge about X in order to pretend to be X.

by noting that young children do not possess concepts like MENTAL STATE or REPRESENTATION in light of their possession of the concept PRETEND. A child can possess the concept of a PHONE without possessing the concept of SOUND WAVES, or as Stich and Tarzia (2015, p. 9) put the point, the concept of COW without possessing the concept of VERTEBRATE.

Before I offer a more detailed defence of Leslie's view, it will be helpful to explain why behavioural theories struggle to explain motivation and recognition.

4.2 Broad & Narrow

Friedman and Leslie (2007) raise two related worries about behavioural theories of pretence recognition, which also bring out some worries about behavioural theories of pretence motivation. The first is that behavioural theories are too *broad*: they predict that children will mistakenly categorise many ordinary behaviours as pretend behaviours. The second is that these theories are too *narrow*: it is difficult to account for certain forms of pretence in behavioural terms.

Let's begin with the too broad objection. Thinking in terms of the idea that pretence is a game, Friedman and Leslie charge that if the behavioural theory was correct, young pretenders would make systematic errors about what sorts of behaviours indicate a pretence game is being played. To see why, we can recall that Nichols and Stich (2003, p. 53) argue that pretence involves behaving in a way that would be appropriate if *p* were the case. This means that for them, recognising pretence behaviour involves recognising that someone is behaving *in a way that would be appropriate if p were the case*. For example, in a banana-telephone scenario, the mother is recognised to be *behaving in a way that would be appropriate if that [banana] was a telephone*.

However, even this careful formulation could accurately describe a wide range of human behaviour. For example, this behavioural description would also cover cases where people merely have a false belief that *p*. Friedman and Leslie set out the worry as follows:

It is often acknowledged that actions based on the false belief that P are instances where one behaves in a way that would be appropriate if P were the case. For example, if Sally's mother mistakenly believes that a candle is an apple then she will behave in a way that would be appropriate if the candle were an apple, and perhaps try to eat it. The Behavioral theory predicts that when Sally witnesses this mistaken action, she will incorrectly consider it to be an instance of pretense. (Friedman and Leslie 2007, p. 111)

Having introduced this worry about false beliefs, they go on to note that the kinds of behaviours a child could construe as behaving in a way that would be appropriate if *p* are even more numerous than we might initially think:

[s]uppose that Sally's mother draws with a piece of charcoal, and that it strikes Sally that Mother is using the charcoal similarly to a crayon. In this case, Sally might well think, MOTHER IS BEHAVING IN A WAY THAT WOULD BE APPROPRIATE IF THE CHARCOAL WERE A CRAYON. That is, Sally will mistake her mother's behavior for pretense. This example might not be so problematic if it were unique or even rare. However, countless other examples of non-pretense behaviors will in like fashion nicely fit the behavioral description. All that is required is that the child should be able to identify a similarity between one thing and another. (Friedman and Leslie 2007, p. 111)

As such, the defender of a behavioural theory will need to find a way of explaining why children don't make systematic errors about when people are pretending thanks to their frequently classifying ordinary behaviours as pretend behaviours.

The narrow objection proceeds by noting that there are three related examples of pretence behaviour that are not easily explained by Nichols and Stich's theory and other behavioural accounts. These are *object substitution* pretence, *sound effects* pretence and *pretend speech* pretence.⁹⁵

An example of object substitution pretence would be a child pretending that a pencil is a car by pushing it along a table. An example of sound effects pretence would be where the child pushing the pencil along the table also makes 'vroom' noises while moving their pencil-car along. A pretend speech example would be where someone sits behind a teddy

⁹⁵ The pretend speech worry is developed in a later paper (Friedman *et al.* 2010), but they take it to illustrate a similar worry to these first two examples of pretence.

bear and begins to lift up its arms and talk in an exaggerated manner, making it seem as if the bear is talking. Friedman and Leslie note that these examples cannot be easily explained in terms of appropriate behaviours if *p* were the case. If the child were behaving in the ways that were appropriate if the pencil were a car:

[t]hen [the child] would hardly push it across a table or make engine noises! Handling, pushing, and making “vroom” noises are not appropriate behaviors when dealing with a real car. Instead, appropriate behaviors for dealing with a real car include opening its doors, getting inside or, if one is very young, being placed inside, sitting still, and looking out the window. (Friedman and Leslie 2007, p. 115)

This quote helps to reveal why the worries here are so potent for the behavioural account of the motivation and recognition of pretence behaviour. It is not only that the supposed behavioural description fails to capture what other people are doing; it also fails to capture what the pretender takes themselves to be doing.

If a behavioural theory is going to encompass these sorts of cases it will need to be modified. Friedman and Leslie suggest adding the addendum that sometimes when pretending children will make *X* move in ways that would be appropriate if *X* were a *Y*. If we made this move, we would need to give a disjunctive account of pretence where we have (Friedman and Leslie 2007, p. 116):

1. The child acts in a way that would be appropriate if *X* were a *Y*
- OR
2. The child makes *X* move in a way that would be appropriate if *X* were a *Y*

This means that the behavioural account of recognition will also have to be disjunctive. For example, Friedman and Leslie propose that if a child sees mother is pushing a car they'll have to determine whether (Friedman and Leslie 2007, p. 116):

1. MOTHER BEHAVES IN A (NON-SERIOUS) WAY THAT WOULD BE (NON-SERIOUSLY) APPROPRIATE IF THE PENCIL WERE A CAR
- OR
2. MOTHER IS MAKING THE PENCIL MOVE AS IF THE PENCIL WERE A CAR

However, they argue that this addendum is going to take us back to the too broad worry, and indeed will make this problem harder to deal with:

This expansion of the behavioral description, however, has the unsavoury consequence of leading the Behavioral theory to predict that children will treat as the same – that is, as “pretense” – all cases where a person makes one object move as if it were another object. What this comes down to again is simply whether the child perceives some degree of similarity between the motions of the object Mother is handling and some other motion the child knows about. Therefore, such examples will be damagingly ubiquitous. For example, ... Sally will treat Mother (seriously) drawing with charcoal the same as Mother pretending to draw with, say, a spoon (namely, as pretending the charcoal/spoon is a crayon). Or, if Mother (seriously) throws a ball, then Sally might notice that her mother is making the ball move as if it were an airplane, and so on. (Friedman and Leslie 2007, p. 116)

In other words, the worry is that this expanded behavioural theory is going to predict that children will mistake an even greater range of behaviour as pretence behaviour. This disjunctive account also still won't encompass sound effects or pretend speech pretence. As we noted previously, one example of sound effects pretence is a child making ‘vroom’ noises while pushing along a pencil-car. A problem for the behavioural view can be seen if we take (2) and try to adjust it for this example. For example, we might say that the child is making the pencil sound like a car, and we can then offer a modified (2r) and say that when mother makes a vroom noise the child recognises that MOTHER IS MAKING THE PENCIL SOUND LIKE A CAR (Friedman and Leslie 2007, p. 117).

The issue with this proposal is that this isn't an accurate behavioural description of what's going on: this suggests mother is making the *pencil* sound like a car but mother is making the vroom noise, not the pencil.⁹⁶

⁹⁶ Friedman and Leslie (2007, p. 117) suggest we could instead add a third disjunct to our behavioural theory of pretence:

3. The child produces sounds that X would produce if X were a Y

And correspondingly:

In light of these sorts of difficulties, Friedman and Leslie conclude that:

It is not obvious to us how to modify the Behavioral theory so that it allows Sally to attribute the sound effects that her mother makes to the pencil-as-car, rather than to just her mother. After all, the sound effects are Mother's behavior, not the pencil's, and Behavioral theory only allows the child to describe Mother's real behavior and not what mother is pretending as such. Mother is actually making car sounds on behalf of the pencil because she pretends the pencil is a car. But behaving 'on behalf of' an object is not something that can be made sense of behavioristically. It seems that the child cannot be a behaviorist about pretense and also get it right about sound effects in pretense. (Friedman and Leslie 2007, p. 118)

At this point then we have a strong challenge for any behavioural theory of pretence to respond to. Recent behavioural theories, such as the ones offered by Langland-Hassan and Stich and Tarzia, need to refine their explanation of recognition so that it rules out children getting systematically confused about what behaviours are to be understood as pretend behaviours, at the same time as making sense of all the different varieties of pretence. In turn, they also have to explain how appropriate pretend behaviours are motivated, since the foregoing worries about these accounts being too narrow will also apply to how children are able to recognise appropriate behaviours to perform.

Before we see how they develop their theories to meet this challenge, it will be helpful to look at why the mentalistic account is well placed to step in here and explain the recognition of pretence and other features like set-up, elaboration and effect.

3r. MOTHER PRODUCES THE SOUNDS THAT X WOULD PRODUCE IF X WERE
A Y

The worry remains the same. For recognition to work in this case, the child needs to recognise that mother pretends the pencil is a car *and* pretends of the sounds she is making that the car makes those sounds. This revised behavioural description still says mother produces the sounds in question, not to the pencil-car.

4.3 Mentalistic Pretence

As we have already noted, the two key proposals of Leslie's theory are that children possess the mental state concept PRETEND and that pretence representations are 'decoupled'. This PRETEND concept is supposed to play a key role in the recognition of pretence. To see why, recall the tea party example. In this sort of case, Leslie argues the mental state involved is a combination of an *agent, attitude and representation*. The agent is 'I', the attitude is 'PRETEND' and the representation is something like 'I am having a tea party'. We end up with the idea that this episode of pretence involves forming meta-representational beliefs like:

I PRETEND "[that] I am having a tea party"

Recognition will involve the agent of this complex mental state changing, with the attitude (PRETEND) staying the same and the representation staying largely the same. For example, the child might represent something like:

MOTHER PRETENDS "[THAT] SHE IS HAVING A TEA PARTY"

In order to recognise this, the child will attend to behaviour much as Nichols and Stich suggest, but this is used to infer the mental state that guided the behaviour in question. For object substitution, sound effects and pretend speech cases, these descriptions will have to become more complicated but will be developed in the same general form. For example, Friedman and Leslie (2007, p. 118) (making use of demonstratives) suggest that recognition of sound effects pretence might look something like:

MOTHER PRETENDS (OF) THIS PENCIL_i "IT_i IS A CAR_j" AND (OF) THAT SOUND_k "IT_k IS MADE BY THE CAR_j"

For the mentalistic theory there is no special problem raised by these kinds of pretence; the recognition of object substitution and sound effects pretence is explained in the same way as other forms of pretence. As such, the too narrow worry doesn't arise. Furthermore, since recognition depends on children possessing the mental concept PRETEND, we can also sidestep the too broad problem. Children's possession of this concept will allow them to recognise genuine instances of pretence as compared to behaviour motivated by false

beliefs etc. As we just noted, the child can recognise that pretence is something more than mere ‘as-if’ behaviour and that it stems from particular sorts of mental states.

This theory is well placed to make sense of the three features of pretence that any theory needs to explain.⁹⁷ In the case of *set-up*, the generation of the pretence premise can be explained by Leslie’s notion of ‘decoupling’. In an early paper, Leslie (1987, p. 419-420) argues that decoupling is achieved by three mechanisms. Firstly, we have the *expression raiser*. This can either take a primary representation of the external world, such as *this is a banana*, or a representation from memory, and decouple (disconnects) it from its ordinary logical connections. The quoted representation ‘this is a banana’ is then supplied with the context AGENT PRETEND by the *manipulator*. This gives a representation like ‘I PRETEND that ‘this is a banana’’. Now that this expression has been decoupled and the right context applied, it can be manipulated further thanks to a connection between the manipulator and what Leslie calls the *interpreter*. The interpreter provides information about primary representations from the external world (such as *this is a banana*), as well as from memory. In the present case, this might lead to the expression being transformed into ‘this [banana] is a telephone’ thanks to noticing some similarities between telephones and bananas in terms of their shape or from remembering a previous time they pretended a banana was a telephone. Since this secondary representation is decoupled, there is thus no threat of inferential chaos here and we can explain synchronous processing. And, since this is supposed to be a sort of representational manipulation we can engage in at will, this explains why we are free to choose what we pretend when we generate pretence premises.

In the case of *elaboration*, these manipulated representations still maintain some ordinary logical connections thanks to the interpreter, meaning that children can form regular inferences from a premise such as ‘this [banana] is a telephone’. Non-inferential elaboration and embellishment are possible because the manipulator leaves the child free

⁹⁷ Nichols and Stich (2003, pp. 47-57) note that Leslie is not always explicit about how his account explains various aspects of pretence. In what follows I am largely attempting to independently square what Leslie argues in various papers with various aspects of pretence.

to manipulate their pretence representations as they desire based on information from the interpreter.

As for *output*, motivation is supposed to be explained by saying that because children possess the PRETEND concept, they can recognise what would count as an appropriate pretend action. It is unclear precisely why this is supposed to be the case, and we will return to this point shortly when we discuss Leslie's approach to explaining recognition. This leaves the outstanding question of why children respond *emotionally* to their games of pretence. This is not an issue that Leslie directly addresses to my knowledge. I will say something that bears on this question in chapter 5, where I'll discuss and respond to a general problem that arises when we argue that beliefs with meta-representational contents lead to affective responses.

Finally, it's worth noting that this sort of distinct content approach gives us a useful explanation of clustering, since according to the mentalistic view children are able to reason about the contents of specific episodes of pretence. Weisberg (2015, p. 4) notes that the mentalistic view 'finds support from research showing that children can navigate multiple episodes of pretend play with different partners, an ability which depends on keeping track of these partners' beliefs about what the props represent in the game.' For example, Weisberg and Bloom (2009) report experiments that suggest children are able to separate out multiple pretend worlds (e.g. they realise a pretend prop can have different identities in different episodes of pretence: that a banana is not a telephone in every game of pretend).

The reader might wonder, at this point, whether this account is implausible because it attributes a capacity to form meta-representational beliefs to young children (as we saw in chapter 2, this means to form a representation of a representation, e.g. representing mother as believing X, or mother as pretending X). A common worry raised here is that children start to engage in pretence between 15 to 18 months, an age where they still fail standard false belief tests. In a standard false belief test, which is often called the 'Sally Anne' test, the child watches someone place an object into a container of some sort, such as a box or a drawer. The person who placed the object then leaves the room, and another person removes the object from the container and places it elsewhere, usually

in another container. Upon watching, this scene, the child is then asked where the person who originally placed the object will look for it when they re-enter the room. Before the age of four children tend to reply that they will check the container the object was moved to. After the age of four, they report that they will check the container the object was in originally.

This is often viewed as demonstrating that children under the age of four lack an understanding of belief, and this can be seen as evidence that they cannot form meta-representational beliefs about other person's mental states. To this empirical worry, we can offer two responses. The first comes from Leslie (1994, 2002) himself, who argues that the ability to form meta-representations containing the PRETEND concept arises before the ability to form meta-representations containing the BELIEF concept. This might be, for example because it takes additional development to come to understand the nature of false beliefs (Leslie and German 2001). A second response is to note that recent studies have suggested the ability to pass a false belief test arises much younger than previously thought, perhaps as early as 15 months (Onishi and Baillargeon 2005). These new studies rely on measuring things like looking time to establish whether young children show an understanding of false belief. As such, the claim that children can form these sorts of meta-representations at the point where they start engaging in pretence is empirically contestable, but there is no knockdown argument against granting this ability to young pretenders. As such, I will not worry further about the mere possibility of forming these sorts of beliefs, though we will touch on similar sophistication worries in relation to childhood engagement with fiction in chapter 5.

As offered, there are several ambiguities in this sketch of Leslie's theory of pretence. I will mention two. The first ambiguity concerns what it means to say that PRETEND is a mental state concept. Here are two readings:

1. PRETEND is a concept of a distinct attitude of *pretending-that*, a distinct attitude from believing or desiring
2. PRETEND is a concept of a certain form of belief, beliefs about what one is pretending

The second suggestion here is the one Leslie comes closest to directly affirming. In an article co-authored with Nichols and Stich, we are told that:

According to Leslie, in terms of boxology, there is no such thing as the ‘pretend box’, and thus no such thing as simply ‘having a pretend’. Instead, pretending is a special case of placing a representation in the ‘belief box’, where the representation says in effect, ‘someone is pretending such and such’. (Nichols *et al.* 1996, p. 56)

However, Currie (1998, pp. 39-41) suggests that the most charitable interpretation of Leslie is captured by the first suggestion: that engaging in pretence involves forming representations like ‘I pretend that p ’ and also meta-representational beliefs like ‘I believe that I PRETEND that p ’. This is for two reasons. Firstly, he notes that it’s unclear how the mental state concept PRETEND could literally be a concept of a certain kind of belief. Secondly, and perhaps more crucially, he also notes that we don’t typically think we can form beliefs at will, so it is unclear why we are supposed to be able to form meta-representational beliefs at will while we are pretending. However, if we reject this second reading and embrace the first approach, we will introduce some sort of distinct attitude into our account of pretence. The resulting view would amount to the claim that we need to supplement an account along the lines of the one offered by Nichols and Stich’s with an ability to form meta-representations and would amount to adopting a distinct attitude and distinct content view of the sorts of representations that allow us to engage in pretence.

However, I think we can develop the core of Leslie’s arguments without introducing a distinct attitude. To begin, it will be helpful to note a comment made by Weisberg. She argues that the crucial claim of the mentalistic theory is that:

Engaging in an episode of pretend play crucially requires understanding something about the mental states involved, so that one is aware of what is intended in the game. On this view, playing the banana-as-telephone game requires knowing that one’s partner (or oneself) *intends the banana to represent a telephone*. (emphasis mine) (Weisberg 2015, p. 4)

This is a weaker claim than either (1) or (2), and can be understood as amounting to the idea that children are able to engage in pretence and recognise when others are pretending because they recognise the intentions to pretend that motivate pretence behaviours. In

other words, in order to engage in or recognise pretence, we have to form beliefs like *mother is (or I am) pretending that p* where ‘pretending’ is understood as not being a form of mere behaviour or as a ‘distinct attitude’, but as relating to pretenders’ intentions to represent things.

On this reading, arguing that PRETEND is a mental state concept amounts to arguing that in order to understand pretence children have to recognise something about people’s internal mental states, and this something is related to their intentions and the way they are manipulating their primary representations, rather than something like a distinct attitude that is some sort of counterpart to belief. It has to be recognised that the reason pretenders engage in certain sorts of behaviour has something to do with their intention to pretend, and what they are pretending has something to do with the way they are representing the world internally, rather than their primary representations of the world.

That being said, we might press that the idea of an ‘expression raiser’ and more generally the notion of ‘decoupling’ amounts to the idea that we can entertain propositions in a belief-like way, thus serving to reintroduce belief-like imaginings. However, this piece of architecture enables just one part of the general process of forming a meta-representation, and Leslie leaves it open whether we can merely decouple a representation without relating it to an informational relation such as PRETEND. For Leslie, decoupling occurs whenever we represent someone else’s mental state, such as if we represent that ‘mother BELIEVES “the cat is on the mat”’. Decoupling is a general mechanism that allows us to remove representations from their ordinary input and output relations and in so doing allows us to engage in mindreading. Furthermore, even if one presses that decoupling must involve some sort of distinct attitude, it is unclear whether this will amount to introducing belief-like imaginings, rather than some thin notion of entertaining a proposition that occurs as a precursor to taking a more specific attitude towards the proposition in question.

The crucial point about decoupling is that if the mentalistic view is right, then we cannot explain how pretence behaviour is motivated or recognised solely by allowing that we have an ability to decouple representations. It also has to be specifically recognised

that you or someone else is *pretending*, where this means being in a certain sort of mental state rather than engaging in a sort of mere behaviour.

This still leaves Currie's worry, however, about why we can generate pretence premises at will when we cannot ordinarily believe things at will. For example, when I decouple a representation in an attitude report like 'mother BELIEVES "the cat is on the mat"' I can voluntarily call to mind different propositions mother might believe, but I am not free to choose which of them I take to be an accurate description of what mother believes. In pretence, on the other hand, it looks like you can simply form whatever representations you desire, and we could argue that to make sense of this we will have to introduce something like a distinct belief-like attitude.

However, we can make sense of pretence premise generation without introducing distinct attitudes by drawing on an argument from Langland-Hassan, whose counterfactual theory of pretence we first came across in section 2.4.1 and which we will discuss in more detail shortly. In explaining how we might engage in hypothetical reasoning, he offers the following example:

[s]uppose one wants to hypothetically reason about what will happen if the Cubs win the World Series this year (a hypothetical that promises to remain hypothetical). Call the proposition that the Cubs win the World Series this year 'c'. The desire to know what will happen if c will be enough to cause one to access whatever general beliefs one has about teams that win the World Series. A few likely come to mind: the team jumps for joy ('j'), their fans shed tears of elation ('e'), they take part in a ticker-tape parade in their home city ('t'), shirts are printed up ('s'), and so on. Having brought these generalizations to mind, and believing the Cubs to be the sort of thing that falls under those generalizations (i.e., a baseball team), one then infers that if c then probably j and e and t and s, etc., and forms the corresponding beliefs. There is no need during all of this to put c itself in either the belief or desire "boxes"—or any "box" at all (hence, no need to "quarantine" c). Turning again to the issue of pretence, if one wishes to pretend that the Cubs have just won the World Series, the inferred (and now believed) conditionals just mentioned will be sufficient to guide a sequence of pretend behavior. (Langland-Hassan 2012, p. 167)

We can embrace part of his insight here, and allow that the voluntary generation of pretence premises has something to do with a general ability to think about what it would be like if such-and-such was the case by reflecting on our beliefs, rather than on an ability to enter into special belief-like states.

