



The  
University  
Of  
Sheffield.

# Feeling Like Stories

Empathy and the Narrative Perspective

By:

Alex Baker-Graham

A thesis submitted in partial fulfilment of the requirements for the  
degree of

Doctor of Philosophy

The University of Sheffield

Faculty of Arts and Humanities

Department of Philosophy

May, 2016

## ABSTRACT

In this thesis I present and defend a theory of empathy, and then apply that theory of empathy to understanding how we engage with stories.

I argue that empathy should be understood as a well-grounded demonstrative ascription of the form '[the target] feels like *this*'. I take the well-groundedness of such an ascription to consist in a series of 'proto-empathic' imaginings, which justify our ascription to a target by virtue of being *congruent* with one another. In laying out my conception of empathy I argue against several prominent theories of empathy, including those favoured by Preston and de Waal and Alvin Goldman. I argue in particular against the idea that empathy should be understood as aiming primarily at a matching of affect between an empathiser and their target.

Moving on to narrative engagement, I argue that when audiences engage with stories they empathise with an implied narrator of that story. I make this case by showing how empathy can *prima facie* be employed to solve two outstanding philosophical problems about stories by virtue of its employment of perspective shifting. I sketch a conception of 'perspectives' and go on to argue that every story features what I call a 'narrative perspective', and by process of elimination conclude that the holder of the narrative perspective must be an implied narrating agency. I then show how an empathic theory of narrative engagement can help us understand how stories can help or hinder our moral education.

Finally, I outline a theory of how audiences engage with interactive artworks such as videogames, drawing out the consequences of that view for how we might apply my theory of empathic engagement to furthering the understanding of interactive art.

## ACKNOWLEDGEMENTS

This thesis would not have been possible without the profound cleverness, boundless patience and frankly unreasonable levels of compassion and support of too many people to name. Anything of value in this thesis is due to the virtues of others, and any faults entirely of my own making. Thanks is due to the AHRC for funding this project, and I hope that many more students continue to benefit from their generosity. I also hope that some of those future students may find my work useful. The Department of Philosophy at Sheffield University is, my implicit bias aside, the best place to study in the country. To my friends, students, teachers and colleagues, you made it more fun than it had any right to be.

Rob Hopkins' peerless attention to detail and uncanny ability to simultaneously pick out and suggest solutions to the innumerable flaws of this thesis at the various stages of its life is the principal reason that I managed to corral my ideas into anything resembling cogent philosophical work. It was a privilege to be his student, and I only regret that I did not take fuller advantage of the immeasurable insight that he offered.

Dominic Gregory has my eternal gratitude for his friendly encouragement, invaluable advice on writing things in such a way as to make it possible to read them, and willingness to let me discuss my ideas until they made some kind of sense. Komarine Romdenh-Romluc offered fantastic advice on my views on empathy at a very late stage, without which the first part of this thesis would have been dramatically worse off. I also want to thank Catharine Abell for first making the philosophy of art so maddeningly interesting to me, and also for doing her level best to talk me out of working for a PhD. I'm sure she is gratified that I now understand her advice more fully, and I'm sure she understands why I'm glad I didn't take it. This work is also dedicated to the memory of Peter Goldie, who taught me much more than I realised at the time, about how good philosophy can be joyful in its rigor, and beautiful in its clarity.

To my Mum, who raised me to be the sort of feckless oddball who would think that a philosophy PhD would be a good use of his time, there will never be enough thanks. To my Dad, who made it possible for me to study at all, I owe you more than you will ever know. To Gemma, you are the ally that helped me triumph against my own worst enemy. Thank you.

## CONTENTS

|   |    |
|---|----|
| Abstract .....  | 2  |
| Acknowledgements.....   | 3  |
| Introduction.....   | 6  |
| Part I: Empathy, Emotion and Understanding.....                           | 9  |
| 1. Theories of Empathy.....   | 11 |
| 1.1 Introduction: Conceptions of Empathy.....                             | 11 |
| 1.2 Folk-Empathy and Sympathy .....                                       | 20 |
| 1.3 Theories of Empathy: Higher and Lower.....                            | 33 |
| 1.4 Theories of Empathy: Broad and Narrow .....                           | 36 |
| 1.5 Affective Matching and the History of Empathy.....                    | 39 |
| 1.6 Regularity, Reliability and Usefulness .....                          | 46 |
| 2. The Complex Theory of Empathy .....                                    | 56 |
| 2.1 Introduction: What we need from a New Theory.....                     | 56 |
| 2.2 Making Reconstructive Empathy Reliable.....                           | 61 |
| 2.3 Making Automatic Empathy Useful.....                                  | 66 |
| 2.4 Congruity and proto-empathy .....                                     | 70 |
| 3. Conclusion .....   | 74 |
| 3.1 The Complex Theory of Empathy: Applications and<br>Implications ..... | 77 |
| 3.1.1 The Moral Value of Complex Empathy .....                            | 77 |
| 3.1.2 The Complex Theory, The 'Folk' and Empathy Research .....           | 79 |
| Part II: Empathy and Narrative Engagement .....                           | 82 |
| 4. Perspectives And Narratives.....                                       | 83 |
| 4.1 Introduction: Proto-Empathies and Narrative Engagement....            | 83 |
| 4.2 Being-With and the Self-Other Distinction .....                       | 85 |