Going back to the mentalistic theory, we can allow that during pretence children can consult their beliefs about the thing they want to pretend, and so we can argue that all they need to do to generate a pretence premise is ask themselves what would happen if p were the case. Combined with a suitable desire to pretend this will be enough to get an episode of pretence started. This then leaves the issue of why at this point we should maintain that the relevant notion of ‘pretend’ here is mentalistic rather than behavioural, an issue we will consider in the next two sections of this chapter.

That being said, we cannot conclusively rule out the possibility that some sort of notion of belief-like imagining is at least sometimes involved in pretence. It may well be that sometimes episodes of pretence get started because children entertain or suppose a proposition in a belief-like way. Equally, however, there could be alternative mental states that kick off pretence and this means that belief-like states are *unnecessary* for explaining pretence. The initial intention to pretend could arise from a desire to do something funny (Carruthers 2006, footnote 18, proposes that this is why the child in Gould’s example pretends to be a dead cat), or to play with a sibling. It might also stem from forming a mental image of a scary dragon and then beginning to act as if this dragon were real. The important point for present purposes is that all of these sorts of states are *insufficient* for explaining how we are able to engage in pretence bearing in mind the worries we noted in the previous section about behavioural views of pretence. In order to pretend or recognise that someone else is pretending, you need to form a belief with distinct content like ‘I PRETEND “ p ”’ or ‘mother PRETEND “ p ”’ where this involves some sort of recognition of your (or mother’s) intention to represent p . We will need further beliefs about what would be the case if p and perhaps a desire to pretend that p , but there is no need here for belief-like imaginings. In this way, once we’ve adopted this sort of distinct content view it becomes *unnecessary* to introduce a distinct belief-like attitude into our account of pretence.

Before we further defend this view of pretence motivation and recognition, we should firstly consider some behavioural responses to the worries that we raised about recognition and motivation. The cost of embracing an approach along the lines proposed by Leslie remains that it attributes a relatively high level of conceptual sophistication to

young children and commits us to the view that they can form meta-representations from a young age. As such, we should take care to examine whether a suitably modified behavioural theory can do the explanatory work here to see if we can avoid making these sorts of empirical commitments.

To develop my argument for why a modified behavioural theory will still struggle to explain all aspects of children's engagement in pretence, we can introduce a distinction between recognising *that* someone is pretending, and recognising *what* they are pretending: the *specific content* of their pretence episode. In terms of motivation, we need to explain not only what motivates someone to pretend, but also what allows them to figure out what constitutes an appropriate pretend action.

This issue has often been relatively ignored, but has started to receive a good deal of attention in recent years, with Langland-Hassan attempting to make sense of this issue on his single attitude behavioural view. Stich and Tarzia also offer a distinct attitude behavioural view that seeks to make sense of this issue, and raise a challenge for mentalistic theories when it comes to explaining how children figure out appropriate pretend actions and recognising what others are pretending.

4.4 Defending Behavioural Recognition & Motivation

To begin this discussion, it will be helpful to quote Langland-Hassan's account of how a tea party proceeds (in his example 'P' refers to what he calls a 'perceptual attitude', 'B' to a belief and 'D' to a desire.⁹⁸):

P1: You say, "Let's have a tea party!" and start setting out dishes and cups. You do all of this with a familiar cluster of mannerisms (e.g., knowing looks and smiles, exaggerated movements and intonation, stopping actions short of normal goal points).

⁹⁸ He tells the reader that if they are sceptical of the existence of perceptual attitudes they can read these as being further beliefs. I take no stand on this issue here as it will not affect our assessment of this theory.

B1: (inferred from P1) You are starting a game where we act in ways that would be appropriate if we were at a tea party, even if we're not at one.

D1: I play this game, too.

P2: You are acting as if you are pouring tea out of the teapot and into the cups.

B2: (from D1 and P2) I should act as if you poured tea into the cups.

B3: (from B2 and stored generalizations) If you had poured tea into both cups, they would both now be full.

B4: (D1 causes this to be inferred from B3): I should act as if both cups are full.

P3: You put down the bottle and say “watch this!”; you turn the green cup upside down and then put it back on the table, right side up.

B5: (background beliefs): When cups containing liquid are turned upside down, the liquid spills out. When full cups are not moved, they remain full.

B6: (inferred from P3, B4, and B5): If you had poured tea into both cups and overturned the green one, the green one would now be empty and the other one full.

B7: (inferred from B6, due to D1) I should act like the green cup is empty and the other one is full.

P4: You say, “Show me which cup is empty and which is full.”

D1—an abiding desire to play the game—then leads the child to consult B7 in giving her answer: she points to the green cup to indicate that it is “empty,” and then to the other to indicate it is “full.” (Langland-Hassan 2012, pp. 165-166)

In this way, he rejects a role for any belief-like states in pretence and also for meta-representational beliefs. As such, he would reject my suggestion in the previous section that recognising pretence or engaging in pretence requires one to have, or to recognise, an intention to pretend, where this is understood as entailing that children do not understand pretence as mere behaviour.⁹⁹

In this sketch of a tea party, Langland-Hassan is suggesting a role for two different sorts of beliefs in an explanation of how children can engage in pretence. The first sort of belief is captured by B3 ('If you had poured tea into both cups, they would both now be full.'). We are told this is inferred from a desire to play the game (D1) along with beliefs about what's going on in the game that lead to one asking oneself an internal question (represented by B2). This is the one he readily signposts by telling us that propositionally imagining involves forming counterfactual beliefs based on stored generalisations, as we noted in section 2.4.1. The other is captured by B1 and B2, which have contents of the form 'I should act as if p'. These sorts of beliefs are what we can call self-regarding ones about how we ought to act or behave (or possibly about how others are behaving, in the case of the recognition of pretence). However, how the child can generate these sorts of beliefs is not explained merely by saying that they can ask themselves questions and call upon stored generalisations. Instead, Langland-Hassan implies that these are formed based on a mixture of the pretender's counterfactual beliefs

⁹⁹ We can bring out here that the behavioural view is also mentalistic in some sense. We can contrast here what we can call a *mentalistic mentalistic* view and a *mentalistic behavioural* view. According to the former, children possess a concept of PRETEND and this concept is a mentalistic concept. According to the latter, children still have a concept of PRETEND but this is a concept of a certain sort of behaviour. The disagreement here is not about whether children possess any sort of PRETEND concept but about what sort of concept PRETEND is. The mentalistic view says that pretence has some mental component and has to be understood as involving mental representations in order for children to engage in pretence and recognise others are pretending. The behavioural view says that although pretence is made possible by the possession of mental states, the recognition and the generation of pretence behaviours only requires children to understand PRETEND as a behavioural concept.

and a desire to play the pretend game in question. This raises the question of whether Langland-Hassan can offer a satisfactory explanation of how children recognise what sort of behaviour counts as appropriate during a pretence episode on the basis of these sorts of mental states.

He argues that when a child forms a belief like 'I should act as if you poured tea into the cups' the 'act as if' here amounts to saying the child has a belief that they should act in a way that makes their actions appear saliently tea-has-been-poured-into-cups like. To make sense of pretence recognition he notes the importance of *manner cues*. For example, when we engage in pretence, we often perform exaggerated mannerisms: licking our lips going 'mmm' when we pretend to eat a cake, or winking after saying 'it's for you!' when answering a banana-phone. It would be natural for a behavioural theorist to integrate these manner cues into their account of pretence recognition. Indeed, Friedman and Leslie (2007, pp. 111-113) try to develop this sort of account on Nichols and Stich's behalf, though find it wanting since these sorts of behaviours are *prima facie* not the sorts of behaviours that would be appropriate if *p* were the case (one doesn't usually lick one's lips in an exaggerated manner when eating a cake; a real monster (presumably) wouldn't shout out 'I'm a monster!' while making exaggerated roaring noises).

Langland-Hassan attempts to integrate manner cues into behavioural theory in a subtle way. Langland-Hassan (2012, pp. 174-177) agrees with Friedman and Leslie that manner cues aren't usually behaviours that would be appropriate if *p* were the case, but suggests that children can recognise that these manner cues often are not appropriate behaviours if *p* were the case. From this observation, Langland-Hassan proposes that pretence recognition is explained by children being able to recognise that the pretender's behaviour is making some X saliently Y-like, whilst also recognising various manner cues that make the pretender's behaviour less Y-like. The combination of the two allows children to recognise that someone is pretending. The manner cues don't make the pretend behaviour more Y-like, but they help to focus attention on the fact someone is pretending.

Having made this proposal, he offers a relatively simple way of evading the too broad and too narrow objections. Children don't mistake other kinds of behaviour as pretence behaviour because other kinds of behaviour are not accompanied by manner cues which

serve to render those behaviours less Y-like. In a false belief case, if someone mistakenly believes that a piece of plastic fruit is real, they won't attempt to eat it whilst making exaggerated 'mmm' noises. Instead, they might take a bite and spit it out, or stop short of taking a bite upon realising the fruit is plastic. For the too narrow objection, his response is a bit more complex. Roughly he argues that the behaviours Friedman and Leslie mention (e.g. running a pencil along a table, making vroom noises) count as a way of making some X saliently Y-like. The child recognises that mother, in pushing the pencil along saying 'vroom', is trying to make the pencil saliently car-like. Perfect resemblance isn't achieved here (as already noted, mother is making the noise, not the pencil), but according to Langland-Hassan 'perfect' resemblance isn't necessary for pretending:

Turning to the matter of pretense recognition, in recognizing the pencil/car pretense, the child recognizes that the father is trying to make the pencil saliently car-like (as above, manner cues both direct her attention to the pencil, and allow her to recognize that he is starting a pretense game with respect to the pencil). One way to make the pencil car-like is to cause it to move forward and backward around the table, since a salient feature of cars is that they move forward and backward. Of course, it is not a salient feature of cars that they have hands moving them, but the hand's involvement is necessary to bring about some other salient resemblance. Another salient feature of cars is that their motions are accompanied by Vroom sounds. In making Vroom sounds while the pencil moves around the table, the father makes the pencil car-like in the respect that its movements are accompanied by Vroom sounds. Of course, the Vroom sounds of cars are made by engines, not mouths. *Perfect resemblance* is not achieved—but, fortunately, pretense does not require it. The point of pretense is *to go some distance toward making some x saliently y-like*. (emphases in original) (Langland-Hassan 2012, p. 177)

He sums up his approach by telling us that:

[a] person can be reliably recognized as pretending that p by recognizing that she is acting in some salient ways that would be appropriate if p, while offering some of a familiar cluster of manner cues, some of which involve acting as if not-p and draw attention to the subject matter of the pretense. (Langland-Hassan 2012, p. 175)

Roughly then, Langland-Hassan's view amounts to the idea that recognising pretend actions involves recognising some X has been made saliently Y-like, but is in some other ways not saliently-Y like thanks to the presence of manner cues. However, we should question whether Langland-Hassan's account of children's understanding of pretence can

explain how children recognise *what* someone is pretending in addition to *that* they are pretending. For example, when a child sees mother pretending to sip tea she needs to recognise that mother is pretending but also as more specifically pretending that she is at a tea party. In turn, this also raises the question of how the child figures out what pretend actions they should perform once they've called upon stored generalisations and formed a desire to pretend that *p*.

We have been told that both these issues relate to the pretender making *some x saliently y-like*, but this is puzzling both from the side of the pretender and from the recogniser. In the first instance, how does the pretender know what will count as making X saliently Y-like? In the second instance, how does the recogniser realise whether X has been made sufficiently Y-like? We have been told that perfect resemblance isn't required, but we have no sense of what is needed instead. In this way, it looks like there is a troublesome vagueness at the heart of Langland-Hassan's theory.

Before we consider how we might respond to these issues, we should first develop Stich and Tarzia's related approach to making sense of recognition and motivation. Stich and Tarzia (2015) develop their refined behavioural account along similar lines to Langland-Hassan. The main difference with their account is that they return an imagination box to the picture, since they think that we need this box to account for where the child 'tokens' pretence premises.¹⁰⁰ That being said, they do not argue that recognition and motivation cannot be explained by Langland-Hassan's theory and refrain from criticising his account in detail, simply pointing out in a footnote (Stich and Tarzia, footnote 10) that he will face issues explaining things like pretence deficits in children with autism. These worries need not concern us since both Langland-Hassan and Stich and Tarzia end up giving a similar account of pretence recognition and the generation of appropriate behaviours.

¹⁰⁰ Langland-Hassan would respond that the premise in fact isn't tokened anywhere since it will not be entertained save for as part of a counterfactual belief.

Stich and Tarzia agree with Langland-Hassan that we should accept Friedman and Leslie's criticisms of Nichols and Stich's behavioural view, and drop the idea that recognition involves children merely recognising behaviour that would be appropriate if p were the case. They replace this with the suggestion that children recognise what someone is pretending by recognising that their external actions imitate the contents of their imagination box by being 'similar, in salient ways' to what is being represented in the imagination box (Stich and Tarzia, p. 7).

For example, this means that if a child sees someone running around barking, they recognise that this is similar to the events represented in the pretender's imagination box when the pretence premise is something along the lines of 'I am a dog'. Appropriate pretend behaviours arise because children can form desires to create a series of events that is similar to the ones in their imagination box after they place a pretence premise into it. In this way, they can make sense of the too narrow worry by allowing that instances of pretence like pretend speech do indeed imitate the contents of the imagination box in some respects. To deal with the too broad worry, they rely on manner cues in much the same way as Langland-Hassan:

[w]e propose that manner cues play an important role in alerting the observer that the pretence game is being played. The smiles, knowing looks, winks and nods, exaggerated gestures, unusual tone of voice and stopping short of normal goals that have loomed large in the pretence literature since Piaget, are signals to children (and to adults as well) that the pretence game is being played. (Stich and Tarzia 2015, p. 7)

However, in this passage, it is only argued that these manner cues 'alert the observer that the pretence game is being played', which leaves it open how the pretender recognises which *specific* pretence game is being played. As such, Stich and Tarzia need to respond to similar questions to Langland-Hassan in order to explain how children recognise what someone else is pretending and what would count as an appropriate pretend behaviour when they engage in pretence. In trying to make sense of these issues, both theorists also need to avoid slipping a mentalistic notion of pretending into their accounts, a risk Langland-Hassan (2012, p. 175-176) helpfully highlights.

In the case of recognising the specific content of an episode of pretence, both Langland-Hassan and Stich and Tarzia maintain that pretence content is recognised thanks

to manner cues. Manner cues thus need to do double duty in highlighting both *that* someone is pretending and *what* they are pretending. For example, Langland-Hassan (2012, p. 175) argues that manner cues ‘draw attention to the subject matter of the pretense’. To illustrate the general idea here, both thinkers consider the case of someone pretending to be sleepy who makes an exaggerated yawning noise and declares that they’re very sleepy, which they argue indicates the content of this pretence to the child.

The problem here is threefold. The first is that not all cases of pretence recognition will involve obvious manner cues that draw attention to the subject matter of the relevant episode of pretence. The second is that it is unclear how manner cues are able to allow children to recognise the content of episodes of pretence if they understand pretence as mere behaviour, since pretend actions are not merely conventional. Finally, we still haven’t heard enough about how the child recognises what pretend actions are appropriate ones to perform in their own games of pretence.

In relation to the first worry, Stich and Tarzia argue that in cases where the premise isn’t obvious:

[t]he child’s cognitive system begins providing the [imagination box] with pretense premises and noting similarities between features of the world described in the [imagination box] and salient features of the behavior of the person (e.g. a parent) who has initiated the pretense. (Stich and Tarzia 2015, p. 7)

This risks introducing a sort of regress into their theory of recognition: Stich and Tarzia don’t make it clear how children work out what sorts of premises they need to ‘try’ here or how they recognise when they have stumbled upon the right one. They accept Weinberg and Meskin’s (2006b, p. 182) proposal that pretence premises are placed into the imagination box by an ‘inputter’, but they do not explain how this mechanism knows which premises to input when we recognise that someone is engaging in pretence behaviour. Perhaps more crucially, they also do not tell us how our cognitive mechanisms recognise when the appropriate premise has been placed in the imagination box and in turn how children recognise that they should stop trying out new premises.

In relation to the second worry, part of the issue here is that children are capable of being creative when they pretend to be an elephant, or a dog etc., and can pretend that this is so in novel and imaginative ways. However, when they see someone pretending in

a novel way, it is unclear how mere manner cues could lead the child to recognise the content of the pretence in question, if they have no understanding of the intentions of the pretender. This point leads us on to the third worry, since when children engage in novel pretence behaviour of their own, it is unclear how they would recognise that their behaviour is nonetheless ‘similar, in salient ways’ to the contents of their imagination box.

Similar worries also arise when we reflect on Langland-Hassan’s view, since he doesn’t explain what allows children to recognise what counts as an appropriate pretend action to perform, or what makes them realise when some X has been made sufficiently saliently Y-like.¹⁰¹ Going back to our original question for his theory, this means that he doesn’t tell us enough about what allows the child to form beliefs like ‘I should behave as if *p*’ or ‘mother is behaving as if *p*’.

The answer to these sorts of worries presumably will come from getting clearer on the notions of ‘saliency’ and ‘similarity’ that these behavioural theories rely on. However, neither of these theorists say much about how we should understand these terms. For example, in a footnote we are told by Stich and Tarzia that:

[i]n the pretense game, the player is aware of the representation of events of a certain sort in a component of her mind that she has access to. She need not be aware that that component of the mind is her [imagination box], or that the events represented

¹⁰¹ This point also helps to rule out a further proposal we might introduce to defend behavioural theories of pretence. We could argue that there is a basic of kind of action – ‘pretendings’ – which can be recognised as such. If this were true, recognition could be explained by saying we just need to recognise someone is performing an action of this kind. The difficulty with this approach is it seems that ‘pretendings’ are best thought of as representational actions, as noted by Friedman (2013, pp. 193-194). He suggests, for example, that we could draw a parallel with drawing, where arguably to recognise what someone has drawn you need to recognise the drawer intended to represent something. In much the same way, we could argue recognition of pretence involves recognising that someone intends to represent something and on the basis of this to infer what they are attempting to represent. In this way, a ‘type of action’ based account of pretence may well still end up being a mentalistic account.

are being imagined. She need not even have the concept of imagination. Rather, she has the demonstrative belief that events of a certain sort are occurring there (in her [imagination box]) and she desires to create a saliently similar sequence of events. (Stich and Tarzia 2015, footnote 11)

It is notable here that this sort explanation of pretence gives us no hint of how the child knows what would count as a ‘saliently similar sequence of events’. For example, in response to the too broad objection, Stich and Tarzia argue that:

[t]he child is not looking for behavior that would be appropriate if p were the case. Rather, she is looking for behavior creating a sequence of events that is saliently similar to the events represented in the [imagination box], when p is used as a pretense premise. And, as we have noted, the similarity can be far from perfect, and the manner cues, while typically diminishing *similarity*, will also often heighten *saliency*. We conclude that the over-extension problem is easily handled on the current account. (emphases mine) (Stich and Tarzia 2015, p. 8)

The reason why behavioural theorists have strayed away from attempting to define these sorts of terms is presumably because they wish to avoid the sorts of issues Friedman and Leslie raised initially. If a broad sketch of similarity and saliency were to be offered, this would risk the theory in question being subject to the too broad objection, and if we give a highly specific account of similarity and saliency, it will fail to readily encompass the wide variety of potential pretence behaviours.

As such, Stich and Tarzia and Langland-Hassan both don’t do enough to explain how children work out the specific content of episodes of pretence or to address how children know which pretend actions to perform.¹⁰² For these theories to seem plausible,

¹⁰² Ferreira (2014, pp. 104-108) raises a somewhat similar concern: whether behavioural theorists can develop a notion of ‘making’ that is robustly behavioural and doesn’t just amount to saying that we have a mentalistic notion of pretending. He considers and rejects various proposals, such as that making amounts to transforming, and argues that the most plausible notion of making involved in pretence is something like a notion of *make-believing X has Y-like properties*. However, as he notes, this amounts to attributing something like a mentalistic concept of PRETEND to the pretender. I am sympathetic to his worries in relation to Langland-Hassan’s view, but it strikes me that Stich and Tarzia can respond here by arguing that ‘making’ merely involves recognising the appropriate degree of similarity and saliency has been achieved in relation to a representation in the pretence

we effectively have to take it as given that children *just do* possess a capacity to recognise that an appropriate degrees of saliency and similarity have been reached when it comes to performing their own pretend actions or observing someone else's pretend actions. In turn, we have to accept that they possess these abilities even though they lack a deep understanding of what it is to pretend. On the other hand, if we embrace the mentalistic theory, this provides us with additional resources to explain how children figure out how to perform appropriate pretend actions and recognise the content of pretence episodes, since we can allow that children have some basic understanding of the fact pretenders intend to pretend and that pretence therefore has a mental component.

However, Stich and Tarzia's have developed a recent critique of Leslie's view of pretence motivation and recognition, with their chief objection being that it is overly mysterious how the PRETEND concept is supposed to enable pretence recognition and motivation. To conclude this chapter, I will attempt to respond to some of their concerns and will further elaborate on why a mentalistic view is better placed than the behavioural theory to explain pretence motivation and recognition.

4.5 How much can PRETEND explain?

Stick and Tarzia note that the PRETEND concept is supposed to do a lot of theoretical work for Leslie. For example, they suggest that for Leslie the PRETEND concept:

[plays] a role in “generating and interpreting” pretense behavior; pretend play “issues from” PRETEND representations; both pretending and recognizing pretense “spring from” the concept; the concept “empowers” the child to recognize that someone else is pretending that P. (Stich and Tarzia 2015, p. 9)

box. If Stich and Tarzia are right, this can be done without any awareness of the sorts of mental states involved in pretending. This returns us to our present worry about how a child can recognise this level of similarity has been reached if they possess a merely behavioural understanding of pretence, which I take to be the hardest challenge for behavioural theories to meet.

Their concern relates to *how* the PRETEND concept can do this. How does possession of this concept motivate pretend behaviours, and enable children to recognise pretence in others? Stich and Tarzia develop their challenge by noting that:

[according to Leslie] possession of the PRETEND concept means that children are able to engage in simple episodes of pretense, to recognize them and to share them with others. Our question is how? How does possession of the PRETEND concept enable children to do these remarkable things? How, for example, does possessing the PRETEND concept enable Sally to understand what Mommy is doing when she holds the banana up to her face and says “Hello, Daddy”? And how does it enable Sally to know what to do when Mommy hands her the banana? (Stich and Tarzia 2015, p. 10)

They are then able to raise a parallel worry to mine about recognising content and generating appropriate behaviour, pointing out that although Leslie and collaborators tell us that:

[f]or Sally to pretend that the banana is a telephone simply requires representing the agent of pretend as self . . . or as we if pretense is shared, and using the resulting meta-representation, I (WE) PRETEND THIS BANANA ‘IT IS A TELEPHONE’ in part as a high-level command to the action planning system. (Stich and Tarzia 2015, p. 10)

There is a gap here because Leslie and collaborators do not:

[t]ell us how the action planning system manages to figure out that the right thing to do is to hold the banana up to one’s ear and mouth and talk into it, rather than pushing it around the table saying “zoom, zoom” or pointing it at one’s partner and saying “bang, bang.” (Stich and Tarzia 2015, p. 10)

As such it looks like both the mentalistic and behavioural theories will struggle to explain the recognition of pretend contents and the motivation of appropriate pretend behaviours. If Stich and Tarzia are right, Leslie and collaborators simply fail to tell us how pretend contents are recognised, and if I’m right, then behavioural theories of pretence still need to spell out the notions of similarity and saliency which they introduce to make sense of this issue. We thus seem to be left at something of an impasse.

However, I think that we can defend a mentalistic distinct content view against these sorts of worries. Stich and Tarzia are right to point out that there are some explanatory

gaps in Leslie's account, but I think by reflecting on some of his claims about pretence we will be able to plug these gaps and render the mentalistic account of pretence more plausible than a behavioural one.

Recall that Leslie's theory tells us that pretence is underpinned by beliefs that look something like I PRETEND "I am a dog". What we need is an explanation of how this sort of meta-representation leads to the actions that are performed when a child pretends to be a dog, and an explanation of how children can recognise that someone else has decoupled this sort of pretence premise. Our question then is how do children go from forming these sorts of complex beliefs to running around the garden barking, and when they see someone else running around the garden barking, how do they figure out the decoupled representation which is part of the pretender's mental representation of their behaviour?

To begin, we should consider whether Stich and Tarzia are right that Leslie has 'ignored' this issue. They note in a footnote (Stich and Tarzia 2015, footnote 19) that he has 'briefly' discussed recognising content in an article co-authored with Happé on pretence and autism (Leslie and Happé, 1989). It will be helpful to look at what Leslie says in this article to see how we might go about defending a mentalistic approach to making sense of how children recognise the content of pretence episodes and what allows them to perform appropriate pretend actions when engaging in pretence.

Leslie and Happé (1989, pp. 209-211) suggest that pretence recognition and motivation can be explained by introducing the notion of *communication*. When mother is pretending to be a dog, or a child is pretending to be a dog, they are trying to *communicate* something about their internal representations by manipulating their body and the world around them. As such, our initial questions now become: how does the child know how to communicate the content of their internal pretence representations via external behaviours? And, how does the child know how to recognise what someone else is communicating something about their internal pretend representations?

Stich and Tarzia are right that Leslie and collaborators are often vague in response to these sorts of questions, but they do give us some information. For example, in the case of recognition, they note the relevance of manner cues for helping to focus attention

on what is being pretended, in a similar way to Langland-Hassan and Stich and Tarzia. Leslie and Happé (1989, p. 210) give the example of putting an empty cup near one's mouth and licking one's lips, and suggest that this sort of exaggerated display can 'trigger the meta-representational mechanisms, which may allow the actor's (pretender's) intention to be inferred'. Friedman further develops the thought that pretence involves communication and makes a related point about the importance of intentions when it comes to recognising pretence contents, telling us that:

[r]ecognition of pretend play ... requires people to recognize that certain actions and objects are intended to serve as representations, and to infer what it is the pretender intends to represent. (Friedman 2013, p. 193)

One might worry here that this is no more informative than the behavioural view, since manner cues will still underspecify the contents of at least novel episodes of pretence. However, the key point here is that you cannot recognise what someone is pretending merely by observing their behaviour with no recognition of the intentions behind it. When a child observes mother pretending, the child has to find a way of looking past mother's odd behaviour (such as talking into a telephone) to infer the mental state that is guiding her behaviour. In this way, if we accept a mentalistic view, then manner cues are of more help for figuring out pretend contents, since children will have some understanding that, in exhibiting manner cues, pretenders are communicating something about their intentions to represent things.

That being said, one might wonder how it is possible for children to recognise intentions to pretend and perhaps more crucially how these sorts of considerations help to explain the motivation of pretend actions. To make sense of the recognition question, we can note that Leslie argues that there is evidence that from a young age children can recognise *goal directed* activity, of which he thinks pretence is an example, in that it is something like an intentional mental activity:

Deliberate—that is, goal directed—external, physical actions require the representation of their goal. Likewise, deliberate, internal, mental actions also require a representation of their goal. Deliberately undertaking the external action of tying laces requires representing the goal of that action as one of tying laces and therefore requires having the concept, TIE LACES. Likewise, deliberately undertaking the action

of pretending that P requires representing the goal of that action as pretending that P. And for this reason, the child who deliberately pretends uses the concept PRE-TEND-THAT in his goal representation. Naturally, this is also the concept required for representing the mental state of another person who is pretending-that. (Leslie 2002, p. 112)

One might still worry, however, that this focus on internal representations and intentions still doesn't do enough to explain how children generate appropriate pretence behaviours and recognise what others are pretending, and amounts to merely labelling the difficulties. We can, however, offer a helpful response to the recognition question by reflecting on how we recognise the content of mental states more generally, and the issue of motivation by reflecting on the sorts of desires that might motivate pretend actions.

In the case of recognition, consider how we recognise the contents of mental states in other contexts. For example, let's suppose we try to construct a theory of how people recognise the specific things that other people believe. We already have general overarching theories that seek to explain our mindreading capacities, such as the theory-theory and simulation theory, but it is hard to give specific advice for recognising a particular belief. Recognising exactly what a subject believes is a complicated task to perform, and yet we seem to be able to at least sometimes succeed in figuring out the specific contents of people's beliefs. In light of this, it is hard to spell out precisely how a theory of mind or a simulation process allows us to figure out what mental states we should ascribe to others.

Since mentalistic views argue that understanding someone is pretending involves a degree of mindreading, it is no surprise that we don't have an easy answer to this sort of question about recognition of pretence contents. Figuring out what someone is intending to represent with their pretence is a complex task that will involve reflecting on the intentions of the pretender, the way their actions relate to other actions, manner cues and so on. That being said, recall that the issue with the behavioural approach is that it gives us no easy way of explaining how children figure out the content of pretence based on the limited resources provided by manner cues and noting similarities between real and pretend actions. If we embrace a mentalistic view, we can at least go beyond this by allowing that children can recognise something about the mental states that motivate

pretence behaviour. In so doing, this means that we cannot give a simple account of how children figure out the specific contents of an episode of pretence, but this should not be taken as demonstrating a failing of the mentalistic view as opposed to a natural consequence of it.

As for motivation, we can note that believing that ‘I PRETEND “*p*”’ will not motivate any actions whatsoever unless it is combined with some further mental states. The standard Humean view of action holds that actions should be explained in terms of a *cognitive* and *conative* component. Ordinarily, this will be a belief and a desire: my walking outside with an umbrella can be explained in terms of a desire to not get wet and a belief that it is raining (or perhaps a belief that it will rain shortly, or that it is likely to rain etc.). When we ask why it is that a child runs around the garden barking like a dog, there will thus be a role for desire to play in responding to this question.

Since we have accepted that children can form beliefs related to what is going on in specific episodes of pretence, we can also propose here that children can form desires that make explicit reference to the fact that they are pretending. Indeed, Leslie suggests the goal of pretence is ‘decoupled’:

An obvious hypothesis about pretend actions is that the goal representation is decoupled. If so, this might explain why the movement undertaken typically does not carry through to the point in the real world that it would normally if it were generated by a regular "coupled" goal representation. For example, if I have a normal goal of drinking from a cup, I will lift the cup all the way to my lips ensuring close contact between cup and lip (for obvious reasons). If I pretend to drink from the cup, typically I will stop short of contact. I may even only outline the action of lifting and drinking in a highly truncated manner of gesturing the cup toward my lips. (Leslie 2002, p. 11)

Bearing this sort of consideration in mind, I think we can make a tentative case for the proposal that pretenders are motivated to perform pretend actions based on forming desires to make such-and-such *fictional*.¹⁰³

¹⁰³ I argue for the introduction of these sorts of desires at greater length in my MPhil dissertation (Davis 2015, Ch. 2).

We can motivate this proposal by noting that when engaging in pretence children do not aim to make it literally true that, for example, they are a dog or that they are having a tea party. Furthermore, it doesn't look like children mistakenly think they are trying to make it the case that they are a dog when they engage in this sort of pretence. Drawing on Walton's theory of fictional truth, we can argue that children aim to make these things *fictional*: when a child pretends to be a dog, they aim to make it fictionally true that they are a dog (*I am a dog*).

I do not have space to fully elaborate on this idea, but we can note two benefits of introducing these sorts of desires. Firstly, desires to make fictional are helpful for explaining how novel forms of pretence are generated, since desiring to make something fictional does not have to restrict you to wanting to pretend in obvious or predictable ways. This is because there are various ways of behaving that can satisfy a given desire to make something fictional. For example, a desire to make it fictional that you are dog could be satisfied by barking, woofing, running around on all fours, chasing a bone, or a combination of various behaviours. In this way, forming these sorts of desires can explain why children perform such a wide variety of actions when engaging in pretence.

Secondly, when developing an account of the motivation behind pretend actions, we need to explain how children recognise how manipulating their bodies and the external world can communicate the contents of their pretence. This point is emphasised by O'Brien (2005, p. 60) in relation to belief, when she notes that the representations that guide pretend play can only play a motivational role when they are 'draped' over the external world. We can make sense of this by arguing that having a desire to make something fictional entails having some understanding of how to represent your internal pretend representations via your external actions.

However, at this point, it could be argued that behavioural theories can also introduce desires to make things fictional in order to respond to my challenge about the vagueness of the notions of saliency and similarity. We could argue that children can form desires to make things fictional, but do not have any understanding of them, and that this is still enough to explain how they recognise what pretend behaviours are

appropriate ones to perform during a given episode of pretence.¹⁰⁴ However, although this might help with issues related to motivation, this will not salvage the behavioural account of recognition. Even having introduced these sorts of desires, it will remain unclear how children can work out what others are trying to make fictional unless we attribute some sort of mentalistic understanding of pretence to them that allows them to reflect on the intentions that are guiding episodes of pretence.

This leaves us with an important final question, which is whether the account I've offered is too sophisticated as an account of how young children engage in pretence. To respond to this, it will be helpful to note some comments made by Friedman and Leslie about this sort of worry:

To be clear, we do not believe that children's possession of this concept PRETEND implies that they know much about this or other mental states. In particular, it does not imply that they theorize about mental representation or that they theorize that pretence is an 'internal, subjective, mentally depictive state', as some have supposed. . . Nor does it require that children can report that pretenders 'are thinking' and what they are 'thinking about' while pretending. (Friedman and Leslie 2007, p. 120)

We haven't done much here to complicate this picture. All we need to add to this is that possession of a mentalistic PRETEND concept also means that children can form desires to make things fictional, which allow them to realise their communicative intentions when they engage in pretence. This still doesn't entail that children possess complicated concepts like that of a MENTAL REPRESENTATION. This mentalistic approach also remains more plausible as an account of pretence recognition and motivation as compared to the behavioural approach, since attributing this sort of sophistication to children gives us some tools for making sense of the mysteries of how they are able to engage in and recognise pretence, whereas the behavioural view is forced to rely on vague notions such as 'similarity' and 'salience' to make sense of these abilities.

¹⁰⁴ Indeed, Schellenberg (2013) makes this sort of amendment to behavioural accounts, suggesting children could have a non-conceptual understanding of these sorts of desires.

Conclusion

We saw in this chapter that issues related to pretence motivation and recognition pose a crucial challenge to behavioural theories of pretence, both those that rely on belief-like imaginings and those that do not. In section 4.4, we saw that behavioural theorists have recently developed more elaborate accounts of pretence in order to explain these issues. However, these refined behavioural theories are more plausible when it comes to explaining how children recognise *that* someone is engaging in pretence as opposed to *what* someone is pretending. In turn, this means that these accounts struggle to explain how children figure out which pretend actions they ought to perform when they engage in pretence. As such, I argued that we should prefer a mentalistic view of pretence that introduces meta-representational beliefs, since this sort of view will allow that children have some awareness of pretenders' intentions to represent things with their pretend actions. In turn, this helps to explain how they recognise the content of episodes of pretence and what counts as an appropriate pretend action.

In section 4.5, we saw that Stich and Tarzia argue that the mentalistic view of pretence also suffers from difficulties when it comes to explaining how children recognise pretence contents and figure out what counts as an appropriate pretend action. I introduced some responses to these worries by focusing on the fact that pretence behaviour is communicative, by highlighting that pretence is a goal-directed activity, and by introducing desires to make things fictional. Having done this, we saw that the mentalistic view remains more plausible than the behavioural view as an account of pretence.

As such, we have come to see that belief-like imaginings are *insufficient* for explaining how we are able to engage in pretence. In addition, they are also *unnecessary*, since once we introduce meta-representational beliefs of the form 'I PRETEND "*p*"' we have no need to introduce a belief-like attitude in order to explain our engagement in pretence. We can thus justify rejecting the non-doxastic assumption in the context of pretence. We will now turn to consider whether we should similarly reject this assumption in the context of our engagement with fiction.

Chapter 5: Imagination & Fiction

Introduction

In this chapter, we will consider whether we need to introduce belief-like imaginings to explain how we are able to engage with fictions. I will argue that belief-like imaginings are not *necessary* for explaining our engagement with fiction. This is not to say that they can never be involved in our engagement with fiction; nor that there is no plausible argument for associating belief-like imaginings and fiction. I will instead suggest that we can also explain our cognitive engagement with fiction by introducing beliefs with distinct contents, namely beliefs implicitly subject to a *fictional operator*.

Many philosophers would likely accept that when reading about the exploits of Sherlock Holmes in one of Conan Doyle's novels, we can form beliefs with fictional content such as 'I believe that Holmes is a detective [in the fiction]' (Kripke 2013, Tullmann 2016). If we borrow notation from Walton (1973) and Evans (1981), the idea here is that we can form beliefs like 'I believe *Holmes is a detective*'. This is because beliefs subject to a fictional operator play an important role in the philosophy of language when it comes to explaining why we can use seemingly empty names in ordinary language. Indeed, even defenders of imagining-based views of how we engage with fiction, such as Walton (1990, Ch. 10), allow that we can form these sorts of beliefs. As such, my argument here can be understood as considering whether these beliefs, which are already frequently introduced, can do all the work in explaining how we engage with fiction, or whether we also need to introduce a distinct belief-like attitude. As such, although the view I wish to defend might seem counter-intuitive, it may, in fact, be more parsimonious than distinct attitude views if the defender of these views also accepts that we can form beliefs like 'I believe Holmes is a detective [in the fiction]'.

That being said, one might want a more positive reason for why we should try and develop this sort of distinct content account of our engagement with fiction. One reason is that in embracing this sort of view, we avoid the need to introduce somewhat mysterious involuntary imaginative counterparts in the context of our engagement with

fiction, and then having to explain which states have these sorts of counterparts (belief) and which do not (e.g. the emotions if we reject the idea of quasi emotions). Furthermore, Tullmann (2015, pp. 36-38) argues that having resisted the introduction of imaginative counterparts into our account of our engagement with fiction, we are well placed to explain various important issues that we will not have space to discuss in this thesis, such as the puzzle of imaginative resistance (why we seem to resist ‘imagining’ certain propositions in fictions, such as that ‘killing babies is a good thing’) and sympathy for the devil (why we sometimes sympathise with immoral protagonists like Tony Soprano and Walter White).

To defend the distinct content view, I will begin in section 5.1 by setting out the similarities and differences between fiction and pretence, which will bring out the involuntary nature of the cognitive attitude involved in our engagement fiction. I will argue that this means we cannot offer exactly the same explanation of how we are able to engage with fiction and in pretence. This will conclude our challenge to the uniformity assumption that we introduced in chapter 2.

In section 5.2, I will consider whether there are aspects of our cognitive engagement with fiction that can only be explained by introducing belief-like states. I will suggest that there are not by considering worries related to *sophistication*, *ontology* and *metafictional statements* made in works of fiction. Our discussion of sophistication will also serve to remind us that distinct attitude views struggle to make sense of clustering, which will suggest that we may have some reason to think that belief-like imaginings are insufficient for explaining how we are able to engage with fictions.

In section 5.3, we will discuss whether belief-like imaginings are needed to explain the distinction between fiction and non-fiction. I will argue that they are not and that a better view of this distinction can be found by endorsing Friend’s proposal that fiction and non-fiction are genres, with neither genre involving a distinctive attitude that isn’t involved in our engagement with the other.

In section 5.4, we will consider a debate between Matravers and Friend about whether research related to how we construct situation models when engaging with narratives suggests that an account of our engagement with fiction requires belief-like

imaginings. I will side with Matravers that it does not, and will offer some reasons for why constructing situation models won't necessarily require belief-like imaginings.

In section 5.5, we will consider whether beliefs about what is fictional can explain why we can become immersed in fictions and why we sometimes respond emotionally to them. I take it this is the strongest objection to an approach to explaining our engagement with fiction that dispenses with belief-like imaginings when explaining our engagement with fiction. I will set out several ways that a defender of a distinct content view can explain why we respond emotionally to fictions, and will question why these sorts of beliefs cannot explain immersion.

I will thus conclude that there is no need to associate belief-like imaginings with our engagement with fiction, though I will also tentatively suggest that other forms of imagining might be involved, such as objectual and sympathetic imaginings. We will thus reject the non-doxastic assumption in the context of our engagement with fiction since introducing a distinct belief-like attitude is *unnecessary* for explaining our engagement with fiction.

5.1 Belief-Like Imaginings & Fiction

Before we can establish whether we should prefer an account of fiction that relies on beliefs with distinct contents to one that relies on belief-like imaginings, we will first need to say a bit more about the relevant sort of counterpart to belief that might be involved here. I suggested in chapter 1 that there are two sorts of counterparts: voluntary and involuntary counterparts, and it looks like the counterpart involved in our engagement with fiction is best thought of as an involuntary counterpart.