|  |     |
|--|-----|
| 4.3 Perspectives.....  | 90  |
| 4.4 Perspectives and the Philosophy of Literature: Imaginative<br>Resistance as Transmission Failure ..... | 98  |
| 4.5 Perspectives and the Philosophy of Literature: Disparate<br>Response .....                             | 114 |
| 5. Empathy and the Narrative Perspective .....   | 126 |
| 5.1 Locating the Narrative Perspective .....   | 126 |
| 5.2 Demonstrative Ascription in Narrative Engagement .....   | 150 |
| 6. Empathy and Narrative Engagement: Application and Expansion   | 155 |
| 6.1 Empathic Engagement and Moral Persuasion.....  | 155 |
| 7. Videogames, Empathy and the Role of the Player .....  | 167 |
| 7.1 The Extra Credits View.....  | 168 |
| 7.2 The Role of the Player.....  | 174 |
| 7.3 The Fridge Magnet Poetry Problem.....  | 183 |
| 7.4 Videogames as Artworks .....   | 188 |
| 7.5 Videogames and Empathic Engagement .....   | 190 |
| 8. Conclusion .....  | 192 |
| References .....   | 195 |

## INTRODUCTION

The aim of this project is to give a theory of empathy that is both plausible on its own terms, and useful in forming the basis of a theory of narrative engagement.

**Part I – Empathy, Emotion and Understanding** offers a critical analysis of contemporary theories of empathy and outlines my own suggested view, which I call the ‘Complex Theory of Empathy’.

**Chapter 1** discusses the place of empathy in modern ‘folk’ philosophy and psychology, and finds that many of the disagreements apparent between the ‘folk’ regarding the nature of empathy are shared to some degree by the refined theories offered by empathy researchers. I offer a definition of sympathy as a starting point, noting how it is commonly conflated with empathy in common usage. I then offer some criticisms of the current most popular theories of empathy. These include the ‘Perception-Action Model’ due to Preston and de Waal, who advocate a broad definition of empathy encompassing several observed psychological processes, and the ‘Two Routes View’, due to Alvin Goldman, but widely endorsed, which posits that empathy, principally conceived of in terms of affective matching, can be achieved in two distinct ways. I argue that the Two Routes View fails because the conditions on each route necessary to secure understanding of one’s empathic target make the other route impossible to use.

In **Chapter 2** I spell out the Complex Theory of Empathy, so named because it involves a complex of imaginative activity as well as because it is a view that conceives of empathy as relatively cognitively demanding. According to the Complex Theory, empathy is the making of a well-grounded demonstrative Ascription to a target of the form ‘[the target] feels like *this*’. I detail the conditions that make such an ascription well-grounded and offer some thoughts on how congruence between imaginings can secure justification for an empathic ascription.

**Chapter 3** is relatively short, and concludes my thoughts on empathy specifically. I argue, contrary to most current views, that affective matching should not be considered the primary function or ultimate goal of empathy, and suggest some implications the Complex Theory has for the moral value of empathy and future empathy research.

**Part II – Empathy and Narrative Engagement** details a theory of empathic engagement with stories, according to which an audience engages with stories by empathising with an implied narrator.

**Chapter 4** begins this task by laying out how my argument will proceed, and offering an account of perspectives, which I conceive of as relatively stable sets of dispositions that generate ways of seeing, which is to say that perspectives lead us to characterise events and objects in particular ways. I show that perspectives are promising as a tool for understanding how we engage with stories by showing how they can form the basis of solutions to both the problem formerly known as ‘the puzzle of imaginative resistance’, and what I call the ‘problem of disparate response’, which is the problem of why we respond to events in stories differently than we think we would respond to the same events in other circumstances.

I devote most of **Chapter 5** to exploring the nature of the narrative perspective, by way of examining the different figures who feature in our engagement with stories and exploring how far the narrative perspective could plausibly be held by each. I finish the chapter with a statement of my theory of empathic engagement with stories.

**Chapter 6** comprises some thoughts about how my empathic theory of narrative engagement might impact on related questions around the philosophy of stories, in particular the issue of how we might understand the idea that stories contribute to our moral development. I criticise a proposal by Shen-Yi Liao which would, if it held, make the empathic theory of narrative engagement unproductive on this subject.

**Chapter 7** is something of an oddity in the context of the wider thesis. In this final chapter I offer the beginnings of what, I hope, will form a future research project, and examine the role of the audience of interactive stories, its consequences for the art-status of videogames and how we might apply a theory of empathic engagement to better understanding them.



PART I:  
EMPATHY, EMOTION AND UNDERSTANDING

*I wish that for just one time you could stand inside my shoes.  
You'd know what a drag it is to see you.*

– Bob Dylan



## 1. THEORIES OF EMPATHY

### 1.1 Introduction: Conceptions of Empathy

Empathy is currently enjoying a fascinating prominence in our culture. Articles are frequently published in the online and print media about how our personal capacity for empathy can be improved by, among innumerable other things, reading literature, using more social media, using less social media, meditating, and being poor. Such is the regard in which empathy is held that those articles that make recommendations on how our empathic skills can be improved rarely feel the need to make the case for an increased capacity for empathy being a good thing. An entirely separate raft of articles, however, assures us that increased empathy can help us become successful entrepreneurs, be better parents, provide better healthcare, popularise political parties and create effective advertising campaigns<sup>1</sup>.