The view that there is a connection between imagination in a propositional belief-like sense and fiction has been defended by Currie (1990), Lamarque and Olsen (1994), Sutrop (2002), Livingston (2005), Davies (2007) and Stock (2011b) among others. Defenders of the consensus view tend to embrace something similar to Walton's notion of fictional truth, and argue that what is fictionally true in a given work of fiction is what we are prescribed to imagine in a belief-like way by the work. However, in recent years

the view that there is a distinctive link between fiction and belief-like imaginings has been called into question.¹⁰⁵ Friend (2008) argues that the imagination is often equally involved in our engagement with works of non-fiction; Matravers (2013) goes a step further and argues that engaging with both fiction and non-fiction does not necessarily have anything to do with the imagination; and Tullmann (2016) adopts a ‘single attitude’ view of fiction which denies that any imaginative counterparts whatsoever are involved in our engagement with fiction.¹⁰⁶ I think we can maintain the insight that fictions ask us to take propositions to be fictionally true, but can argue that instead of prescribing us to imagine these propositions in a belief-like way, we are simply prescribed to believe these things are fictionally true.

Most philosophers who are explicit about the nature of the counterpart to belief involved in our engagement with fiction appear to take it to be one in the voluntary sense. This is because they embrace the uniformity assumption and argue that this belief-like attitude is the same as the one that is involved in pretence and perhaps also suppositional reasoning, and both of these activities must involve some sort of attitude that is subject to the will (e.g. Nichols 2004, Currie and Ravenscroft 2002, Goldman 2006b). However, when we look more closely at the role belief-like imaginings are thought to be playing in our engagement with fiction, it becomes apparent that it is ambiguous whether we're discussing a role for a counterpart in this voluntary sense or whether we are discussing an involuntary counterpart. This ambiguity arises because there is an important distinction between engaging with fiction and *creating* a work of fiction. To bring out this ambiguity, it will be helpful to contrast fiction and pretence.

¹⁰⁵ Not all of these thinkers use the phrase ‘belief-like imagination’. All agree, however, that there is a special attitude associated with fiction which is some sort of counterpart to belief.

¹⁰⁶ Tullmann’s argument is broader than mine. I only wish to deny that belief-like imaginings play a role in our engagement with fiction. Tullmann denies that any sort of imaginative counterpart is involved in our engagement with fiction, thus also rejecting a role for quasi emotions, i-desires and so on.

I suggested in chapter 3 that there are three things a theory of pretence needs to explain: set-up, elaboration and output. These three headings parallel – to some extent – what an account of our engagement with fiction needs to explain.

1. Set-Up

The first aspect of set-up, the generation of the initial pretence premise, looks like it has something in common with *fiction-making*, the act of creating a work of fiction. If a belief-like counterpart is involved in fiction-making, it will have to be a voluntary counterpart since the fiction-maker is able to freely generate the contents of their fiction. When J.K. Rowling decided that she wanted to make it fictionally true that Harry Potter is a wizard, this is presumably something she was able to make fictionally true at will by, for example, writing certain words on a computer screen or in a notebook, which would eventually make their way onto a printed page.¹⁰⁷ I will not say much about the attitude (or attitudes) that allows us to create works of fiction in what follows. I suspect that fiction making may well have important features in common with the generation of pretence scenarios and it might also involve *creative* or *constructive* imaginative capacities that cannot be readily reduced to mere belief-like imaginings.¹⁰⁸ When it comes to engaging with a work of fiction, generation is less obviously involved since what we are supposed to take as fictionally true is already given to us on the page or on screen etc.

Before we move on to consider how fiction relates to elaboration, we should note that the issue of recognition which we placed under the heading of set-up (how children are able to recognise that someone else is pretending) is also important when it comes to explaining our engagement with fiction. When I pick up a book, I need some way of

¹⁰⁷ More contentious is whether an author can make things fictionally true without including this truth explicitly in their fictional work. For example, there has been some controversy in the Harry Potter fandom about whether Dumbledore is gay since this is not prescribed as being fictionally true in the novels but J.K. Rowling stated this was true of Dumbledore in an interview.

¹⁰⁸ Van Leeuwen (2013) offers an interesting discussion of constructive imagination and some reasons for why it might not neatly reduce to belief-like imagining.

determining whether it is fiction or non-fiction, and the same is true if I turn on the TV and see that a movie is playing. We might be able to do this by relying on the description on the sleeve of the book, or by checking TV guide, but there is a question about whether we can also determine whether a work is fictional or non-fictional in the absence of this sort of information. My suspicion is that we cannot, since there are no obvious surface features of fictions that serve to distinguish them from non-fictions. In practical terms, this is unlikely to be a frequent concern since we are usually aware of whether we are engaging with a work of fiction or non-fiction. This also helps to bring out a difference between pretence and fiction: while engaging in pretence we perform seemingly odd actions, often accompanied by manner cues, with this helping to draw attention to the pretend status of our actions. With a work of fiction, the sentences contained within can just be ordinary examples of sentences in a given language.

2. Elaboration

Though generation is not essentially involved in our engagement with fiction, elaboration is frequently involved, at least in the sense of inferential elaboration. While we are engaging with works of fiction, we draw various belief-like inferences and these inferences are essential for rendering fictional worlds coherent.¹⁰⁹ For example, few works of fiction take the time specify that their human characters have internal organs, but we nonetheless take this as given.¹¹⁰ We infer that Sherlock Holmes has a liver in the *Sherlock*

¹⁰⁹ Following Walton (1990, Ch.4), the task of figuring out how these inferences are drawn can be described as trying to find the ‘principles of generation’ that govern our engagement with fiction. There is little widespread agreement about how we should set out these principles, but I assume that the way we set out these principles will be equally compatible with either a distinct attitude or distinct content view of how we engage with fiction. Friend (2017) offers the interesting proposal that we assume everything that is true in reality is also true in a given fiction and only depart from this assumption when the fiction gives us reason to do so.

¹¹⁰ At least in most cases: we won’t infer this if a fiction explicitly tells us its human characters do not have organs. This deviation, however, can be explained by the fact that works of fiction are free to depart from at least some truths about the actual world.

Holmes novels, for example, and also that he requires food for energy. Sometimes, and indeed probably for the most part, these sorts of inferences will not involve forming occurrent representations. If a realistic novel tells us a character has travelled from London to New York, we may not consciously represent that they have travelled by plane, but if the next page tells us they teleported between the two cities we will be surprised.¹¹¹

On the other hand, non-inferential elaboration does not seem to be associated with engaging with fiction, though it can be involved in fiction making. We might perform non-inferential elaborations to try to figure out why certain fictional events happened, or how they might have transpired differently, but works of fiction do not ordinarily prescribe us to take propositions to be fictionally true when they are not explicitly stated in the work or cannot be inferred from the explicitly stated fictional truths of the work.

This is because we have less control over the fictional events described in a work of fiction as compared to the fictional events that make up an episode of pretence. When engaging in pretence, children sometimes make valid inferences and follow instructions about what to pretend, but they are also free to break the rules of their pretence. For example, a child can declare that all the tea at their tea party has evaporated and this can be made fictionally true because of their declaration. On the other hand, when engaging with a novel, I cannot simply declare that whatever I imagine is fictionally true. As noted in chapter 1, it is not within the power of a mere reader of *Sherlock Holmes* to make it fictionally true that Holmes is a wizard.¹¹²

A slight complication about the role of non-inferential elaboration arises when thinking about interactive fictions. When engaging with an interactive fiction, such as a

¹¹¹ Walton (1990, pp. 16-18) offers an interesting discussion of whether there can be non-occurrent imaginings, and argues that there can be. Regardless of whether we endorse a distinct content or distinct attitude view, we need to allow that not everything we take to be fictionally true will be occurrently represented by us.

¹¹² As we noted in chapter 1, they might be able to create their own fictional world in which Sherlock Holmes is a wizard. The point here is they cannot make him a wizard in the world of the novels.

videogame, players often have the power to make at least some things fictionally true based on their own volition, even if these things aren't strictly entailed by previous events in the fictional world. The extent of this power will vary greatly depending on the game in question. Some games allow for a great deal of player control, such as sandbox games like *Minecraft*, while others are more linear, such as *Uncharted 4*. Even in more linear games, the player still has the power to, for example, make it fictionally true that their avatar stood in a certain spot, climbed a wall in a certain way, picked up a certain power-up, and so on.¹¹³

We could argue that the reason this sort of non-inferential elaboration is possible when engaging with interactive fictions is because the gap between fiction-maker and consumer is not clearly demarcated in these sorts of fictions. If we return to the distinction that Walton makes between *work worlds* and *game worlds*, it might look like the reason videogames allow for this sort of non-inferential elaboration is because players can sometimes directly affect what is fictionally true in work worlds. Tavinor (2005, 2009) defend this view, suggesting that videogames render Walton's distinction ambiguous. Meskin and Robson (2012) respond that the work world and game world distinction is still applicable to videogames, and that players in fact cannot ordinarily directly affect the work world.¹¹⁴ Instead, they argue that we should think of a player as being similar to a *performer* in a play. The player-as-performer has a great deal of control over what happens in their playing, but this control is not absolute and they cannot directly affect the work world, only the world of their specific performance. To see how this idea works in practice, consider the final boss fight in an adventure game such as *Uncharted 4*. It

¹¹³ One might also think this means that generation plays a role in our engagement with interactive fiction. However, generation is supposed to refer to the initial generation of a premise in a game of pretence, and it's not obvious that someone engaging with an interactive fiction can generate this sort of initial premise. For example, with a videogame the initial premises will be set by the development team, as will much of the subsequent story in narrative-based videogames.

¹¹⁴ One exception is perhaps when playing massively multiplayer online games where there is a single fictional world every player interacts with at the same time.

might be true in the work world that your avatar defeated the final boss, but the player cannot make it true in the work world that the boss was defeated in precisely the manner they directed their avatar. These specific details are only fictionally true in the context of your specific performance

I am inclined to side with Meskin and Robson here, since the videogame player is more limited in what they can make fictionally true as compared to the fiction-maker or someone engaging in pretence, since they are still constrained by the rules set by the developer. In *Crash Bandicoot*, I can choose whether Crash moves to the right or the left, or whether he jumps over the hole or into the hole, but I cannot choose to make it fictionally true that he is a thylacine rather than a bandicoot. In a game of pretence, these sorts of limits will only be present if explicit rules are agreed upon, and even then, these rules can be transgressed if the pretenders decide to reject them or alter them. With a videogame, I can only completely change the rules by either hacking the game to change its code (thus arguably simply creating a new game) or by having the fiction-maker perform the change. We do not have the space to consider these intricacies in any further detail, but it's worth bearing in mind that an explanation of the mental goings on when we engage with interactive fiction may not perfectly match an explanation of the mental goings on when we engage with non-interactive fiction.¹¹⁵

¹¹⁵ That being said, I doubt that this means we need to introduce belief-like imaginings to explain how we engage with interactive fictions if we accept that they are not involved in our engagement with non-interactive fictions. It is unclear why we would maintain that just because someone makes the decision to move Crash Bandicoot from point A to point B this means that they imagine in a belief-like way that he is at point B, while if they see a movie character move from point A to point B they do not imagine that he is at point B but instead just believe this to be fictionally true. I expect that if there is a difference here it will concern the role of desire in our engagement with interactive fictions and the opportunities for fictional action that this opens up.

3. Output

Output is also relevant to our engagement with fiction. In the case of affect, we appear to exhibit emotional (or emotion-like) responses to works of fiction. Action and motivation are less obviously related to how we engage with non-interactive fiction since we cannot perform ‘pretend’ actions when engaging with non-interactive fictions. One could argue that our tears at the death of Anna Karenina constitute a sort of action, or that covering our eyes during a gory horror film constitutes an action, but these are not the same sorts of actions that are involved in pretence. Covering one’s eyes is an action in the real world to block our certain images, whereas the child running around the garden like a dog is trying to make something *fictional* with their real-world actions.¹¹⁶

There are thus both similarities and differences between what an account of pretence needs to explain and what an account of our engagement with fiction needs to explain. Perhaps the key difference is that because the generation of fictional truths is only related to fiction-making, it’s not so obvious that the sort of mental attitude involved in engaging with fiction is under the control of the will. The relevant state looks to be a largely automatic response to what we read on the page or see on screen etc. There is an important distinction between the fiction-maker, who presumably is able to create works of fiction thanks to some sort of voluntary attitude, and the individual engaging with a work of fiction. This distinction is not usually present in the pretence case, since often the child who is pretending is also in some sense the creator of their pretence.¹¹⁷ As Sainsbury puts a similar point:

¹¹⁶ A follower of Walton might argue that covering one’s eyes is an event with the ‘game world’ and so is a fictional action in some sense. It still seems to me that there is an important difference here in that covering one’s eyes doesn’t constitute a separate fictional action in the way a child raising their arm can constitute fictionally raising their trunk.

¹¹⁷ Another philosopher who emphasises the importance of associating something like an involuntary attitude with fiction is Cooke (2014, p. 324), who argues that: ‘the act of fictively imagining x is not identical to the act of non-fictively imagining x. This is either because the propositional attitudes ... associated with the two are not the same or because

Like belief, but unlike pretence, make-believe is often involuntary. To open a novel with a normally receptive mind is to start make-believing. Likewise, to engage in conversation with a normally receptive mind is to start believing. (Sainsbury 2010, p. 12)

As such, if we need to introduce a counterpart to belief to explain our engagement with fiction, it will be an *involuntary* counterpart rather than a voluntary one. In this way, we have now developed our challenge to the uniformity assumption.

To defend the proposal that an involuntary counterpart to belief is involved in our engagement with fiction, we will need to follow my third principle and argue that either: a) there are constraints on forming beliefs which prevent beliefs subject to a fictional operator from doing all the explanatory work when it comes to explaining our engagement with fiction, or b) that there are puzzling aspects of our engagement with fiction that cannot be explained by mere beliefs. As such, we will need to consider if there is reason to think either that we cannot form beliefs about what is true in fiction, or that there are aspects of our engagement with fiction that these beliefs cannot explain. This is because, as we noted in my second principle, we cannot justify introducing this counterpart solely by noting that the relevant attitude here is formed in response to engaging with something we recognise as being fictional.

That being said, as noted in chapter 2, my distinction between voluntary and involuntary counterparts does not entail that if there is a counterpart to belief involved in fiction, it necessarily has to be a distinct counterpart from one that might be involved in hypothetical reasoning or pretence. It might be that sometimes belief-like imaginings are under the control of the will and sometimes they are not: they could be an example

the imaginings have different content.’ His first suggestion here is broadly in line with my starting proposal in chapter 2: we can argue that there is one sort of imaginative attitude associated with entertaining or supposing, and another associated with fiction (and perhaps also pretence). The second would be to introduce a *distinct attitude distinct content* view of propositional imaginings, where there is one distinct attitude here but some of them have special contents, of the form ‘I imagine that p [in the fiction]’. As already noted, this doesn’t strike me as having any obvious explanatory benefit compared to adopting a distinct content view.

of a counterpart that can either be voluntary or involuntary depending on context. In section 1.5, the issue with making this suggestion in the case of imagining seeing was that the formation of mental images cannot be what makes things fictionally true about what we see in paintings. In the belief case, we do not have this sort of *prima facie* reason to rule out the proposal that a voluntary counterpart to belief can sometimes behave in an involuntary manner when engaging with fiction.¹¹⁸

However, while this is a possible approach, we will need positive reasons for extending the functional role of belief-like imaginings in this way and for associating them with fiction. This will still involve showing that constraints on beliefs are violated when we engage with fiction, or that puzzles arise if we don't introduce belief-like imaginings into our account. The difficulties for the uniformity assumption this discussion has revealed suggest that we should resist moving too quickly from establishing a voluntary belief-like attitude is involved in pretence (or can be associated with a notion like 'supposing') to arguing that this same attitude is involved in our engagement with fiction.

Having set out in a bit more detail the sort of counterpart that might be involved here, we can now turn to consider whether constraints on forming beliefs are present when thinking about our engagement with fiction. I will consider three worries; worries about conceptual sophistication, an ontological worry and a worry about the contents of fiction.

5.2 Cognitive Engagement

In this section, I will consider three arguments for resisting the idea that beliefs with distinct content can explain our engagement with fiction. These worries will concern for

¹¹⁸ For example, Weinberg and Meskin (2006b, p. 196) mark a distinction between *streaming* and *punctate* inputs to make sense of why belief-like imaginings are sometimes subject to the will and sometimes they are not. We could argue that engaging with fiction involves belief-like imaginings, but that fictions only provide streaming inputs.

the most part constraints on believing, but will also include aspects that are better understood as puzzles. The first issue is whether these beliefs are too sophisticated for young children to form, the second issue is whether ontological and linguistic issues point towards a belief-like state being involved in our engagement with fiction, and the final issue is whether there are some sentences found in works of fiction that we cannot believe to be true in fiction.

5.2.1 Sophistication

Let's begin with the sophistication worry: whether young children can form beliefs about what is true in fiction. We noted in passing that from a young age, children don't appear to mistake fictional characters for real individuals (Harris 2000, pp. 60-65). In light of this, children are able to engage with fiction from a young age. However, one could question whether we can explain this ability in terms of them forming beliefs about what is *fictional*, on the grounds that young children presumably lack this sort of adult concept. If this concern has merit, it at least shows that children cannot engage with fiction by forming beliefs about what is fictional, and this will also be problematic when thinking about adults if we seek a unified account of how people engage with fiction. On the other hand, it seems less conceptually onerous for children to simply imagine that *Harry Potter is a wizard* or that *Peppa is a pig*.

One way to motivate this worry would be to accept what Bermúdez and Cahen (2015) call the *conceptual constraint*.¹¹⁹ They set out this constraint as follows:

¹¹⁹ This relates to a debate about whether there can be *nonconceptual* mental content. Bermúdez and Cahen suggest that whether there can be nonconceptual content turns on how one thinks we ought to make sense of something like the following two claims:

1. In specifying what a thinker believes, what a perceiver perceives or what a speaker is saying by uttering a certain sentence in a particular context one has to be as faithful as possible to how that thinker, perceiver or speaker apprehends the world.
2. How a thinker, perceiver or speaker apprehends the world in having beliefs about it, perceiving it or speaking about it is a function of the concepts he possesses.

Specifications of the content of a sentence or propositional attitude should only employ concepts possessed by the utterer or thinker.

If this is right, this would give us reason to resist saying that young children engage with fiction by forming beliefs about what is fictional, since it is unlikely they possess this adult concept. In the case of PRETEND, we were able to say that this is a concept that was part of a theory of mind, but it is not so clear how we could justify introducing a *fiction* concept. That said, since I am allowing children can form meta-representational beliefs about pretence, this means that we shouldn't *prima facie* rule out the possibility of children forming what appear to be something like meta-representational beliefs about fiction (we can describe them in this way since we believe that such-and-such is the case *according to fiction*, which is perhaps in some ways similar to believing that such-and-such is the case *according to mother*). Indeed, as I will now go on to argue, it is important at the very least that children recognise that some of the things they engage with are fictional, since otherwise both the distinct attitude and distinct content views will struggle to explain how children engage with fiction.

To begin, it will be helpful to think about how distinct attitude theories are supposed to avoid sophistication concerns. Presumably this is because, as noted above, we might think children can form imaginings such as 'Peppa is a pig' from a young age. Thinking in terms of Nichols and Stich's theory, this representation is placed in the imagination box and can be placed in this box from a young age. But how exactly does the child recognise that this proposition should be placed in the imagination box? In the case of pretence recognition, Nichols and Stich defer to recognising certain sorts of behaviour, and Stich and Tarzia add to this the importance of manner cues. But there are no similar features of fiction that children could notice to trigger putting content into their imagination box, since even if we think producing a fiction involves pretending in some sense, there is no behaviour that children could observe to recognise this. As such, the child has to in some way figure out that the fiction in question is not be literally believed before they can imagine its contents in a belief-like way.

One way to make sense of this is to argue that children are somehow able to (at least sometimes) recognise when they are engaging with content that is 'made up'. From a

relatively young age, they can recognise that Peppa isn't a real, existing, pig even if they don't explicitly recognise Peppa is a *fictional* pig. However, once we allow that this sort of recognition, this does not have to entail that children then go on to form belief-like imaginings about Peppa. Instead, it can mean that they form beliefs about her that are different in some sense from their ordinary beliefs, thanks to their recognising that Peppa doesn't actually exist.

Furthermore, we noted in chapter 3 that the representations involved in our engagement with fiction exhibit clustering: we do not represent that *Harry Potter is a wizard simpliciter* but that he is a wizard relative to the world of the Harry Potter novels. Research suggests that young children are also able to recognise this sort of clustering. Skolnick and Bloom (2006) performed an experiment to investigate whether five-year-old children could distinguish between things that are true in the world of SpongeBob Squarepants and things that are true in the world of Batman. They found that children were indeed able to separate out these two sets of fictional truths, e.g. recognising that Batman does not believe SpongeBob exists.

This suggests that children are able to *compartmentalise* their representations about fictional worlds, a notion we will say something more about when we discuss discourse processing in section 5.4. Children are in some sense aware not only that fictional entities do not exist, but also that things are only true of them relative to particular works of fiction. This shows an impressive degree of sophistication in how young children understand fictional worlds, since it suggests that they understand their attitudes towards fictions – be they imaginings or beliefs about the fiction – only relate to some fictional worlds and not others. If we say that children have belief-like imaginings about fictions, this does nothing to directly explain why they are able to distinguish between what is true in one fictional world and false in another fictional world. As we noted in chapter 3, for philosophers like Nichols (2004), the belief-like imaginings involved in our engagement with fiction are not supposed to have distinct contents and are meant to be isomorphic to ordinary beliefs.

Returning to our initial examples, imagining that *Peppa is a pig* does not explain how children are able to keep this representation separate from an imagining that *Harry*

Potter is a wizard. Indeed, we saw in section 4.3 that Weisberg (2015, p. 4) argues that the fact that children can compartmentalise in this way when engaging in pretence offers *prima facie* evidence for a mentalistic view. In a similar way, we can argue that these sorts of experiments show that we don't have to rule out the possibility children form something along the lines of beliefs about what is true in specific fictions, even if they will perhaps have to be represented internally via somewhat different concepts. As such, issues related to sophistication appear to show some benefits of adopting a distinct content view of how we engage with fiction, and reveal at least one way in which a belief-like attitude based approach is potentially *insufficient* for explaining our engagement with fiction.

That being said, this debate relates difficult empirical issues about children's understanding of fiction, and it is hard to offer an easy answer to these sorts of questions. The important point to note here is that much like with pretence, saying children take a distinct belief-like attitude towards fictional contents does not make sense of all the mysteries related to how they are able to engage with fiction. This remarkable ability to engage with fiction is not readily explained simply by supposing that they have a capacity to form belief-like imaginings, since this leaves open the more basic question of how they know they should form these sorts of imaginings rather than straightforward beliefs.

If we endorse a distinct attitude view, we thus still have no explanation of why children are able to recognise at least some of the rules and conventions of fiction from a young age. As such, regardless of whether we think a distinct attitude is involved in children's engagement with fiction, we still might need to attribute a surprising amount of sophistication to them since they show a surprisingly nuanced understanding of the nature of fictions and fictional worlds from a young age.

5.2.2 Ontology

The second worry about whether constraints on forming beliefs suggest we need to introduce belief-like imaginings into our account of how we engage with fictions arises when thinking about the ontology of fictional characters. The notion of 'fictional characters' can be understood in a broad sense as relating to failed scientific posits (such

as Vulcan), fictional events, fictional places, and perhaps other seemingly non-existent things (Friend 2007, p. 142). Philosophers of language have argued that it is potentially troublesome for us to talk about fictional characters in ordinary language, making statements like ‘Holmes lives at 221B Baker Street’. A common view of how names function in language says that they are referring expressions and this raises the question of what ‘Holmes’ refers to in this sort of sentence, and also what ‘221B Baker Street’ refers to.

There are two main approaches we can take in response to this question about reference. Firstly, we can take a realist approach and argue that Holmes and 221B Baker Street exist in some sense – perhaps as abstract objects (van Inwagen 1977) or possibilia (Lewis 1983) – and that when we use these names, we refer to these abstracta or possibilia. Secondly, we can take an anti-realist approach and argue that Holmes does not exist, and that we are doing something like pretending to refer to someone or something when we use an empty name like ‘Holmes’ in a sentence (Walton 1990, Ch.10).

Both the realist and anti-realist proposals raise questions for a distinct content view of how we engage with fiction. If the realist approach is right, this might entail that we have straightforward beliefs about fictional characters, rather than beliefs about what they’re up to in fiction. If the anti-realist approach is right, one might worry that pretending to refer necessarily involves a distinct attitude, and that it is only after imagining in a belief-like way that ‘Holmes exists’ that we can go on to form beliefs like ‘Holmes is a detective [in the fiction]’.

This ontological debate is nuanced and hard to settle. This issue becomes rapidly more complex when we consider the sheer variety of sentences that appear to refer to fictional characters. For example, as well as utterances like ‘Holmes is a detective’, we can make:

1. Cross-fictional utterances like ‘Holmes is a better detective than Shaggy and Scooby Doo’
2. Utterances that appear to ascribe real-world properties to fictional characters like ‘Bart Simpson is a pop culture icon’

3. Negative existential claims like ‘Holmes does not exist’.

Regardless of whether one takes a realist or anti-realist position, it is hard to explain how these seemingly empty names function in all these contexts. In particular, negative existential claims prove tricky. If we adopt a realist view such as the abstract object view, this claim about Holmes seems to be false, since it asserts that something which exists does not exist. If we take an anti-realist view where we say sentences about fictional characters should be understood as subject to something like a fictional operator, then a negative existential claim about Holmes would amount to saying, ‘Holmes does not exist [in the fiction]’, yet this seems false, since according to the fiction Holmes does indeed exist.

To streamline discussion, it is worth noting that many philosophers who defend abstract object views of fictional characters argue that we do often attach an implicit *in the fiction* operator to many of our beliefs about fictional characters. This is because abstract objects cannot possess the properties that works of fiction ascribe to fictional characters: to be a wizard, Harry Potter has to be a concrete individual of some sort. Defenders of these sorts of mixed views (e.g. Thomasson 1999) argue that when we say something like ‘Harry Potter is a wizard’, we are expressing a claim about what is fictional: we are saying ‘Harry Potter is a wizard [in the fiction]’, since an abstract object cannot be a wizard. If we say, ‘Harry Potter is a pop culture icon’, this is instead supposed to be a straightforwardly true claim about an abstract object. Since the former sort of representations are more relevant for engaging with works of fiction, adopting a realist view does not entail that we should abandon a distinct content view unless we have reason to think that pretending to refer involves a distinct attitude.

At this point then, there are two issues we need to consider. The first is whether the sort of pretence involved in referring to fictional characters (or at least making internal statements about them) necessarily involves a distinct attitude. The second is whether it is problematic to argue that something we know to be non-existent is the intentional object of one of our beliefs. We will begin with this second concern.

The best response to make to the concern about whether fictional characters can be the intentional objects of our beliefs, is to note that disallowing this entails adopting a restrictive view of reference, and it is not obvious why we should commit ourselves to this sort of view. For example, it seems uncontroversial to allow that we can have beliefs such as ‘I believe that according to John, Santa Claus exists’ and presumably this means that we can also have beliefs such as ‘I believe that according to folklore, Santa Claus exists’.¹²⁰

This leaves open the possibility, however, that we cannot refer to a fictional character without doing something like pretending or make-believing that they exist, and that pretending and make-believing necessarily involve a distinct belief-like attitude. If this is right, then when I pretend to refer to Santa Claus (such as when I say to a young child ‘make sure to leave out some milk and cookies for Santa’), this means that I must have some sort of belief-like imaginings with propositional content related to Santa. For example, perhaps I can only make this utterance if I imagine in a belief-like way that ‘Santa Claus exists’.

However, it looks like we can nonetheless refer to something non-existent without first imagining it to exist. To see why, we can return to the Santa Claus case. When I knowingly using a name that has no referent – such as ‘make sure to leave out some milk and cookies for Santa’ – this may well involve a sort of pretence, but we can argue that I am able to pretend to refer in this context because I believe Santa Claus exists according to the child in question. In this way, we do not have to accept that imagining in a belief-like way that Santa exists is a precursor to being able to pretend to refer to him.

As such, worries issues concerning the ontological status of fictional characters are unlikely to settle the debate about what sort of cognitive attitude is involved in our engagement with fiction. Puzzles about the ontological status of fictional characters and our practice of using seemingly empty names in ordinary discourse do not go away if we introduce a notion of belief-like imagining, and they do not reveal any constraints on

¹²⁰ Thanks to Stacie Friend for this example.

belief that give us a reason to reject the possibility that beliefs subject to a fictional operator can explain our engagement with fiction.¹²¹

5.2.3 Fictions Are Weird

At this point, one might object that although it looks like a distinct content view can explain how we engage with works of fiction, it will in fact struggle to explain how we understand some of the utterances found in works of fiction. Picciuto (2015, p. 70) points out that sometimes fictions contain odd passages, such as the following from *The Lion the Witch and the Wardrobe*:

‘Let’s go home,’ said Susan. And then, though nobody said it out loud, everyone suddenly realized the same fact that Edmund had whispered to Peter at the end of the last chapter. They were lost.

Picciuto suggests that it is strange to say we believe that *in the fiction* these characters realised what Peter had whispered at the end of the last chapter since this would mean these characters are aware they live in a merely fictional world.

To respond to this worry, we should first widen the scope of this objection since it captures one of many aspects of what we might call the *Fictions Are Weird* problem. Fictions often experiment with storytelling devices and this can raise tricky questions about how we should understand their fictional worlds and what they prescribe us to take as fictionally true. For example, in the comic *Bomb Queen*, a character called Editor Girl has the power to alter the content of speech bubbles (Cook 2012, p. 177). Ordinarily speech bubbles in comics are mere genre tropes that are not supposed to be viewed as part of the fictional world, yet here we have to explain whether they do indeed exist as a tangible part of this particular fictional world.

¹²¹ There are many further issues we could have discussed here, such as whether fictional names are understood to refer *de re* or *de dicto*. I have chosen to leave the discussion at this point because since I find it unlikely we will settle the debate about the cognitive attitudes involved in fiction with further discussion of ontological and linguistic issues.

However, the fact that fictions are weird does not have to be seen as ruling out a distinct content view. Picciuto is highlighting a question about what we ought to do when we are presented with *inconsistent* or *incoherent* aspects of a fiction. Walton (Ch. 4, pp. 174-182) worries about some of these issues, calling them *silly questions*. For example, in *Othello*, should we take it as fictionally true that Othello can speak in iambic pentameter despite the fact he is supposed to be a somewhat brutish individual? These sorts of questions about fictional truths arise for any account of the attitude we take towards fictions, and they do not tell in favour of a belief-like imagining based view. The notion of imagining associated with fiction is supposed to track fictional truths, and saying that we imagine that ‘in the last chapter ...’ is no less puzzling as compared to believing that *in the fiction* metafictional claims are true.

One response here is to simply declare that such things are indeterminate in the fiction. It is neither fictionally true nor fictionally false that they recalled what had been said at the end of the last chapter. If one finds this unsatisfying, Matravers (2013) offers several further ways to respond to these sorts of questions:

Faced with an incoherence a reader can do one of four things. He or she could take the story to be a misreport; that is, take the narrator to be mistaken or lying. If this were the case the reader would reject the claim as false; both false in the narrative and, if the events were being reported as actual, false per se (I shall call this ‘the rejection strategy’). The reader could think hard and find a way in which the narrative could be made coherent (I shall call this ‘the reconciliation strategy’). The reader could reclassify the narrative; that is, attribute to the world it describes a principle in which the situation is no longer contradictory (I shall call this the ‘weird world strategy’). Finally, the reader could simply ignore that part of the narrative or put it aside as a flaw, and try to make sense of the rest without it (I shall call this the ‘disregarding strategy’). (Matravers 2013, p. 131)

In the passage from *The Lion the Witch and the Wardrobe* highlighted above, we can use the ‘reconciliation’ approach. The initial reference is to ‘the same fact that Edmund had whispered to Peter’ and we could take this passage as saying everyone recalled the whispered fact itself, rather than that they recalled this fact and also recognised it was whispered at the end of the previous chapter. These sorts of metafictional examples are thus compatible with a distinct content approach to explaining how we engage with

fictions. It will sometimes be hard to explain what is true in a given fiction, but these complexities should not push us into accepting a role for belief-like imaginings here.

As such then, these initial concerns about sophistication, ontology and weird fictions haven't given us reason to introduce belief-like imaginings into our account of how we engage with fiction.

We will now turn to discuss something like a puzzle that philosophers have argued can only be solved by introducing belief-like imaginings. This puzzle concerns how to spell out the distinction between fiction and non-fiction. Several philosophers have suggested this distinction is best explained by supposing fictions invite imaginings whereas non-fictions invite belief, so we should consider whether this issue forces us to introduce belief-like imaginings into our account of how we engage with fiction.

5.3 Fiction & Non-Fiction

The question of how we should distinguish works of fiction from works of non-fiction has received a great deal of attention in contemporary philosophy of fiction. Much of this debate has focused on verbal fictions, such as novels, though it is usually thought similar arguments and issues arise when thinking about plays, films and other visual fictions.¹²²

Lamarque and Olsen emphasise this distinction, suggesting that:

[t]he classification of narrative into fiction and non-fiction is of the utmost significance; not only is it a precondition of making sense of a work, but it determines how we should respond in both thought and action. (Lamarque and Olsen 1994, p. 30)

Whether we classify a work as fiction or non-fiction will have important implications for how we engage with it. If you picked up a dystopian novel such as Orwell's *1984* and didn't realise it was a work of fiction, it would be a disaster for your representations of

¹²² As we will go on to see, this is because much of the debate has focused on how we should understand the propositions that makeup works of verbal fiction but it has been argued this account might also apply to visual fictions (Currie 1990, p. 39)

the world. Suddenly, you could find yourself forming various beliefs about wars that have taken place between Oceania, Eurasia and Eastasia, grossly distorting your understanding of world history.

In recent philosophical work, it has been suggested we should distinguish fiction and non-fiction with reference to the attitude they prescribe readers to take towards their content.¹²³ Defenders of this sort of view argue that works of fiction present propositions that we should *make-believe* or *imagine*, whereas works of non-fiction present propositions that we should *believe*. In this way, we might be forced to introduce belief-like imaginings into an account of how we engage with works of fiction in order to capture this crucial distinction.

This sketch of the attitudes we take towards fiction and non-fiction is too simplistic, however. Most works of non-fiction will contain at least one proposition that we are not supposed to straightforwardly believe, and many works of fiction contain propositions that we are supposed to straightforwardly believe. For example, a work of history might ask us to consider what it would be like if Christianity had never become the world's dominant religion, and a work of fiction can also contain utterances we are supposed to believe. Friend (2012, p. 184) gives an example from *Mary Barton* 'There are some fields near Manchester, well known to the inhabitants as "Green Heys Fields", through which runs a public footpath to a little village about two miles distant'. This looks like an

¹²³ This sort of view is contrasted most frequently with a view where one argues there is something about the form, syntax or semantics of fiction that distinguishes it from the form, syntax or semantics of non-fiction. These views are unpopular because it looks like works of non-fiction and fiction can be written in much the same way (e.g. a recent review in the *London Review of Books* (<https://www.lrb.co.uk/v39/n15/sheila-fitzpatrick/good-communist-homes>) of the history book *The House of Government: A Saga of the Russian Revolution* argues it has much in common stylistically with *War and Peace*). As such, philosophers have attended to the *pragmatic* factors that might distinguish fiction from non-fiction. This relates to my earlier point about recognition: there are no obvious surface features of works of fiction that serve to distinguish them from works of non-fiction (e.g. both a work of fiction or a work of literary non-fiction can begin with the sentence 'Once upon a time ...').

utterance that is supposed to invite straightforward belief, and the same can be said for various other descriptions of places and locations found in works of fiction. There are also more general claims found in fictions that seem to invite belief, such as Tolstoy's 'All happy families are alike; each unhappy family is unhappy in its own way.' Again, this is not something we are supposed to take to be only true in the fiction: it looks like Tolstoy is asserting this to be straightforwardly true.

The most popular way of setting out the difference between fiction and non-fiction is what Stock (2016, pp. 205-209) calls the 'imagining plus' view. This view maintains that prescriptions to imagine are *necessary* for distinguishing fiction from non-fiction, but are not sufficient. These prescriptions are supplemented with reference to *authorial intentions* and some sort of additional condition.

This intention-based view was first set out by Currie (1990) drawing on work from Grice (1957). Currie argues that authors of works of fiction put forward what have come to be known as *fictive utterances* (Lamarque and Olsen 1994, pp. 32-33), and that authors intend for readers to imagine these utterances and to do so *because* they recognise this intention. To see how this works in practice consider a novel like *Harry Potter*. This novel contains various utterances about wizards, wizarding schools and other magical goings on. According to Currie and his followers, the starting point for explaining what makes *Harry Potter* a work of fiction is that these utterances express propositions which are supposed to be imagined, and we are supposed to imagine them because we recognise J.K. Rowling's intention to present these as propositions to be imagined.

Thinkers who have been drawn to this sort of view have also added a second necessary condition to set a demarcation between fictive and non-fictive utterances, driven by several hypothetical examples introduced by Currie (1990, pp. 42-44) For Currie, this condition is that a fictive utterance must be either false or non-accidentally true, which is supposed to rule out, for example, a case where someone writes what they take to be a work of fiction based on repressed childhood memories. Lamarque and Olsen

(1994, p. 44) and Davies (2007, p. 46) offer subtly different twists on this formula but the general idea of introducing an additional condition remains.¹²⁴

However, so far, we have not seen how this view accounts for the difference between *works* of fiction and non-fiction. We have introduced the theoretical notion of a fictive utterance, but as we pointed out earlier a work of non-fiction can also contain fictive utterances, since a fictive utterance is merely one that we are supposed to imagine and not believe. In practice, this means that for Currie and his followers both works of fiction and works of non-fiction are what we can call 'patchworks' (Currie 1990, p. 49) of fictional and non-fictive utterances. This leaves a question of what makes a given patchwork fictional rather than non-fictional.

This focus on fictive utterances also introduces another patchwork problem, the 'patchwork of attitudes' problem (Stock 2011b, Friend 2011, p. 167). If Currie and his followers are right, when we engage with a fiction we will sometimes believe the propositions expressed by utterances that we read on the page, and sometimes imagine them. Indeed, we might even switch between these two attitudes over the course of a single paragraph. However, reading a work of fiction or non-fiction feels like a unified psychological experience, at least introspectively. When reading a work of fiction, there is no immediate phenomenological difference between reading a sentence that sets out a truth as opposed to a fictional truth.

As such, introducing a counterpart to belief to explain the distinction between fiction and non-fiction may well serve as an example of where introducing an imaginative

¹²⁴ A good recent summary of this debate and various existing views can be found in Kajtar (2017). I will not focus on these developments here, since my claim that we don't need to introduce belief-like imaginings to explain the distinction between fiction and non-fiction does not depend on adopting any specific imagining-plus view. For what it's worth, I have some sympathy with Davies' (2007, p. 46) 'fidelity constraint', which an author follows 'if they include only the events she believes to have occurred, narrated as occurring in the order in which she believes them to have occurred', since I would agree with him that there is more liberty for authors of fiction to depart from truths about the actual world in their narratives as compared to authors of non-fiction.

counterpart not only fails to solve a purported puzzle (how to demarcate fictions from non-fictions), but also serves to deepen the puzzle. This is because this approach looks like it will lead to us maintaining that there is a distinct attitude involved in reading fictions as compared to non-fictions, when this is not phenomenological evident on reflection.

This, of course, does not rule out an involuntary counterpart being involved here since an involuntary counterpart to belief could be phenomenologically identical to genuine belief. However, this also doesn't rule out an approach where we recast being prescribed to imagine in terms of being prescribed to believe something to be fictionally true. The difficulty here, however, is that we would still face the initial patchwork problem: we have no simple explanation of how to move from noting the presence of utterances that express propositions we are supposed to believe to be fictional to categorising a work as fiction or non-fiction.

As such, I think a better sort of approach to distinguishing fiction and non-fiction is offered by Friend, who dispenses with any reference to prescriptions to imagine. Friend (2012) suggests that we should think of fiction and non-fiction as being *genres*. Drawing on Walton's categories of art (1970), she argues that there are standard, contra-standard and variable features associated with being a work of fiction or non-fiction. Standard features are the features a member of a category normally exhibit, such as being on a canvas for a member of the painting category. Variable features are those that can vary amongst members of a category, such as the colours used to paint a particular painting. Contra-standard features are features that normally rule out membership of a given category, such as being observable in the round for a painting.¹²⁵

¹²⁵ In terms of appreciation, this has important implications. This does not mean that there cannot be a painting that is observable in the round. What it means is that if we categorise something observable in the round as a painting, we will see this as a striking feature of the work. If we categorise something observable in the round as a sculpture, this will normally be less striking since this is a standard feature of statuary.

Friend (2012, p. 190) argues that containing many statements we are not supposed to straightforwardly believe is a standard feature of fiction. This explains why we find some works hard to classify. She gives the example of the Reagan autobiography *Dutch*, which features various obviously made up segments. This sort of material is contra-standard for non-fiction, and this is why it is unclear how we should classify this work. In response to this sort of problem case, the imagining-plus view needs to find a way of moving from noting the presence of these apparently fictive utterances, to establishing whether the work is fiction or non-fiction. On the other hand, Friend can simply point out that we ought to be surprised by the presence of this content, but can nonetheless reflect on factors like how the work was classified on release, what sort of work the author took himself to be writing, and so on.

At this point, it will be helpful to take a step back from our discussion of the distinction between fiction and non-fiction. The orthodox view of the fiction and non-fiction distinction explains it in terms of the attitude that we are prescribed to take towards the utterances that make up the work in question. This raises various patchwork worries, and in particular, does not rule out that fictions prescribe us to believe things are fictionally true. In light of these patchwork concerns, I prefer Friend's approach which doesn't rely on making a move from classifying an utterance as fictional to classifying a work as a whole as fictional. However, endorsing this genre approach will not allow us to simply conclude that there is no role for belief-like imaginings in our engagement with fiction. Indeed, Friend argues that belief-like imagining can play a role in our engagement with both fiction and non-fiction. She relies on considerations from discourse processing to establish this view, and we will now turn to consider whether this research suggests that we should associate belief-like imaginings with fiction.