As one might expect given empathy's apparent place in 'folk' philosophy as a moral panacea, the conception of empathy appealed to varies greatly between, and sometimes within, articles such as these. Empathy in popular discourse is variously conceived of as a capacity to feel as another person feels, a disposition to have morally appropriate responses to suffering, a way of understanding another person's emotional states, or some combination of these. The following are some definitions of empathy from the popular press:

---

<sup>1</sup> Those claims are found in Bury (2013), Brosterman (2015), Quenqua (2013), Kristof (2015), Reynolds (2015), McBride (2011), Ficarra (2011), Mulgan (2015) and Barrow (2015) respectively.

**E1:** “Empathizing with [someone] is feeling what they are feeling and acknowledging those feelings...it is the art of compassion and sensitivity, as well as the ability to give moral support.”<sup>2</sup>

**E2:** “Empathy is walking a mile in another's shoes, or trying to understand the feelings and perspectives of an individual. It isn't pity or kindness, and it certainly isn't treating someone as you would like to be treated.”<sup>3</sup>

**E3:** “Empathy means not just that we care about what another person is feeling. Empathy means that we understand and feel what they're feeling. It means hurting like they hurt.”<sup>4</sup>

**E4:** “Empathy involves being moved by another's experiences...[and/or] the act of correctly acknowledging the emotional state of another without experiencing that state oneself”<sup>5</sup>

These examples are instructive for a several reasons. They confirm that ‘the folk’ don’t agree on what empathy is and also indicate potential areas of disagreement. However, one of the things I am interested in exploring in this thesis is whether and how far our best theories of empathy can perform the functions that the ‘folk’ credit it with. With that in mind, the examples above, contradictory though they may be, also give us some idea of what a refined concept of empathy needs to look like in order to be recognisable as empathy under a common conception.

As far as the above popular definitions of empathy agree, I think it is fair to offer the following as an initial suggestion of a working Folk Theory of empathy:

---

<sup>2</sup> Burg (2011)

<sup>3</sup> Reynolds (2015)

<sup>4</sup> Barron (2015)

<sup>5</sup> Ficarra (2010) (*N.B.* The original article is unclear as to whether the first and second parts of this definition are intended to complement or contrast one another).

**Folk Theory of Empathy:** Empathy is the activity of coming to understand another person, in particular (though perhaps not only) their feelings in a way that makes distinctive use of one's own feelings.

While the suggested Folk Theory doesn't by itself serve as a satisfactory definition of empathy for the purposes of a philosophical enquiry, it is nonetheless important to keep the Folk Theory in mind when discussing the various conceptions of empathy proposed by philosophers and scientists. I take it that it is a desideratum for a theory of empathy (and indeed for any theory of a concept commonly deployed by the 'folk') that a refined theory should provide a picture of empathy recognisable to those familiar with the folk theory alone (or, if it is severely at odds with the Folk Theory, that the refined theory gives compelling reasons why it is so). A further desideratum of a Refined Theory of empathy is that it settles any confusion or disagreement between various strands of Folk Theory of empathy, as illustrated in E1-E4 above. I suggest that a Refined Theory of empathy should be able to shed light on the following questions raised by the individual definitions above:

**Q1.** Does empathy entail compassion, pity or sympathy for the target of the empathy? (E1, E2, E3)

**Q2.** Does empathy require that the empathiser feels the same as their target? If so, the same in what sense? (E3, E4)

**Q3.** What kind of understanding does empathy provide, and what capacities does that understanding grant the empathiser? Must it, for instance, allow the empathiser to respond appropriately, or to lend moral support to the target? (E1, E3)

**Q4.** How is empathy distinct from cognitive mind-reading (deploying a theory of mind to determine a target's mental state)? (E1, E4)

So how close are we to having a Refined Theory of empathy capable of answering the above questions? The world of academic research has devoted a great deal of energy to understanding the potential benefits

and applications of empathy, and discovering how we might be able to increase people's capacity for empathy (indeed, many of the public-facing articles I have cited are inspired by particular pieces of academic research). This being the case, it is remarkable that there remains significant disagreement about what precisely empathy is. In fact, many of the disagreements between strands of the Folk Theory of empathy also appear in some form in the academic literature. As a result, those who wish to contribute to our knowledge of how empathy can best be developed or employed are still in some danger of speaking at cross-purposes. Although there has been much rigorous work, especially from philosophers and cognitive scientists, devoted to developing a sound concept of empathy, there continues to be no standard view either across or within disciplines. As well as the debates that echo disagreements between the strands of Folk Theory, there remain outstanding questions particular to the academic research community.

Debates of the former kind derive from more general questions about how empathy should be defined, and the latter from more esoteric concerns about which psychological or neurological processes underlie or constitute empathy. One interesting feature of the former debate, about how empathy should be defined, is that empathy is a rather new psychological concept. Although contemporary understandings of empathy certainly owe something to the motivator of moral sentiments that Adam Smith and David Hume discussed under the name 'sympathy' in the 18<sup>th</sup> century, empathy as a distinct concept has its deepest roots in the 19<sup>th</sup> and early 20<sup>th</sup> century works of the likes of Theodor Lipps, Vernon Lee and Edith Stein.<sup>6</sup> The relative newcomer status of empathy as a philosophical concept can perhaps partially explain the dramatically different approaches to defining empathy that have been taken even in recent times.