5.4 Discourse Processing

Psychologists have developed accounts of discourse processing to explain how we engage with all kinds of texts, and as such psychologists working in this area do not tend to

argue that there is one way in which readers engaging with fiction and another way in which they engage with non-fiction.¹²⁶

The psychological consensus is that there are multiple levels of memory representation that we form when engaging with texts. A representative example of these different levels – adapted from a recent review article (Graesser and Forsyth 2013, p. 477, Table 30.1) – is as follows:

1. Words and Syntax: a representation of the actual words that make up the discourse along with the syntax of the propositions that make up the discourse.
2. Textbase: paraphrased representations of the propositions in a discourse.
3. Situation model: a model of the situation described in a text, supplemented by pre-existing knowledge.
4. Genre: information about the genre of a text and relevant classificatory issues (e.g. is it a work of a fiction? A newspaper article?).
5. Pragmatic communication: information about the goals of the author and the intentions behind their discourse.

These levels of representation shouldn't be thought of as being entirely independent of one another. Instead, the idea is that these levels constitute different dimensions to our engagement with a discourse and reflect the different psychological constructs and cognitive processing involved in our engagement with texts.¹²⁷ That being said, not all psychologists accept the existence of these different levels of representation. Most would at least accept the existence of the textbase (perhaps combining this with the words and

¹²⁶ Matravers offers a helpfully summary of much of this research in chapters 4-7 of his *Fiction and Narrative* (2013). His overview is admittedly biased since he takes this psychological work to show that the consensus view that fiction involves a distinct imaginative attitude is mistaken.

¹²⁷ Various psychological methods are used to show these levels exist, such as verbal reports and attention tests (Graesser *et al.* 1997).

syntax level) and the situation model level.¹²⁸ Roughly, this marks the distinction between our representation of the propositions that make up a text and of the events described in a text.

For present purposes, we need to consider whether any of these levels involve belief-like imaginings, either when engaging with any sort of discourse or when engaging with works of fiction. So far as I'm aware, it has not been argued that the first two levels of comprehension involve belief-like imaginings. These levels relate to our initial viewing of words on a page and recognising them as making up sentences, followed by us forming a representation not of the literal form of the sentence, but a paraphrase of its meaning. These two levels are supposed to be sub-personal and automatic, with this processing occurring in the same way for all kinds of discourse (Sparks and Rapp 2010). Later levels, represented by (4) and (5) also don't seem to require belief-like imaginings. Weinberg and Meskin (2006a, pp. 229-231) explicitly allow that when engaging with fiction we form beliefs about things like genre, and they would presumably say something similar about pragmatic issues related to the goals of the author.¹²⁹

This leaves (4), the situation model, as the level where belief-like imaginings might be involved and this is where Friend (n.d.) draws a connection between fiction and belief-like imaginings. The situation model is the eventual result of our engagement with a text and represents the various happenings in a story. Johnson-Laird (quoted in Matravers 2013) tells us that:

¹²⁸ We can also call these 'mental models' (Johnson Laird 1983), a name which appears to be more frequently used in philosophical discussions.

¹²⁹ They consider this in a discussion about, among other things, how beliefs about actors and actresses can colour our viewing of a visual fiction and how beliefs about genre can influence our engagement with all kinds of fiction. The distinct content view is well placed to make sense of this. You just do believe that James Stewart is an all-around nice guy, and this interacts with your belief that in a given film, a character played by him is not such a nice guy. This is an issue where simple beliefs alone might struggle to capture how we engage with fiction, but once we accept that some of our beliefs will have content about what is true in fiction, this can allow for interesting interactions between our standing beliefs and beliefs about fictions.

[situation models] play a central and unifying role in representing objects, states of affairs, sequences of events, the way the world is, and the social and psychological actions of daily life. They enable individuals to make inferences and predictions, to understand phenomena, to decide what action to take to control its execution, and above all to experience events by proxy; they allow language to be used to create representations comparable to those deriving from direct acquaintance with the world; and they relate word to the world by way of conception and perception. (Johnson-Laird 1983, p. 397)

One way to motivate the suggestion that situation models involve forming belief-like imaginings is to note that the models that we produce when we engage with texts are ordinarily *compartmentalised* (Potts *et al.* 1989). That is to say, we do not automatically integrate information from these models into our pre-existing structures of belief. Studies show that this happens with non-fictions as well as fictions. This is obvious on reflection, since so long as we are being appropriately critical, we don't come to believe that everything we read in a work of non-fiction is true.¹³⁰

This introduces the question of what sort of attitude we take towards the compartmentalised propositions found in a situation model. Matravers argues that we take no particular attitude towards them whatsoever: we merely regard them as 'representations in a situation model'. Some of these propositions we will go on to believe, and this is the only point at which we take a specific cognitive attitude towards them. As he puts this:

If a proposition in a fictional narrative does not become a belief, the proposition's role in our cognitive economy is only that of a proposition that forms part of the content of a narrative. It has this role in common with the other propositions that are part of the [situation] model we form on engaging with a narrative, whether non-fictional or fictional. If a proposition in a fictional narrative does become a belief, it has a role

¹³⁰ Interestingly, there is some evidence that we are in fact less discriminatory when reading works of fiction as to what we come to believe after (Prentice and Gerrig 1999, cited in Matravers 2013 p. 97). One reason for this might be that when we read a work we believe to be non-fictional (or which we believe to purport to offer a faithful account of the actual world) we pay close attention to suspect claims, whereas with a fiction (or something we don't believe to offer a faithful account of the actual world) we are less discriminating of suspect claims.

as a proposition that forms part of a narrative and it becomes a belief. (Matravers 2013, p. 96)

Friend (n.d.), on the other hand, argues that we imagine these propositions in a belief-like way and that this means belief-like imaginings play a key role in our engagement with fiction. She associates forming a situation model with the notion of ‘imagining a world’ and argues that imagining a world involves certain forms of non-propositional imaginings, such as mental imagery and imagined perspectives, along with propositional, belief-like imaginings.¹³¹ This means she accepts that situation models have an *analogical*, rather than *propositional* structure. Friend defends this view by adopting a particular approach to understanding how we construct at least some situation models: the ‘event indexing’ model (Zwaan *et al.* 1995).¹³² Accepting for the sake of argument that this a plausible account of how we form models of the situations described in works of fiction, we should consider whether Friend is right that we have to be understood as imagining the propositions found in a situation model. If she is, then we necessarily have to introduce belief-like imaginings into our account of how we engage with works of fiction. Indeed, this may even make sense of my clustering worries in relation to the distinct attitude view. If one thinks that we can explain clustering in terms of the formation of situation models, and situation models often contain belief-like imaginings, then the distinct attitude theory will have at least some sort of response to my concerns about how to explain clustering.

A first point to make is that psychologists do not argue that the propositional contents in situation models are marked as relating to ‘fiction’ at the level of content.

¹³¹ It is perhaps interesting that this is similar to the notion of an *imaginative exercise* which Everett used in the initial quote of this thesis about the intuitive link between fiction and imagination.

¹³² This event indexing approach is supposed to help explain how we form models of discourses where the events are described in a different order to the order in which they actually occurred. This is often necessary when engaging with fiction, since not all fictions give a neat and ordered timeline. Even when they do, we still have to do some work to figure out how the different described events fit together.

Friend notes that situation models themselves may well be marked with a source ‘tag’ such as ‘in *Harry Potter*’, but the propositions within the model are supposed to be represented *simpliciter*. We could deny the scientific consensus on this matter, but it will be more productive to consider whether we are forced to accept that these compartmentalised propositions are imagined in a belief-like way.

To respond to Friend’s proposal, it will be helpful to quote Matravers again on this point:

In as much as we have an attitude to them [compartmentalised propositions we do not believe] at all, it is merely one of them being part of the content of whatever particular representation we are reading or remembering. We may have an attitude to the representation as to it being fictional and non-fictional ... but that is a different matter. (Matravers 2013, p. 79)

I take it part of what Matravers wants to bring out here is that often we do not pay much attention to the contents of situation models, so it might seem implausible to think that all the propositions within them are either believed or imagined in a belief-like way. As such, these compartmentalised propositions don’t seem to exhibit the richness that philosophers like Nichols and Currie seem to have in mind when they associate belief-like imaginings with our engagement with fiction:

Unless we have reason to think that the psychological processes of reading texts such as histories, biographies, and novels should differ completely from those that operate when reading the texts that feature in psychology experiments (a claim that would be empirically incredible) we have reason to think that the subsequent mental models are going to be rather vague and sketchy. In short, what goes on in our heads when we read is, generally, a great deal less exciting than we might have thought. (Matravers 2013, p. 73)

Furthermore, it’s worth reminding ourselves that not believing p does not mean we necessarily imagine p in a belief-like way. This is particularly relevant in the present discussion, since construction of situation models goes on when we read any sort of narrative, including works of non-fiction. It would be counter-intuitive to argue that whenever we take a claim to be false in a work of non-fiction we imagine it in a belief-like way, as opposed to just regarding it as part of the content of what we are reading.

For example, when I read a newspaper article I disagree with, it would be odd to describe myself as taking a belief-like imaginative attitude towards its contestable claims.

At this point, we can anticipate the following response from someone defending the idea that we should associate fiction and belief-like imaginings. It could be argued that although belief-like imaginings are not involved in *constructing* situation models, they are nonetheless *generated* on the basis of constructing models related to works of fiction. We could suggest, for example, that we form belief-like imaginings when we reflect on the events described in fictions, or on what is true in a fiction. However, we can also suggest that when we reflect in these ways beliefs about what is fictional are generated. As of yet, we have no reason to think that beliefs subject to a fictional operator are unable to serve this role in our reflection on works of fiction.

As such, we have yet to find any constraints on belief, or puzzles related to our engagement with fiction that necessitate the introduction of belief-like imaginings. Indeed, in the case of sophistication, we saw that clustering might reveal a way in which the distinct attitude view is insufficient for explaining how we are able to engage with fictions. This leaves a final important issue to consider, which is whether beliefs about what is true in fiction can explain the way we can become *immersed* in works of fiction.

5.5 Immersion & Affect

The most important objections to a distinct content view of fiction arises when thinking about what we can call *immersion*. I will consider two issues here, one related to what Gerrig calls ‘transportation’, and another related to affect raised by Weinberg and Meskin.

Gerrig points out that fictions have a remarkable ability to engage our attention. He illustrates this with the metaphor of transportation:

Readers become ‘lost in a book’; moviegoers are surprised when the lights come back up; television viewers care desperately about the fates of soap opera characters; museum visitors are captivated by the stories encoded in daubs of paint. In each case, a narrative serves to *transport* an experiencer away from the here and now. (emphasis mine) (Gerrig 1993, p. 3)

This notion has since been made more formal by Green and Brock (2000), who introduce a 15-point scale for transportation. Transportation is thought to partly explain many different aspects of our engagement with fiction, such as why we find it enjoyable to engage with fictions, why we end up being persuaded by merely fictional works, and why we have emotional responses towards them.¹³³

This allows us to introduce a worry about whether beliefs about fictions are sufficient for explaining this sort of transportation. This is because they look to be *external* rather than *internal* attitudes. Perhaps we need internal attitudes – namely belief-like imaginings – in order to be transported by a work of fiction.

However, this worry is not entirely compelling. I have two objections here. Firstly, it is unclear why beliefs about a fiction are unable to explain transportation and immersion while belief-like imaginings can. This mirrors a debate about whether there is a role for desire-like imaginings when accounting for childhood pretence. Velleman (2000, p. 257) argues that an explanation of pretence in terms of desires to pretend in certain ways is *depressingly un-childlike* and cannot explain immersion because it renders children as like actors following a script. However, Funkhouser and Spaulding (2009, pp. 12-13) rightly respond that this trades on a stipulation about what sort of states can and cannot explain immersion. In the case of beliefs about the fiction, having these sorts of beliefs does not have to entail that you're constantly reflecting on the fact that you are engaging with a fiction, with this somehow preventing immersion. Instead, it just means that when we engage with fictions we form beliefs *implicitly* subject to a fictional operator. We would require a further argument for why it matters whether the relevant attitude is 'internal' or 'external'. Weinberg and Meskin perhaps offer the relevant sort of argument here in relation to affect, and we will discuss their argument shortly.

The second objection to make here is that works of narrative non-fiction are equally capable of transporting us (Green and Brock 2000), which suggests that belief-like

¹³³ A good recent summary of these connections, among others, can be found in Green and Donahue (2009).

imaginings are *unnecessary* for transportation, since the cognitive attitude involved in our engagement with non-fiction is usually agreed to be belief. That being said, we may well need to associate some other aspects of imagination with transportation in order to explain why we are sometimes transported by works of non-fiction. For example, while reading a piece of long-form journalism about a war zone, I might begin to form imagery of a war-torn environment, or what it would be like to be one of the subjects of the piece. As such, in the case of non-fiction, it looks like beliefs are either redundant for explaining transportation, or are able to interact with objectual and sympathetic imaginings to allow for transportation. If the former is right, then even if we introduce belief-like imaginings, this reveals that they would presumably also be redundant when it comes to explaining transportation when engaging with works of fiction. If the latter is correct, then we can allow that beliefs about fiction can also interact with non-propositional imaginings to transport us when engaging with works of fiction.

Having noted why these initial worries about immersion shouldn't make us endorse a distinct attitude view of how we engage with fiction, we can now turn to consider Weinberg and Meskin's argument for why beliefs about fiction cannot prompt affective responses.¹³⁴

The argument for why distinct content views struggle to explain our affective responses to fictions is helpfully set out Weinberg and Meskin when they note, after having considered and rejected an illusion-based theory of fiction, that:

The failures of the illusion theory suggest a different belief-based theory: to imagine the fictional content *p* is simply to hold the belief about the fiction that in-the-story-it-is-true-that-*p*. Although the "meta-representational" theory clearly entails behavioral circumscription (since such beliefs about a fictional story will rarely prompt us to an action), it also seems unable to explain phenomenological/physiological robustness. Fearful affective responses, for example, seem to require representations of the form *S* is in danger, where *S* is someone we care about. But it is not enough for the representation to be a subpart of another representation. For example, if you believe,

¹³⁴ Boruah (1988, pp. 59-64) also raises some related worries here. I take them to have been adequately dealt with by Neill (1993), so I will not elaborate on them for the purposes of this discussion.

not that the slime is threatening, but that a friend think the slime is threatening, your only fear will be for that friend's sanity; similarly for a belief that it is metaphysically possible that a slime is threatening. So a belief like in the novel it is true that S is in danger is not of the right form to generate correct affective responses. (Weinberg and Meskin 2006a, p. 225)

As such, Weinberg and Meskin argue that beliefs about the fiction are *insufficient* for generating affective responses. They then go on to argue that they are also *unnecessary*, since mere imaginings can generate affective responses:

A further problem for the meta representational theory is that we often have affective responses to imaginings that are not derived from any fiction. Merely imagining a close friend's being in great pain may be enough to produce a pang of pity, without there being any work of fiction at all concerning your friend and his or her suffering While it might seem tempting to suggest that in such cases a minimal story is created, such a view leaves open the question of what it is to be a story. And the meta representational theorist cannot explain what a story is in terms of imagining, since that would lead to vicious circularity. (Weinberg and Meskin 2006a, p. 225)

We will begin with this second argument. As we have already seen, I think we have reason to be sceptical that when we talk about a notion of imagination related to fiction it is the same as the one related to imagining things about a close friend. For one thing, the latter seems to involve a voluntary, rather than involuntary attitude. As we noted, it could be that belief-like imaginings behave in different ways in different contexts but we will need a positive argument for this claim, and as such whether this second worry needs to concern us depends in part on whether we find the insufficient argument compelling.¹³⁵

Having made this response to their 'unnecessary' argument, we can now consider their first argument, the 'insufficient' argument. We can break down Weinberg and Meskin's argument as being something like the following:

¹³⁵ It's also worth bringing out here that they refer to imagining 'your sister being in pain'. On the face of it, this is not a propositional imagining, but something like an imagined experience. One response here would be to say that mere propositional imaginings, even if they do exist, are not enough to generate affective responses in the first place and that we need to refer to other forms of imagining.

P1: When you engage with a fiction, you form beliefs about what is true in the fiction

P2: When you engage with a fiction, you exhibit emotional responses towards the events depicted in the fiction

P3: In order to exhibit an emotional response to something, you need a representation of the form X is F¹³⁶

P4: Beliefs about what is true in fiction are not representations of the form X is F

C: Beliefs about what is true in fiction cannot be what generates our emotional responses to fictions

We can note first of all that P1 and P2 are hard to deny. P1 is a truism unless one has a restrictive view of reference, which we saw in section 5.2.2 we do not need to commit ourselves to. P2 one might prefer to rephrase as ‘emotion-like’ responses if they think our responses to fiction involve quasi emotions, but otherwise seems like a truism. P4 is also hard to deny. If I believe that S is threatened [in the fiction] this is not a representation of the form X is F since it is a representation of the form X is F [in the fiction].

As such, we should see if we have reason to challenge P3. A first question is how strong Weinberg and Meskin intend their claim to be. One way of reading this claim would be that *all* emotional responses require a representation of the form X is F. The other is that *some* emotional responses require a representation of the form X is F. I am going to take it they mean ‘some’ here, since the ‘all’ claim will quickly run into difficulties

¹³⁶ This claim is different to the belief claim we used to introduce the paradox of fiction in chapter 1. Their claim here is only that we need to take some sort of attitude towards a representation of the form X is F to generate an emotional response. This does not commit them to the stronger view that the relevant attitude has to be belief, which is of course a crucial distinction for them since they think belief-like imaginings are what prompt emotional responses to fictions (Meskin and Weinberg 2003)

when thinking about automatic reflexes, such as being *startled* by loud noises. Indeed, one can also be startled by events depicted in a work of fiction even if one doesn't have a representation of the form 'X is startling'; at least in the case of visual fictions like films that sometimes make use of sudden images and sounds.¹³⁷ That being said, there are some important emotional responses to fiction that look like they might require this sort of representation. If I feel pride for Superman, presumably this means I do have some sort of representation about Superman, as opposed to my pride being a mere automatic bodily response.

A second point to make in this respect is that having a representation of the form X is F (be it a belief or a belief-like imagining) is not *sufficient* for generating an affective response. At the very least, we will need some sort of desire to accompany the relevant representation. We saw this in chapter 3, where we noted that Nichols (2006) argues that sometimes belief-like imaginings seem to lead to affective responses and sometimes they do not. Indeed, if at this very moment I imagine that I am being chased by a monster, this doesn't seem to make me fearful.

It is also worth noting at this point that, as I suggested in response to worries about transportation, imaginative aspects of our engagement with fictions – aside from belief-like imaginings – likely play an important role in generating our emotional responses to fictions. For example, Van Leeuwen (2011, p. 66) emphasises the importance of imagery in generating affective responses, arguing that 'There is reason to think that imagining in, say, visual or auditory detail has far greater emotional impact than bare, propositional imagining'. Indeed, this is backed up by psychological work on transportation, where a higher degree of transportation is associated with the formation of imagery.

The association of these further kinds of imaginings with fiction is compatible with my preferred distinct content view. If we need to make reference to our objectual and sympathetic imaginings to explain our emotional responses to fiction, then the reason that Weinberg and Meskin's worry arises is because they're asking beliefs about the

¹³⁷ On the other hand, I doubt that one can be 'startled' by a novel.

fiction to do too much explanatory work. However, this worry will equally arise if we seek to explain our affective responses to fictions solely in terms of belief-like imaginings.

In response, Weinberg and Meskin could argue that although belief-like imaginings are insufficient for generating affective responses, at least they don't rule them out: they could argue that beliefs about what is fictional shouldn't be able to play a role in generating affective responses regardless of which further imaginings might accompany them. In other words, they could move to argue that a representation of the form X is F can combine with other mental states to produce affective responses, whereas beliefs about fiction cannot.

Whether this is right will turn on their suggestion that in order for a representation to lead to an emotional response, 'it is not enough for the representation to be a subpart of another representation'. However, they do not make it explicit what a *subpart* of a representation is supposed to be. Their examples of the sort of representations that cannot lead to affective responses because they involve subparts are 1) if you believe that a friend thinks a slime is threatening, and 2) if you believe that it is metaphysically possible that a slime is threatening. But, suppose I believe that my sister will be threatened *in two weeks*. Does this mean 'my sister will be threatened' is a mere subpart because my representation contains a temporal operator? Likewise, if we stretch the notion of a subpart, we could argue that an everyday representation that leads to a fear response is of the form X is threatened *in reality*. Neill develops a similar point:

[t]o the extent that my belief that fictionally Shylock is a victim of injustice can be construed as "a mere recognition that fictionally, something is the case," my belief that many Guatemalan refugees are victims of injustice can be construed similarly as "a mere recognition that, actually, something is the case." And there is no reason to suppose that "mere recognition" of what is fictional is any less causally efficacious with respect to emotion than "mere recognition" of what is actual. (Neill 1993, p. 3)

What then, is supposed to be special about the subpart involved in thinking that X is threatened *in the fiction* which is supposed to rule out this representation provoking an emotional response? Indeed, we seem to care about the trials and tribulations of fictional characters, and in light of this, it is perhaps only to be expected that we exhibit emotional responses towards them.