---

<sup>6</sup> Hume (1751), Smith (1759)

Contemporary research tends to take one of two broadly different approaches to empathy, seeing empathy as primarily an epistemological or a psychological phenomenon.<sup>7</sup> Proponents of the epistemological approach to empathy conceive of empathy primarily as a route to or form of knowledge about another person.<sup>8</sup> The psychological approach to empathy is concerned principally with defining empathy as a particular kind of psychological or neurological process that allows people to share mental states with another person.<sup>9</sup> Neither of these approaches are exclusive; a view that empathy is a special form of knowledge about another person will need to accommodate plausible theories about the psychological and neurological processes that make that knowledge possible. Conversely, a view that empathy is a neurological process that allows one person to share the mental state of another will be strongest when it can allow that the neurological process in question can perform the epistemic functions that are commonly ascribed to empathy.

#### *1.1.2 Introducing the Complex Theory of Empathy*

I will be arguing for an epistemic conception of empathy; empathy as a way of gaining a special kind of knowledge about another person, the knowledge of what it's like to be another person in a particular situation. To be precise, I will be arguing for what I call the *Complex Theory* of empathy:

**Complex Theory of Empathy:** Empathy is the making of a well-grounded demonstrative ascription to a target of the form '[the target] feels like *this*'.

The details and motivations for the Complex Theory of Empathy will be discussed throughout the first two chapters of the thesis, but it will

---

<sup>7</sup> Smith (2015)

<sup>8</sup> *Ibid.*, Coplan (2011)

<sup>9</sup> Preston and de Waal (2002)

be helpful to first offer some clarification on what the Complex Theory involves. A 'target' is the object of empathy; typically the person with whom I am empathising.<sup>10</sup> When I say that empathy is a demonstrative ascription, I mean that it is a matter of ascribing a property on the basis of a demonstrative act, which for now we can think of, following Walton, as using one's own feelings as a *sample* of how the target feels (in a sense of 'feeling' that includes the broad phenomenal character what it's like to be that person, and not just emotions or affective states).<sup>11</sup> A demonstrative ascription is *well-grounded* when we are the various imaginative activities that we must undertake to justify that ascription are congruent with one another. The imaginative activities that we must undertake to justify empathies I name *proto-empathies*, and I will identify three that I believe are jointly sufficient to ground an empathic ascription: *feeling-with*, *being-with* and *doing-with*. Feeling-with is a way of imagining that I share the affective state of another, being-with is a way of imagining that I share the same character as another, and doing-with is a way of imagining that I am in the same situation as another.

In arguing for this position I will describe some popular approaches to empathy, both those in ordinary use and from empathy researchers, and offer some criticisms of them; in particular I will argue that they do not offer an empathiser sufficient understanding of their target to perform the minimal epistemic role empathy ought to play. I will also argue against the common view in modern empathy research, that empathy is primarily or solely a matter of affective matching, that is, of coming to share an emotional or other affective state with another person.

---

<sup>10</sup> Which is to say that sometimes the target of my empathy is not a person.

<sup>11</sup> Walton (2015). My view shares many similarities with Walton's, though there are significant differences between us in what we take to be sufficient grounding for an empathic ascription.



There is one other significant difference between the Complex Theory and most other contemporary views on empathy that it is worth pointing to before I proceed further. The first is that the Complex Theory does not view empathy as a ‘success term’ in the way many other views do.<sup>12</sup> This is largely due to my rejection of the centrality of affective matching in empathy. According to most views of empathy, empathy occurs when the subject and object share affective states, and where the affective states do not match there is no empathy.<sup>13</sup> I also take empathy to be a success term, but I attach the success to the well-groundedness of the empathic ascription rather than the similarity of feeling between the empathiser and her target. It could well be that this is not much more than a semantic distinction, but I think the point is interesting nonetheless.

Empathy is a way to gain a special kind of knowledge about another person, the knowledge of what it’s like to be them. I take it to be analogous with other ways of gaining knowledge, in that for a method of knowledge acquisition to successfully occur (for example, reading a book or conducting an experiment), one does not need to reach a justified true belief, but merely to have properly used the method in question. If I read a book by a usually reputable author, but it turns out the claims in her book were false, I may not have successfully gained knowledge, but I have successfully followed a method that generally produces it (no one would say that I had failed to read the book!). Similarly, to successfully empathise is to attempt to gain empathic knowledge of another person in the proper way. Empathy ought to usually result in knowledge (since any epistemic view of empathy needs to hold that it is a reliable way of gaining knowledge), but if my well-grounded demonstrative ascription turns out to be false, I take it to be in spite of a successful empathising, not to constitute a failure. It is more helpful, when thinking of empathy in

---

<sup>12</sup> Walton (2015), Smith (2015)

<sup>13</sup> E.g. Vignemont and Singer (2006)

epistemic terms, to consider empathy to have failed when the mechanisms that justify the beliefs that empathy produces have failed, and not when the beliefs that empathy produces turn out to be false.