Weinberg and Meskin's intuition here might relate to what is entailed by one's representations of the world. If I believe that *in two weeks* my sister will be threatened, this entails that I do in fact believe that my sister is threatened. Likewise, if I believe that my sister is threatened *in reality*, this implies that my sister is threatened *simpliciter*. The fiction case is different. If I believe that my sister is threatened *in the fiction*, then I don't believe she is actually threatened: all that is entailed by this belief is that she is threatened in the fiction. A similar point applies to both further examples of subparts they give. If you believe that your sister thinks she is threatened, you won't believe that she is actually threatened – and if you believe a slime is metaphysically possible, you won't believe that you are actually threatened by one.

To respond to this way of developing the worry about subparts, it will be helpful to return to consider how distinct attitude theorists make sense of our affective responses to fiction. In chapters 2 and 3, we saw that the distinct attitude theorists argue that belief-like imaginings and beliefs can interact with our affective systems in much the same way, and as such can prompt affective responses. One way to explain why this might happen is offered by Harris (2000, Ch. 4), who argues that emotional responses are the product of *appraisals*, a view associated with Lazarus (1991). According to appraisal theory, our emotional responses result from how we appraise situations across various dimensions, such as how a given situation relates to our goals. To explain why we exhibit emotional responses to the merely fictional and pretend, Harris (2000, pp. 65-67) suggests that this appraisal process may not discriminate between fictional and real inputs.

Now, this sort of suggestion can be understood as entailing that there is an insensitivity between belief and imaginings in our appraisal processes, but we can equally view this as entailing that there is an insensitivity between different sorts of *content* in our appraisal processes. As such, we could argue that whether or not we represent something as fictional or real at the level of content has no (or limited) bearing on the affective responses we exhibit to it.

This gives us a way of responding to Weinberg and Meskin's worry about subparts. When you merely believe that your sister thinks she is threatened, you will not feel fear because you don't appraise this as a situation that merits a fear response. In a fiction

case, you do perform this sort of appraisal because this appraisal process is neutral as to whether a representation concerns reality or fiction. In the sister and metaphysical possibility examples, you merely think someone isn't actually threatened, while in the fiction case you do believe someone is threatened, albeit in fiction. If this is right, then although our affective systems are not always neutral between cases where we take someone to be threatened and someone to not be threatened, they nonetheless are somewhat neutral in the special case of beliefs about fiction.

One might object, however, that this view ignores the fact that differences in content ordinarily do alter our emotional responses. If you have a representation of the form X is annoying, this will play a role in making you angry; if you instead represent that X is scary, this will play a role in making you fearful, and so on. Similarly, we might object that judging that X is annoying [in the fiction] should lead to a different affective response as compared to judging that X is annoying *simpliciter*. However, as we noted in our discussion of quasi emotions in section 1.6, our response to fictions often are indeed somewhat different to the ones we exhibit to ordinary stimuli. For example, Charles, who clutches his arm rests and yelps as an on-screen slime glares at him, is not motivated to respond in the usual way to feeling fear and makes no attempt to flee his cinema seat.

As such, we could take Weinberg and Meskin's worry as an argument for the suggestion that beliefs about the fiction lead to us feeling quasi emotions, rather than emotions *simpliciter*. This might seem like an odd response bearing in mind that I am denying a role for the perhaps more intuitive notion of belief-like imaginings in our engagement with fiction. However, it's worth emphasising that suggesting that we form belief-like imaginings when we engage with fiction does not mean quasi emotions are rendered explanatorily redundant. Walton's motivation for introducing these counterparts was to account for some puzzling features of our emotional responses to fiction, in particular that they seem to have different connections with motivation as compared to ordinary responses. We can remain neutral about this issue for present purposes, since the more important point to bring out here is that if one is concerned that we ought to respond differently to beliefs about fiction as compared to

straightforward beliefs, there is some evidence that we do respond differently once we look beyond mere phenomenology and bodily sensations.

At this point, Weinberg and Meskin might want to remind us that the paradox of fiction introduces not only a causal worry, but also a worry about rationality. They could argue that even if there is some way of explaining how beliefs about fiction can *cause* emotional responses, it is nonetheless *irrational* for us to exhibit these responses. In turn, they could argue that it would be more rational to respond emotionally to our belief-like imaginings, and that as such we have reason to prefer a distinct attitude account of our emotional responses to fiction.

To see how we might respond to this rationality worry it will be helpful to introduce some responses developed by distinct attitude theorists. One common response to the rationality worry made by defenders of distinct attitude views, is that emotional responses to fictional scenarios help to inform our future actions and behaviour in much the same way as our emotional responses to actual scenarios. For example, Robinson (1995) claims that fictions can help educate our emotions so we exhibit appropriate emotional responses to real world happenings. In a somewhat related vein, Nussbaum (1984) argues that engaging with fictions and responding emotionally to them can develop our moral sensibilities.

However, all of these sorts of considerations are equally applicable to an account founded on the possibility of our having emotional responses to beliefs about what is fictional. These responses trade on pointing out how our emotional responses to fiction might be useful in our everyday life, and the utility of these responses remains if we say they are produced in response to beliefs about what is true in fiction.

It is also worth noting that Davies (2009) offers some helpful responses to the rationality worry on behalf of a belief-based view. He argues that it is, in fact, perfectly rational to feel emotional responses towards what you believe to be fictional, since for the fictional characters we feel emotions towards, the relevant issue is what is true in the fiction. Put simply, if we have any reason to feel pity for Anna Karenina (and not another emotion), it will be because of what happens to her *in the fictional world*, not in reality:

The belief relevant to Diana's pitying Anna Karenina is not that Anna exists in the actual world but that she suffers in the world of the fiction. ... so long as she is aware of the context that places Anna in a fictional or possible world, Anna's non-existence in the actual world should not inhibit Diana's pity for Anna. And notice that Diana's response of pity now can be seen as rational, because she has the pity-relevant belief, which is that Anna suffers, whereas by contrast, amusement at Anna's plight and satisfaction at her suicide would not be appropriate. (Davies 2009, p. 270)

Furthermore, he draws on Moran to argue that emotional responses to fictions are not unique in being somewhat disconnected from existence beliefs. Moran notes:

Relief, regret, remorse and nostalgia are, after all, among the paradigm cases of emotional response; and although they are essentially backward-looking, they are not commonly thought to present any special puzzle among the emotions. But at the same time from within everyday psychology we have also to confront the proverbial injunction against crying over spilt milk, which can itself seem very puzzling. After all, if we can't cry after the milk is spilled, when can we cry? Presumably not on confronting the milk still safely in the bottle. (Moran 1994, p. 78-79)

We could argue in response to this sort of argument that these cases are somewhat different to the fiction case because you believe something existed in the past, not that something simply doesn't exist. However, the important point here is that we have no principled reason for saying that it is rational to have affective responses as a result of past-directed belief, while maintaining that emotional responses to fiction-directed beliefs are irrational.

It is thus unclear why we should think it would be rational to respond emotionally to belief-like imaginings, yet irrational to respond emotionally to a belief about what is true in fiction. This is not to say that this rationality worry is easily dealt with and none of the points canvassed above will conclusively respond to this worry (a helpful summary of this debate can be found in Friend 2016, pp. 223-227). For example, we haven't considered whether it is rational for Charles to either fear or quasi fear for himself, bearing in mind he doesn't think he is actually threatened: Davies (2009, p. 281) concludes Charles is, in fact, irrational, and presumably he would maintain that this is true regardless of whether we endorse a distinct content or distinct attitude view of our emotional responses to fiction. The key point here is there is no clear-cut reason why introducing a distinct attitude makes it easier to respond to rationality worries, as

compared to allowing that our emotional responses to fiction are caused by beliefs with distinct contents.

As such, it looks we can respond to both causal and rationality worries about beliefs about the fiction leading to affective responses. This will also be helpful for explaining our affective response to pretence, since the meta-representational beliefs introduced by Leslie have a similar form to beliefs subject to a fictional operator. That said, there are important differences between the two, in particular the fact these sorts of representations can be formed at will in the case of pretence. This raises interesting questions about why, for example, someone can respond emotionally to something they voluntarily represent, an issue that does not arise when we reflect on our engagement with fiction. We do not have the space to discuss this matter further, but this is a good example of where separating out our theories of pretence and engagement with fiction can provide interesting avenues for future research.

Conclusion

By comparing fiction and pretence, we saw in section 5.1 that it looks like the counterpart to belief involved in our engagement with fiction may well be an involuntary counterpart, since it is supposed to be an automatic response to reading words on the page or seeing images on the screen, and so on. This gave us reason to question the uniformity assumption and to argue that justifying the involvement of a counterpart to belief in our engagement with fiction must be done separately from justifying the involvement of one in hypothetical reasoning or pretence. In light of this, and in line with my second and third principles about imaginary counterparts, we considered from sections 5.2 to 5.6 whether constraints on belief that arise when thinking about our engagement with fiction suggest that beliefs cannot be the only cognitive attitude involved in our engagement with fiction, or whether puzzles arose if we maintained that beliefs are the only cognitive attitude associated with fiction.

We saw that issues related to ontology, the fact fictions are weird, conceptual sophistication and discourse processing did not reveal any constraints on belief that give

us reason to think that beliefs about what is fictional will fail to do all the explanatory work when it comes to explaining our engagement with fiction. Indeed, when we considered issues related to sophistication, we saw, as we did in chapter 3, that we still require some sort of account of how distinct attitude approaches explain clustering. This suggests perhaps one way in which a distinct attitude is insufficient for explaining our engagement with fiction. Furthermore, puzzles related to the fiction and non-fiction distinction, and why we respond emotionally to fictions, cannot be resolved solely by introducing belief-like imaginings, and are not rendered harder to solve by supposing that only beliefs are involved in our engagement with fiction.

As such, we have found no reason to necessarily introduce a belief-like attitude into our account of how we engage with works of fiction, and so can reject the non-doxastic assumption in the context of fiction. That being said, this does not have to mean rejecting any link between imagination and fiction. This should instead point us towards developing an account of how we engage with fiction that looks beyond the mere propositional elements of our engagement.

Concluding Thoughts

It is commonly argued that when reflecting on our engagement with fiction and in pretence, we should embrace the uniformity and non-doxastic assumptions, according to which the same cognitive attitude is involved in both of these activities: an imaginative counterpart to belief. By marking a distinction between voluntary and involuntary counterparts in chapter 1, we were able to challenge the first of these two assumptions, the uniformity assumption. In chapter 2, I suggested that if a belief-like counterpart is involved in our engagement with fiction, it will have to be an involuntary counterpart, whereas if one is involved in pretence, it will have to be a voluntary counterpart. We further developed this challenge in chapter 5, by reflecting on the similarities and differences between engaging with fiction and in pretence, and in doing so I argued that we have good reason to reject the uniformity assumption once we note the differing explanatory demands of these two activities.

We also challenged the second of these two assumptions, the non-doxastic assumption, by arguing that a belief-like attitude is neither necessary nor sufficient for explaining our engagement with fiction and in pretence. We first questioned sufficiency in chapter 3, where we saw that introducing belief-like imaginings does not make sense of why our representations related to fiction and pretence exhibit clustering. In chapter 4, I argued that to explain pretence recognition and motivation we need to introduce meta-representational beliefs of the form ‘I PRETEND “*p*”’, since introducing belief-like imaginings is insufficient for making sense of these two issues. I also argued that once we introduce these beliefs, belief-like imaginings are rendered unnecessary for explaining our engagement in pretence. Finally, in chapter 5, I argued that beliefs subject to a fictional operator can explain the cognitive elements of our engagement with fiction, and that in light of this introducing a distinct belief-like attitude is also unnecessary in the context of our engagement with fiction.

In challenging these two assumptions, we have also seen why distinct content based approaches can potentially offer compelling accounts of how we engage with fiction and in pretence. I do not anticipate that I will have been able to convince every philosopher

to give up on belief-like imaginings when thinking about fiction and pretence, but I hope to have at least shown that further arguments are needed to justify the introduction of these sorts of imaginings, at least in the context of our engagement with fiction and in pretence.

However, this leaves a residual question about whether the fact that I have defended distinct content accounts of how we engage with fiction and in pretence entails that I accept a version of the uniformity assumption. On the one hand, I have argued that both these activities involve the same cognitive attitude: belief. On the other, I have argued that the relevant beliefs have somewhat different contents in each case. As such, I am concerned that it will be potentially misleading to count myself as adopting a version of the uniformity assumption, since this terminology risks failing to reflect the fact that I am arguing that the same *attitude* is involved in our engagement with fiction and in pretence, but different *contents*.

Finally, I would like to make some suggestions about potential directions for future research. It is important to remind ourselves that even if engaging with fiction and in pretence doesn't involve belief-like imaginings, engaging in these activities might still involve some non-propositional varieties of imagination, such as objectual and sympathetic imaginings. I also suspect that the creative imagination will play an important role in fiction-making, and perhaps also in pretence when it comes to generating novel behaviours. An interesting issue to pursue in future research is exactly what role these forms of imagination play in our engagement with fiction and in pretence.

Bibliography

- Arcangeli, M. (2014). Against Cognitivism About Supposition. *Philosophia* 42 (3): 607-624.
- Aristotle, *De Anima II and III* (1999). Translation by Hamlyn, D, Oxford: Oxford University Press.
- Baron-Cohen, S; Leslie, A. M. & Frith, U (1985). Does the autistic child have a “theory of mind”? *Cognition* 21 (1): 37-46.
- Bermúdez, J and Cahen, A, Nonconceptual Mental Content, *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2015/entries/content-nonconceptual/>>.
- Boruah, B. (1988). *Fiction and Emotion*. Clarendon Press.
- Bosco, F. M., Friedman, O., & Leslie, A. M. (2006). Recognition of pretend and real actions in play by 1-and 2-year-olds: Early success and why they fail. *Cognitive Development* 21 (1): 3-10.
- Botterill, G & Carruthers, P. (1999). *The Philosophy of Psychology*. Cambridge University Press.
- Brooks, L. R. (1967). ‘The Suppression of Visualization by Reading’. *The Quarterly Journal of Experimental Psychology* 19 (4): 289-299.
- Buckwalter, W. & Tullmann, K. (forthcoming). The Genuine Attitude View of Fictional Belief. In Bradley H., Sullivan-Bissett E. & Noordhof P. (eds.), *Art and the Nature of Belief*. Oxford University Press.
- Budd, M. (1989). *Wittgenstein's Philosophy of Psychology*. Routledge.
- Budd, M. (1993). How Pictures Look. In Dudley Knowles and John Skorupski (eds.), *Virtue and Taste*, Blackwell. pp. 154-175.
- Camp, E. (2007). Thinking with maps. *Philosophical Perspectives*, 21(1): 145-182.
- Carroll, N. (1990). *The Philosophy of Horror: Or, Paradoxes of the Heart*. Routledge.
- Carruthers, P. (1996). Simulation and self-knowledge: A defence of the theory-theory. In Peter Carruthers & Peter K. Smith (eds.), *Theories of Theories of Mind*. Cambridge University Press. pp. 22-38.
- Carruthers, P. (2003). Review of Gregory Currie, Ian Ravenscroft, *Recreative Minds: Imagination in Philosophy and Psychology*. *Notre Dame Philosophical Reviews* 2003 (11).
- Carruthers, P. (2006) Why Pretend. In Shaun Nichols (ed.), *The Architecture of the Imagination*. Oxford University Press. pp. 89-109.
- Carruthers, P. (2009a). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences* 32 (2): 121.
- Carruthers, P. (2009b). Review: Simulation and the First-Person. *Philosophical Studies* 144 (3): 467-475.

- Chalmers, D. J. (2002). Does conceivability entail possibility? In Tamar S. Gendler & John Hawthorne (eds.), *Conceivability and Possibility*. Oxford University Press. pp. 145-200.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.
- Coleridge, S. T. (1984). *Biographia literari*. Princeton University Press.
- Cook, R. T. (2012). Why Comics Are Not Films: Metacomics and Medium-Specific Conventions. In Aaron Meskin, & Roy T. Cook (Eds.), *The Art of Comics: A Philosophical Approach*. Wiley-Blackwell. pp. 165-87.
- Cooke, B. (2014). Ethics and Fictive Imagining. *Journal of Aesthetics and Art Criticism* 72 (3): 317-327.
- Crane, T. (2009). Is perception a propositional attitude? *Philosophical Quarterly* 59 (236): 452-469.
- Crane, T. (2012). What is the Problem of Non-Existence? *Philosophia* 40 (3): 417-434.
- Crane, T. (2013). *The Objects of Thought*. Oxford University Press.
- Currie, G. (1990). *The Nature of Fiction*. Cambridge University Press.
- Currie, G. (1995). Imagination as simulation: Aesthetics meets cognitive science. In Martin Davies & Tony Stone (eds.), *Mental Simulation*. Blackwell.
- Currie, G. (1998). Pretence, pretending, and metarepresenting. *Mind and Language* 13 (1): 35-55.
- Currie, G. & Ravenscroft, I. (2002). *Recreative Minds: Imagination in Philosophy and Psychology*. Oxford University Press.
- Damasio, A. R., Tranel, D., and Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience*. 13: 89-109.
- Davidson, D. (1973). Radical interpretation. *Dialectica* 27 (1): 314-328.
- Davies, D. (2007). *Aesthetics and literature*. Continuum.
- Davies, S. (2009). Responding Emotionally to Fictions. *Journal of Aesthetics and Art Criticism* 67 (3): 269 - 284.
- Davis, J. F. (2015). *Imaginary Desires* (MPhil Stud dissertation, University College London).
- Dennett, D. C. (1987). *The Intentional Stance*. MIT Press.
- Descartes, R. (1642/1984). Meditations on First Philosophy. In *The Philosophical Writings of Descartes*, vol. 2, edited by J. Cottingham et al. Cambridge: Cambridge University Press.
- Doggett, T. & Egan, A. (2007). Wanting things you don't want: The case for an imaginative analogue of desire. *Philosophers' Imprint* 7 (9): 1-17.
- Doggett, T. & Egan, A. (2012). How We Feel About Terrible, Non-existent Mafiosi. *Philosophy and Phenomenological Research* 84 (2): 277-306.
- Dummett, M. (1993). Could There Be Unicorns? In *The Seas of Language*. Oxford: Clarendon Press. pp. 328-348.

- Egan, A. (2008). Imagination, delusion, and self-deception. In Tim Bayne & Jordi Fernandez (eds.), *Delusion and Self-Deception: Affective and Motivational Influences on Belief Formation*. Psychology Press. pp. 263-280.
- Estes, D., Wellman, H. M., & Woolley, J. D. (1989). Children's understanding of mental phenomena. *Advances in child development and behavior*, 22: 41-87.
- Evans, G. (1982). *The Varieties of Reference*. Oxford University Press.
- Everett, A. (2013). *The Nonexistent*. Oxford University Press.
- Ferreira, M. J. (2014). *A Pluralistic Account of Propositional Imagination* (Doctoral dissertation, The Ohio State University).
- Fitzpatrick, S. (2017). Good Communist Homes. [Review of the book *The House of Government: A Saga of the Russian Revolution*.] *London Review of Books*, 39 (15): 3-7. Retrieved from <https://www.lrb.co.uk/v39/n15/sheila-fitzpatrick/good-communist-homes>.
- Frege, G. (1956). The thought: A logical inquiry. *Mind* 65 (259): 289-311.
- Frege, G. (1948). Sense and reference. *The Philosophical Review*, 57(3): 209-230.
- Friedman, O & Leslie, A. M. (2007). The conceptual underpinnings of pretense: Pretending is not 'behaving-as-if'. *Cognition* 105 (1): 103-124.
- Friedman, O; Neary, K. R.; Burnstein, C. L. & Leslie, Alan M. (2010). Is young children's recognition of pretense metarepresentational or merely behavioral? Evidence from 2- and 3-year-olds' understanding of pretend sounds and speech. *Cognition* 115 (2): 314-319.
- Friedman, O. (2013). How do children represent pretend play? In *The Oxford Handbook of the Development of Imagination*, pp. 186-195.
- Friend, S. (2003). "How I really feel about JFK." In Matthew Kieran (ed.), *Imagination, Philosophy and the Arts*. Routledge. pp. 35-53.
- Friend, S. (2007). Fictional characters. *Philosophy Compass* 2 (2): 141-156.
- Friend, S. (2008). Imagining Fact and Fiction. In Kathleen Stock & Katherine Thomsen-Jones (eds.), *New Waves in Aesthetics*. Palgrave-Macmillan. pp. 150-169.
- Friend, S. (2011). II—Stacie Friend: Fictive Utterance and Imagining II. *Aristotelian Society Supplementary Volume* 85 (1): 163-180.
- Friend, S. (2012). Fiction as a Genre. *Proceedings of the Aristotelian Society* 112 (2): 179-209.
- Friend, S. (2016). Fiction and Emotion. In Amy Kind (ed.), *The Routledge Handbook of Philosophy of Imagination*. Routledge. pp. 217-229.
- Friend, S. (2017). The real foundation of fictional worlds. *Australasian Journal of Philosophy*, 95 (1): 29-42.
- Friend, S. (N.D.). The World of The Story (draft of chapter 4 of her forthcoming book on fiction).
- Funkhouser, E, and Shannon S (2009). Imagination and other scripts. *Philosophical Studies* 143 (3): 291-314.

- Gaiman, N. (2013) 'Police in Britain fired their guns just seven times in the last year', *The Guardian*. Available at: <https://www.theguardian.com/books/2013/oct/15/neil-gaiman-future-libraries-reading-daydreaming>
- Gendler, T. S. (2000). The puzzle of imaginative resistance. *Journal of Philosophy* 97 (2): 55-81.
- Gendler, T. S. (2006). Imaginative contagion. *Metaphilosophy*, 37(2): 183-203.
- Gendler, T. S. (2011). Imagination. In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. Center for the Study of Language and Information, Stanford University.
- Gendler, T. S. & Kovakovich, K (2006). Genuine Rational Fictional Emotions. In Matthew Kieran (ed.), *Contemporary Debates in Aesthetics and the Philosophy of Art*. Blackwell. pp. 241-253.
- Gendler, T. S. & Hawthorne, J. (2002) Introduction. In *Conceivability and Possibility*. Oxford University Press. pp. 1-70.
- German, T. P., & Leslie, A. M. (2001). Children's inferences from 'knowing' to 'pretending' and 'believing'. *British Journal of Developmental Psychology*, 19 (1): 59-83.
- Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*. Yale University Press.
- Goldie, P (2004). Review of *Recreative minds* by Currie and Ravenscroft. *Philosophy* 79: 331-335.
- Goldie, P (2005). Imagination and the distorting power of emotion. *Journal of Consciousness Studies* 12 (8-10): 127-139.
- Goldman, A. (1989). Interpretation psychologized. *Mind and Language* 4 (3): 161-85.
- Goldman, A. (2006a). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford University Press.
- Goldman, A. (2006b). Imagination and Simulation in Audience Responses to Fiction. In Shaun Nichols (ed.), *The Architecture of the Imagination: New Essays on Pretence, Possibility, and Fiction*. Oxford University Press. pp. 41-56.
- Goldman, A. (2008). Hurley on Simulation. *Philosophy and Phenomenological Research* 77 (3): 775-788.
- Gopnik, A. & Meltzoff, A. N. (1998). Theories vs. modules: To the Max and beyond: A reply to Dubois and to Stich and Nichols. *Mind and Language* 13 (3): 450-456.
- Gopnik, A. & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language* 7 (1-2): 145-71.
- Gopnik, A. (2009). *The philosophical baby: What children's minds tell us about truth, love & the meaning of life*. Random House.
- Gordon, R. M. & Barker, J. A. (1994). Autism and the "theory of mind" debate. In George Graham & G. Lynn Stephens (eds.), *Philosophical Psychopathology*. MIT Press. pp. 163-181.
- Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language* 1 (2): 158-71.

- Gould, R. (1972). *Child Studies through Fantasy*. Quadrangle Books.
- Graesser, A. C. and Forsyth, C. (2013). Discourse Comprehension. In *The Oxford Handbook of Cognitive Psychology*: Oxford University Press. Retrieved 13 Sep. 2017, from <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780195376746.001.0001/oxfordhb-9780195376746-e-30>.
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual review of psychology*, 48 (1): 163-189.
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of personality and social psychology* 79 (5): 701-721.
- Green, M. C., & Donahue, J. K. (2009). Simulated worlds: Transportation into narratives. In *Handbook of imagination and mental simulation*, pp. 241-256.
- Grice, H. P. (1957). Meaning. *Philosophical Review* 66 (3): 377-388.
- Grzankowski, A. (2012). Not All Attitudes are Propositional. *European Journal of Philosophy* 23 (3): 374-391.
- Harris, P. (2000). *The Work of the Imagination*. Blackwell.
- Harris, P. L., & Kavanaugh, R. D. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development* 58, 1-107.
- Harris, P. L., Brown, E., Marriott, C., Whittall, S., & Harmer, S. (1991). Monsters, ghosts and witches: Testing the limits of the fantasy—reality distinction in young children. *British Journal of Developmental Psychology*, 9 (1): 105-123.
- Heal, J. (1986). Functionalism and replication. In Jeremy Butterfield (ed.), *Language, Mind and Logic*. Cambridge University Press.
- Heal, J. (2003). *Mind, reason and imagination: selected essays in philosophy of mind and language*. Cambridge University Press.
- Hurley, S. (2005). The Shared Circuits Hypothesis: A Unified Functional Architecture for Control, Imitation, and Simulation, In Susan Hurley & Nick Chater (eds.), *Perspectives on Imitation: From Neuroscience to Social Science, Volume 1: Mechanisms of Imitation and Imitation in Animal*. MIT Press, pp. 177-193.
- Jackson, F. (1999). All that can be at issue in the theory-theory/simulation debate. *Philosophical Papers* 28 (2): 77-96.
- James, W. (1984) What is an emotion? *Mind* 34: 188-205.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.
- Kajtár, L (2017). Fiction cannot be true. *Philosophical Studies* 174 (9): 2167-2186.
- Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*. 17: 4302-4311.
- Kind, A. (2001). Putting the image back in imagination. *Philosophy and Phenomenological Research* 62 (1): 85-109.

- Kind, A. (2011) The puzzle of imaginative desire. *Australasian Journal of Philosophy* 89 (3): 421-439.
- Kind, A. (2013). The Heterogeneity of the Imagination. *Erkenntnis* 78 (1): 141-159.
- Kind, A. (2016). Introduction. In Amy Kind (ed.), *The Routledge Handbook of Philosophy of Imagination*. Routledge. pp. 1-11.
- Kripke, S. A. (1980). *Naming and Necessity*. Harvard University Press.
- Kripke, S. A. (2011). Vacuous Names and Fictional Entities. In *Philosophical Troubles. Collected Papers Vol I*. Oxford University Press.
- Kripke, S. A. (2013). *Reference and Existence. The John Locke Lectures*. Oxford University Press.
- Lamarque P, & Olsen S. H. (1994). *Truth, Fiction, and Literature: A Philosophical Perspective*. Oxford University Press.
- Lang, P. J., Levin, D. N., Miller, G. A., & Kozak, M. J. (1983). Fear behavior, fear imagery, and the psychophysiology of emotion: the problem of affective response integration. *Journal of abnormal psychology*, 92 (3): 276-306.
- Langland-Hassan, P. (2012). Pretense, imagination, and belief: the Single Attitude theory. *Philosophical Studies* 159 (2): 155-179.
- Langland-Hassan, P. (2014). What It Is to Pretend. *Pacific Philosophical Quarterly* 95 (1): 397-420.
- Langland-Hassan, P. (2016). On Choosing What to Imagine. In *Knowledge Through Imagination*. Oxford University Press. pp. 61-84.
- Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American psychologist*, 46 (8): 819.
- Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind.". *Psychological Review* 94 (4): 412-426.
- Leslie, A. M. (1994). Pretending and believing: issues in the theory of ToMM. *Cognition* 50 (1-3): 211-238.
- Leslie, A. M. (2002). Pretense and representation revisited. *Representation, memory, and development: Essays in honor of Jean Mandler*. Psychology Press. pp. 103-114.
- Leslie, A. M. & Happé, F. (1989). Autism and ostensive communication: The relevance of metarepresentation. *Development and psychopathology*, 1 (3): 205-212.
- Leslie, A. M. & Scholl, B. J. (1999). Modularity, development and 'theory of mind'. *Mind and Language* 14 (1): 131-153.
- Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, 43 (3): 225-251.
- Lewis, D. (1983). Postscript to truth in fiction. In *Philosophical Papers*. Oxford University Press, pp. 276-280.
- Liao, Shen-yi & Doggett, T. (2014). The Imagination Box. *Journal of Philosophy* 111 (5): 259-275.

- Liao, Shen-yi & Gendler, T. S. (2011). Pretense and Imagination. *Wiley Interdisciplinary Reviews* 2 (1): 79-94.
- Lillard, A. S. (1993). Young children's conceptualization of pretense: Action or mental representational state? *Child Development*, 64 (2): 372-386.
- Livingston, P. (2005). *Art and Intention: A Philosophical Study*. Oxford University Press.
- Martin, M. G. F. (2002). The transparency of experience. *Mind and Language* 4 (4): 376-425.
- Matravers, D. (1991). Who's afraid of Virginia Woolf? *Ratio* 4 (1): 25-37.
- Matravers, D. (2010). Why We Should Give Up on the Imagination. *Midwest Studies in Philosophy* 34 (1): 190-199.
- Matravers, D. (2013). *Fiction and Narrative*. Oxford University Press.
- McGinn, C. (2005). *Mindsight: Image, Dream, Meaning*. Harvard University Press.
- Meskin, A. & Robson, J. (2012). Fiction and Fictional Worlds in Videogames. In J. R. Sageng, T. M. Larsen & H. Fossheim (eds.), *The Philosophy of Computer Games*. Springer. pp. 201-18.
- Meskin, A. & Weinberg, J. M. (2003). Emotions, fiction, and cognitive architecture. *British Journal of Aesthetics* 43 (1): 18-34.
- Moran, R. (1994). The expression of feeling in imagination. *The Philosophical Review* 103: 75-106.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review* 83: 435-50.
- Neill, A. (1993). Fiction and the Emotions. *American Philosophical Quarterly* 30 (1): 1-13.
- Nichols, S, Stich, S. P., Leslie, A, M. & Klein, D. B. (1996). Varieties of off-line simulation. In Peter Carruthers & Peter K. Smith (eds.), Cambridge University Press. pp. 39-74.
- Nichols, S. (2004). Imagining and believing: The promise of a single code. *Journal of Aesthetics and Art Criticism* 62 (2): 129-39.
- Nichols, S. (2004b). Review of Currie and Ravenscroft, recreative minds. *Mind*, 113 (450): 329-334.
- Nichols, S. (2006). Just the imagination: Why imagining doesn't behave like believing. *Mind and Language* 21 (4): 459-474.
- Nichols, S. & Stich, S. P. (2003). *Mindreading. An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford University Press.
- Nussbaum, M. C. (1983). Flawed crystals: James's The Golden Bowl and literature as moral philosophy. *New Literary History*, 15 (1): 25-50.
- O'Brien, L. (2005) Imagination and the motivational view of belief. *Analysis* 65: 55-62.
- O'Craven, K. M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience* 12: 1013-1023.
- O'Shaughnessy, B. (2000). *Consciousness and the World*. Oxford University Press.

- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308: 255-258.
- Peacocke, C. (1987). Depiction. *Philosophical Review* 96 (3): 383-410.
- Picciuto, E. R. (2013). *The experience of fiction*. (Doctoral dissertation, University of Maryland, College Park).
- Potts, G. R., John, M. F. S., & Kirson, D. (1989). Incorporating new information into existing world knowledge. *Cognitive Psychology*, 21(3): 303-333.
- Prentice, D. A. and R. J. Gerrig (1999). 'Exploring the Boundary between Fiction and Reality'. In S. Chaiken and Y. Trope (eds.). *Dual Process Theories in Social Psychology*. Guilford Press, pp. 529-546.
- Prior, A. N. (1971). *Objects of Thought*. Clarendon Press.
- Radford, C. & Weston, M. (1975). How Can We Be Moved by the Fate of Anna Karenina? *Aristotelian Society Supplementary Volume* 49 (1): 67-93.
- Ravenscroft, I, Folk Psychology as a Theory, *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2016/entries/folkpsych-theory/>>.
- Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14- and 18-month-olds. *Developmental psychology*, 33 (1): 12-21.
- Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the brain: How our minds share actions and emotions*. Oxford University Press.
- Robinson, J. (1995). L'éducation sentimentale. *Australasian Journal of Philosophy* 73 (2): 212 – 226.
- Russell, B. (1905a). On Denoting, *Mind*, 14: 473-493.
- Russell, B. (1905b). Critical Notice of: A. Meinong, Untersuchungen zur Gegenstandstheorie und Psychologie, *Mind*, 14: 530-538.
- Savile, A., & Wollheim, R. (1986). Imagination and pictorial understanding. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 60: 19-60.
- Schaper, E. (1978). Fiction and the suspension of disbelief. *British Journal of Aesthetics* 18 (1): 31-44.
- Schellenberg, S. (2013). Belief and Desire in Imagination and Immersion. *Journal of Philosophy* 110 (9): 497-517.
- Schroeder, T & Matheson, C. (2006). Imagination and Emotion. In Shaun Nichols (ed.), *The Architecture of the Imagination: New Essays on Pretence, Possibility, and Fiction*. Oxford University Press. pp. 19-39.
- Schwitzgebel, E, Belief, *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2015/entries/belief/>>.
- Scruton, R. (1974). *Art and Imagination: A Study in the Philosophy of Mind*. St. Augustine's Press.

- Shah, N. & Velleman, D. J. (2005). Doxastic deliberation. *Philosophical Review* 114 (4): 497-534.
- Sinhababu, N. (2013). Distinguishing belief and imagination. *Pacific Philosophical Quarterly*, 94 (2): 152-165.
- Sinhababu, N. (2016). Imagination and Belief. In Amy Kind (ed.), *The Routledge Handbook of Philosophy of Imagination*. pp. 111-123.
- Skolnick Weisberg, D., & Bloom, P. (2009). Young children separate multiple pretend worlds. *Developmental Science*, 12(5): 699-705.
- Skolnick, D & Bloom, P. (2006). The Intuitive Cosmology of Fictional Worlds. In Shaun Nichols (ed.), *The Architecture of the Imagination: New Essays on Pretence, Possibility, and Fiction*. Oxford University Press. pp. 73-86
- Smith, M. (1995). Film spectatorship and the institution of fiction. *Journal of Aesthetics and Art Criticism* 53 (2): 113-127.
- Smuts, A. (2009). Art and negative affect. *Philosophy Compass* 4 (1): 39-55.
- Sparks, J. R., & Rapp, D. N. (2010). Discourse processing—examining our everyday language experiences. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(3): 371-381.
- Spaulding, S. (2012). Mirror neurons are not evidence for the Simulation Theory. *Synthese* 189 (3): 515-534.
- Stevenson, L. F. (2003). Twelve conceptions of imagination. *British Journal of Aesthetics* 43 (3): 238-59.
- Stich, S & Nichols, S. (2000). A cognitive theory of pretense. *Cognition* 74 (2): 115-147.
- Stich, S & Tarzia, J. (2015). The pretense debate. *Cognition* 143: 1-12.
- Stock, K. (2011a). Cognitive Theory of Imagination and Aesthetics. In Elisabeth Schellekens & Peter Goldie (eds.), *The Aesthetic Mind: Philosophy and Psychology*. Oxford University Press. pp. 268-282.
- Stock, K. (2011b). I–Kathleen Stock: Fictive Utterance and Imagining. *Aristotelian Society Supplementary Volume* 85 (1): 145-161.
- Stock, K. (2016). Imagination and Fiction. In Amy Kind (ed.), *The Routledge Handbook of Philosophy of Imagination*. Routledge. pp. 204-216.
- Strawson, P. F. (1970). *Meaning and Truth: An Inaugural Lecture Delivered Before the University of Oxford on 5 November 1969*. Clarendon Press.
- Sutrop, M. (2002). Imagination and the act of fiction-making. *Australasian Journal of Philosophy* 80 (3): 332 – 344.
- Tavinor, G. (2005). Videogames and interactive fiction. *Philosophy and Literature* 29 (1): 24-40.
- Tavinor, G. (2009). *The Art of Videogames*. Wiley-Blackwell.
- Taylor, M. (1998). *Imaginary companions and the children who create them*. Oxford University Press.
- Thomasson, A. L. (1999). *Fiction and Metaphysics*. Cambridge University Press.

- Tullmann, K. (2015) *Problems, Puzzles, and Paradoxes for a Moral Psychology of Fiction*. (Doctoral Dissertation, City University of New York).
- Tullmann, K. (2016). Taking the fictional stance. *Inquiry: An Interdisciplinary Journal of Philosophy* 59 (6): 766-792.
- van Inwagen, Peter (1977). Creatures of Fiction. *American Philosophical Quarterly* 14 (4): 299 - 308.
- Van Leeuwen, Neil (2011). Imagination is where the Action is. *Journal of Philosophy* 108 (2): 55-77.
- Van Leeuwen, Neil (2013). The Meanings of "Imagine" Part I: Constructive Imagination. *Philosophy Compass* 8 (3): 220-230.
- Van Leeuwen, Neil (2014). The Meanings of "Imagine" Part II: Attitude and Action. *Philosophy Compass* 9 (11): 791-802.
- Velleman, J.D. (2000) On The Aim Of Belief. In *The Possibility of Practical Reason*. Oxford University Press. pp. 244-282.
- Walton, K. L. (1970). Categories of art. *Philosophical Review* 79 (3): 334-367.
- Walton, K. L. (1973). Pictures and make-believe. *Philosophical Review* 82 (3): 283-319.
- Walton, K. L. (1978/2015). Fearing fictions. Reprinted *In Other Shoes: Music, Metaphor, Empathy, Existence*. Oxford University Press. pp. 252-272.
- Walton, K. L. (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Harvard University Press.
- Walton, K. L. (1997/2015) Spelunking, Simulation, and Slime: On being moved by fiction. Reprinted in *In Other Shoes: Music, Metaphor, Empathy, Existence*. Oxford University Press. pp. 273-287.
- Walton, K. L. (2006/2008) On The (So-Called) Puzzle of Imaginative Resistance. In *Marvellous Images: On Values and the Arts*. Oxford University Press. pp. 47-59.
- Walton, K. L. (2008). Pictures and Hobby Horses: Make-Believe beyond Childhood. In *Marvellous Images: On Values and the Arts*. Oxford University Press. pp. 63-78.
- Walton, K. L. (2015) Fictionality and Imagination Reconsidered. In *In Other Shoes: Music, Metaphor, Empathy, Existence*. Oxford University Press. pp. 17-35.
- Walton, K. L. (1994). Morals in Fiction and Fictional Morality. *Aristotelian Society Supplementary Volume* 68 (1): 27-66.
- Weinberg, J. M. (2013). All Your Desires in One Box. In Greg Currie, Matthew Kieran, Aaron Meskin & Jon Robson (eds.), *Aesthetics and the Sciences of Mind*: Oxford University Press. pp. 181-204.
- Weinberg, J. M. & Meskin, A. (2006a). Imagine that! In Matthew Kieran (ed.), *Contemporary Debates in Aesthetics and the Philosophy of Art*. Blackwell. pp. 222-235.

- Weinberg, J. M. & Meskin, A. (2006b). Puzzling over the imagination: Philosophical problems, architectural solutions. In Shaun Nichols (ed.), *The Architecture of the Imagination: New Essays on Pretence, Possibility, and Fiction*. Oxford University Press, pp. 175-202.
- Weisberg, D. S. (2015). Pretend play. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(3): 249-261.
- Wellman, H. M., & Estes, D. (1986). Early understanding of mental entities: A re-examination of childhood realism. *Child development* 57 (4): 910-923.
- White, A. R. (1990). *The Language of Imagination*. Cambridge: Blackwell.
- Williamson, T. (2005). I-Armchair philosophy, metaphysical modality and counterfactual thinking. *Proceedings of the Aristotelian Society* 105 (1): 1-23.
- Williamson, T. (2016). Knowing by Imagining. In *Knowledge Through Imagination*. Oxford University Press. pp. 113-123.
- Wiltsher, N & Meskin, A. (2016). Art and Imagination. In Amy Kind (ed.), *The Routledge Handbook of the Philosophy of Imagination*. Routledge. pp. 179-191.
- Wollheim, R. (1973). *On Art and The Mind*. Harvard University Press.
- Wollheim, R. (1980). *Art and its Objects*. Cambridge University Press.
- Wollheim, R. (1991). A Note on Mimesis as Make-Believe. *Philosophy and Phenomenological Research* 51 (2): 401--406.
- Wollheim, R. (1998). On pictorial representation. *Journal of Aesthetics and Art Criticism* 56 (3): 217-226.
- Woodward, R. (2012). Fictionalism and Incompleteness. *Noûs* 46 (4): 781-790.
- Yablo, S. (1993). Is conceivability a guide to possibility? *Philosophy and Phenomenological Research* 53 (1): 1-42.
- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological science*, 6 (5): 292-297.