### 1.1.3 Introducing psychological approaches to empathy

When we think of empathy in psychological and/or neurological terms, the broadest divide in current theories comes down to whether empathy should be thought of as a sub-personal, automatic process, a higher-level, cognitive process, or a disjunctive concept operating over both of these other options. Several observed psychological phenomena have been suggested as being identifiable with empathy; these are *Emotional Contagion*, *Low-Level Simulation*, *High-Level Simulation*, *Perspective Shifting* and *In-His-Shoes Imagining*.

*Emotional Contagion*: a phenomenon whereby the emotional state of one person is transferred to another who, by virtue of some perceptual process, automatically adopts the perceived emotion as their own and thus their own emotional state converges with that of the first person.<sup>14</sup>

*Low-Level Simulation*: A neurological process whereby the perception of a mental state in another causes, in the perceiver, the activation of neural systems (*mirroring systems*) that would activate if they were in possession of the mental state perceived, leading the perceiver to adopt a state isomorphic to that of the target. Low-level simulation can be used to recreate a class of mental states broader than emotion-states, including pain-states and action-planning states, which distinguishes it from emotional contagion.<sup>15</sup>

---

<sup>14</sup> Davies, S (2011); Dimberg (1982); Bavelas et al. (1987); Hatfield, Rapson, and Le (2009); and Hatfield, Cacioppo, and Rapson (1994).

<sup>15</sup> Goldman, A (2011)

*High-Level Simulation:* The creation in one's own mind of a selected mental state through the faculty of imagination. The mental state in question can be one identified in another person, but need not be.<sup>16</sup>

*In-His-Shoes Imagining:* An imaginative project in which we experientially and centrally imagine from the inside (imagine in various sense modalities, from the perspective of someone within the imagined scene) that we are in the situation of another, but not that we *are* that other.<sup>17</sup>

*Perspective Shifting:* An imaginative project in which we experientially and centrally imagine from the inside that we are both in the situation of another *and* that we are that other.<sup>18</sup>

Those who offer definitions of empathy tend to use some or all of the above processes as a way of underpinning it. Hence, the definitions of empathy one tends to find prefacing contemporary papers on the subject do not often stray far from the mould exemplified by the following, due to Hoffman, Iacobani and Vignemont and Singer respectively:

Hoffman: “[*Empathy is any emotional response involving*] any processes that make a person have feelings that are more congruent with another’s situation than with his own situation.”<sup>19</sup>

Iacobani: “*Empathy is commonly defined as the ability to understand and share the feelings of another*”<sup>20</sup>

---

<sup>16</sup> Goldman (2006) and (2011), Currie and Ravenscroft (2002).

<sup>17</sup> Some philosophers have used the phrase ‘in-his-shoes’ imagining to refer to what I have called ‘perspective shifting’ (e.g. Gaut 2006 p.118). I suppose the difference in usage can be put down to whether one takes the metaphor of the other person’s shoes to refer to only their situation or also to relevant aspects of their psyche. There is in any case obviously enough of a distinction between the two imaginative projects to warrant two labels. See Wollheim (1984) pp. 71-3 and *Art and the Mind* p. 54) and Ninan (2010) for usage similar to mine.

<sup>18</sup> Wollheim (1984) pp. 71-3, Goldie (1999), Coplan (2011)

<sup>19</sup> Hoffman (2000)

<sup>20</sup> Iacobani (2011)

and,

Vignemont and Singer: *“There is empathy if: (i) one is in an affective state; (ii) this state is isomorphic to another person’s affective state; (iii) this state is elicited by the observation or imagination of another person’s affective state; (iv) one knows that the other person is the source of one’s own affective state.”*<sup>21</sup>

There will be much to say about these accounts of empathy as we progress. The rest of this section will consist in a critical overview of how these different approaches to empathy can answer our questions Q1-Q4 above, with the dual purpose of introducing the different dominant approaches to empathy and the principal issues and concepts that will become recurring themes throughout the first half of this thesis.

## 1.2 Folk-Empathy and Sympathy

I want to begin by getting as clear an idea as possible of what Folk Theory of empathy we want to operate with before finding out which, if any, Refined Theory can be best mapped onto it. I want to clarify that the idea here is not that having something in common with a concept that closely reflects ordinary usage by itself makes any Refined Theory more plausible than any other. However, I take it that having something in common with a clear Folk Theory is nonetheless a beneficial feature of a candidate for a Refined Theory, and a Refined Theory which looked substantially different to our best Folk Theory would be under an obligation to explain why it was so.

Abstracting from the sample definitions of empathy from the popular press, I suggested the following as a rough outline of what a Folk Theory of empathy might look like:

---

<sup>21</sup> Vignemont and Singer (2006)

**Folk Theory of Empathy:** Empathy is the activity of coming to understand another person, in particular (though not only) their feelings in a way that makes distinctive use of one's own feelings.

As it stands, the Folk Theory fails to distinguish between two commonly conflated concepts in popular discourse: empathy and sympathy. Empathy is typically thought of as a way of understanding another person by sharing their emotions, sympathy as a way of feeling particular emotions about another person that indicate that the object of sympathy is in need. Clarifying the phenomenon of sympathy and distinguishing it from empathy is a good first step towards developing a Refined Theory of Empathy and addressing some of the questions raised in my discussion of the Folk Theory in the introduction; whether empathy entails sympathy and whether empathy entails feeling, in some sense, the same as the target of the empathising.

Although people do commonly confuse, conflate, or just synonymise empathy with sympathy, there are some clues even in ordinary language that the terms refer to different things (although these clues are not as straightforward as some have maintained). Expressions of sympathy are often reserved for situations in which the object of the sympathy are grieving or have experienced a loss (and, I would tentatively suggest, where the object of sympathy is blameless for their own situation). It's often suggested that empathy can be distinguished from sympathy in common usage by virtue of being applicable to a much broader variety of circumstances, both happy and not.

However, I do not think this distinction between common use of the terms holds up under scrutiny. Nor does another frequently suggested distinction, that sympathy can be commonly spoken of as a distinct emotion alongside others such as joy, jealousy and frustration whereas empathy is commonly held as a way of feeling other emotions, as

opposed to an emotion in its own right.<sup>22</sup> Against the former claim, it is rare to find people proclaiming their empathy with people whose situation would not merit sympathy; people are generally only liable to declare empathy with those who are having a bad time of things.

Against the latter, it is also very common to hear people speak of ‘feeling empathy’ in the same manner in which they would speak of feeling sympathy, or other more straightforward emotional states.

This is not to say that these distinctions do not manifest themselves between Refined Theories of empathy and sympathy, only that they are not reflected in common usage. Discussing these and other possible distinctions between sympathy and empathy will help us clarify both concepts shortly, but for now I want to draw attention to an interesting distinction that I believe *does* hold between the ordinary usage terms of empathy and sympathy. That is, that sympathy entails an attitude that the sympathetic object is pitiable. Pitiable, that is, in a way that invokes the idea of pity as demonstrated in high-handed judgement (in the way that Mr. T is habitually said to pity ‘fools’) more than pity as mere sorrow for the misfortunate of another. In *Jane Eyre*, Rochester describes the variety of pity I have in mind in the following way:

*"Pity, Jane, from some people is a noxious and insulting sort of tribute, which one is justified in hurtling back in the teeth of those who offer it; but that is a sort of pity native to callous, selfish hearts; it is a hybrid, egotistical pain at hearing of woes, crossed with ignorant contempt for those who have endured them."*

—Rochester, *Jane Eyre*

Charlotte Bronte, through Rochester, is nicely capturing the thought that pity is not merely a feeling that someone suffers, but also expresses a power relation between the pitier and pitied. The implied

---

<sup>22</sup> Both claims are maintained in Harrold (2000), Prinz (2010) and Gibson (2015)

superiority of the pitier explains why, as Rochester says, people often feel slighted at being offered pity. The one who pities affirms not just that they recognise the suffering of another, but also that they are in a superior position. Indeed, often the power imbalance implicit in pity is created merely by the declaration of pity, as the pitied person becomes aware that the pitying person considers them vulnerable.<sup>23</sup> The power imbalance implicit in a declaration of pity also features in the ordinary use concept of sympathy, demonstrated by the fact that people are often insulted to be on the receiving end of sympathy in just the same way as pity, and the fact that those with whom people sympathise are often in an inferior position to those who do the sympathising.

Pity is the emotion that we most often associate with a sympathetic state, but pity and sympathy are not coextensive, even in common usage. It may be an unusual occurrence, but we can still recognise pity as pity when it is directed towards somebody for whom one has no personal concern or affection, whereas sympathy generally requires concern at least. Although this sense of pity has fallen out of style, it is perfectly coherent to suggest that you can feel pity towards one who has wronged you, and whom you freely, rightly and happily despise, but nonetheless pity out of a concern for, among other things, justice, the pleasure of another, your own sense of superiority or the security of your own salvation as reckoned against the damnation of another. A feeling of pity does not imply that the subject does not wish suffering on the pitied object (as in the King who pities the subjects he sentences to death), nor does it come packaged with a desire that that person's condition improve (as in cases where the pitied object's condition cannot improve).

Such feelings, I contend, do not capture our ordinary sense of 'sympathy', and so while pity might be the canonical sympathetic

---

<sup>23</sup> Stramondo (2010)

emotion, and lend to sympathy the power imbalance that so often drives resentment by those with whom we sympathise, sympathy clearly involves more than only pity. I will go on to give a fuller account of sympathy in the next section, but the preceding points give us a good starting point for that discussion, in that they mark a distinction between empathy and sympathy reflected in common usage far more reliably than the other two that were suggested (that empathy is thought of as applicable to a broader variety of circumstances than empathy, and that sympathy is commonly conceived of as an emotion, whereas empathy is not).

To summarise, then, when commonly speaking of sympathy we think of it as being given from a position of relative superiority (whether superiority of social position, assumed moral superiority, or simply a greater degree of security and comfort), whereas empathy need not be. The one who sympathises is most often having a better time of things than the one with whom they sympathise. The same is not true of empathy; empathy is commonly conceived of as, among other things, a way for people to demonstrate that they are sharing in another's pain. Sympathy, on the other hand, is not commonly conceived as a matter of sharing, but of offering solace, support, or even impersonal disapproval of circumstances, from a position of relative security.

#### *1.2.1 'Feeling' sympathy and empathy*

Now that I have suggested one way in which the ordinary usage of the terms 'sympathy' and 'empathy' can be distinguished from one another, I want to briefly discuss another common confusion in ordinary discourse which occasionally appears in empathy research. That is, that sympathy and empathy are feelings or emotions (or clusters of feelings or emotions) of the type we would usually recognise alongside fear, pity, joy and so on. In what follows I will argue that if we are to preserve the common notion that affective



matching is necessary for empathy,<sup>24</sup> we must dispense with the less common, but still present, idea that empathy names one or a disjunction of several distinct emotions.

If Avon claims that he “feels empathy with Omar”, what exactly should we take him to mean by this? If Omar had recently lost his job and was angry about it, Avon might (in line with some common usage) be conveying that he has a feeling consistent with an overriding interest in Omar’s wellbeing, i.e. a feeling of concern. Despite the fact that there are both laypersons and empathy researchers who would be happy to call such concern a state of empathy<sup>25</sup> there are good reasons to resist such a commitment.

In fact, it might be tempting to describe Avon’s state in such a case as an example of *feeling sympathy* rather than empathy. This is more correct than the claim that Avon feels empathy, but it is by no means a completely accurate picture of what is going on here, and we must be precise.

In what sense is the claim that Avon is feeling sympathy incorrect? It is for the very same reason that Avon is not feeling empathy, in that neither empathy nor sympathy are, properly speaking, the kinds of thing one feels; they are not kinds of feelings that we can distinguish by what they feel like. To explain what I mean, let us suppose that Avon and Omar are both angry at Omar’s unemployment. This, some might say, is sufficient for Avon to be feeling empathy, but it is a simple matter to show that this is not true. Omar and Avon are both feeling sad, and the quality of Omar’s feeling as the target – the affective component of Omar’s ‘sadness’ state – exhausts the affective features of the so-called feeling of empathy that Avon has alongside him. The point is this: the supposition is that Omar is *feeling* sad, but that Avon is *feeling* ‘empathy’, and consequentially that empathy is the

---

<sup>24</sup> As seen in the definitions offered by Jacobani and Vignemont and Singer in section 1.1.

<sup>25</sup> For instance, Davis (1980) and Batson (2011)

kind of thing that can be *felt*, can only be justified if there is some relevant qualitative difference in Omar's feeling from Avon's. What does Avon feel over and above sadness that marks his feeling as a feeling of empathy with Omar? It is difficult to even know what to make of this question.

To build a positive response we might look at the problem a different way. Suppose that Omar is again sad, and Avon feels sadness alongside Omar so that their states are isomorphic in the relevant ways. Further suppose that Lester is triumphant, and Jimmy feels triumphant with Lester so that their states are also isomorphic in the relevant ways. What are those relevant ways? I will cover the subject of what isomorphism should or must amount to for these purposes at a later stage. I think we can assume, however, that they would at least have to include certain perceptible affective properties of the feeling. For now, it will suffice if Avon's feeling is of the same kind (that we would ordinarily recognise; happy, sad, angry, jealous, petulant etc.) as Omar's, and that Lester's is of the same kind as Jimmy's. Questions of whether this isomorphism ought to involve sameness of intensity, duration, intentional object or the like I will leave open for now.

So, assume following the above that Avon feels sadness with Omar and Jimmy feels triumph with Lester. We know that there is no affective component of Avon's or Jimmy's states that marks them out as empathic. So what else could Avon's empathy Omar have in common with Jimmy's empathy with Lester? What is the common feature that makes it the case that Avon and Jimmy are both empathising?

The intuition that empathy is a feeling might be defended by claiming that when Avon feels anger with Omar, there is a feeling of concern for Omar's wellbeing in concert with that feeling of anger that marks out Avon's feeling as one of empathy, rather than simple sadness. This cannot be right either, in the first place because it is not plausible that we should require a feeling of concern to infringe upon Avon's sadness

in order for it to count as empathy. It is perfectly conceivable that, although Avon's sadness might have been partly caused by a concern for Omar's wellbeing, that concern need not be occurrently felt alongside, or somehow mixed in with, Avon's sadness. For further evidence that empathy cannot simply mean a 'concern-involving emotion', consider a different case in which Omar is feeling concerned rather than sad that he has lost his job. Avon's empathy for Omar here will manifest in a feeling of concern, the affective component of which will itself exhaust the affective component of Avon's concern-state, leaving no affective work for the empathy to do. While it could be true that such compounded concern feelings might result in a concern feeling of unusual intensity, the resulting affect could not be reliably distinctive in kind. Not only this, but in fact to claim that empathy is something Avon *feels* would deny the possibility of empathy occurring in its most complete and successful form from the outset. If empathy involves Avon feeling, in some suitably defined sense,<sup>26</sup> the same thing as what Omar is feeling, then the claim that empathy itself carries some affective character implies that Avon could never feel Omar's sadness in empathy – he could only feel sadness-plus-empathy, which is certainly not what Omar is feeling. In summary: what Avon feels by empathising with Omar is only that state which Omar feels.

It follows from the above that we are mistaken if we say that Avon feels sympathy for when he feels concern any more than that he feels empathy when he feels sadness. Just as his feeling of anger exhausts the affective component of the so-called empathy felt by Avon, so his concern exhausts the affective component of Avon's 'sympathy'. The

---

<sup>26</sup> It might be claimed that the 'suitably defined sense' in which Avon feels the same as Omar *is just* Omar's emotion-plus-empathy. However, my previous discussion has shown that empathy cannot be isolated by its affect alone, so I do not believe such an assertion can be seriously justified.

confusion apparent in ordinary language, giving rise to the need for the preceding discussion, stems from the habit of using the term ‘sympathy’ to describe a form of pity, which is indeed a feeling of the kind we would usually recognise.<sup>27</sup> As we have just seen, however, the terms ‘sympathy’ and ‘empathy’ are explanatorily impotent if they are deployed to describe the feelings felt as a part of an affective state – the affective component of my pity for one with whom I sympathise can be exhaustively described without reference to a feeling of sympathy. This is not because sympathy and empathy are unanalysable terms; they are not. Rather, it is because empathy and sympathy have no distinctive kind of affective state with which they are associated. The same argument applies to any claim that empathy or sympathy name a cluster of emotions, such as the ‘constellation’ of emotions such as sadness, compassion, tenderness, upset, concern and grief, which have been suggested to disjunctively constitute ‘empathic concern’.<sup>28</sup> For any instance of sympathy or empathy, it isn’t the type of emotion that marks the experience as sympathetic or empathic.

### *1.2.2 Sympathy as Feeling-for*

We have now established that there is a difference in the ordinary language application of ‘sympathy’ and ‘empathy’, in that they imply different power relationships between the subject and object, and that neither sympathy nor empathy should be identified with a particular feeling or emotion, nor a cluster of feelings or emotions.

So what does sympathy consist in? A typical definition has it that sympathy is “feeling for’ someone,” and “often involves feelings of concern.”<sup>29</sup> There is nothing unexpected in this, save perhaps that it ought to be emphasised that feelings of concern have a strong role to play, in that pity, as we have seen, doesn’t entail concern for its object.

---

<sup>27</sup> We have also seen that it is inadequate to think of ‘sympathy’ as co-extensive with ‘pity’ because pity does not necessitate concern for the pitied.

<sup>28</sup> Batson (2011)

<sup>29</sup> Eisenberg and Strayer (1987)

As remarked in the previous section however, it is implausible to suggest that a concurrent feeling of concern is definitive of a feeling's being either empathic or sympathetic.

This definition, of sympathy as 'feeling for someone', points us in the right direction of what kinds of mental phenomena empathy and sympathy are, but needs a little unpacking. Sympathy is 'feeling for', but as we have seen is not a feeling in itself, so sympathy must be some means by virtue of which a feeling becomes a 'feeling-for' – it is a *mode* of feeling. What do I mean by a 'mode of feeling'? To answer, let us consider what makes Avon's feeling of concern a *feeling-for* Omar, and not just a feeling as he might have if he were concerned about how he was going to pay his utility bills.

The most obvious difference is in the intentional object of the feeling of concern – that object which the concern is, to borrow Peter Goldie's term, "felt towards".<sup>30</sup> Omar's feeling of concern about his bills is a form of his thinking about his bills, and his concern for Avon a form of thinking about Avon. This alone does not make Omar's concern for Avon sympathetic, however. 'Feeling-for' and 'feeling-towards' are markedly different concepts, the former (we are assuming) being the hallmark of sympathy, the latter being a feature of emotions more generally; it is the way emotions are about things. It cannot be the case that all emotion-states, or even all emotion-states that are felt-towards another person (or features of that person or their situation) are sympathetic – Omar's envy of Avon's success could hardly be described as such. Nor is it the case that only *certain* emotion-states felt-towards another person will always count as sympathetic, although this is closer to the mark. This is not just because some emotion-states that are felt in the sympathetic mode can also be felt in non-sympathetic contexts - we might feel embarrassment on behalf of someone whose flies are undone, unknown to them, either

---

<sup>30</sup> Goldie (2002)

sympathetically (which might lead us to warn them), or not (which might lead us to joke about it). That is enough to prove my point that certain emotion states are always sympathetic. However, as we have already seen that pity, certainly the emotion most associated with sympathy, is not co-extensive with pity, we can make with confidence the stronger claim that no emotion is, by itself, sufficient for or otherwise definitive of sympathy.

Although sympathy, feeling-for, is not to be identified with a feeling of concern or pity for another, there can be no doubt that the disposition to have those kinds of emotional states is closely correlated with the disposition to feel-for another. The common thread between these two dispositions seems self-evidently to be a concern for the well-being of the person about whom we feel pity, or feel-for in general. With this in mind, I propose the following:

S feels-for O when:

- (1) S has some feeling F,
- (2) S values the wellbeing of O to some non-negligible degree, such that, all things being equal, S prefers that O is doing well.
- (3) (1) is true because (2) and
- (4) S believes that (3).

So, Avon feels-for Omar just in case his feeling springs from a general concern for Omar's wellbeing and he knows that to be the case. The belief condition might give some readers pause: why should Avon be required to believe that his feeling was caused by his valuing of Omar's wellbeing in order for it to count as sympathetic? It is a common thought that sympathy has a particular conative force,<sup>31</sup> in that it motivates us to alleviate the suffering, or generally promote the welfare, of the person with whom we sympathise. In order to preserve that property of feeling-for, it is necessary that the subject be aware that his feeling-for is prompted by a concern for the wellbeing of his

---

<sup>31</sup> Goldie (2002)

object. Avon, unaware that his sorrow for Omar is caused by a concern for his wellbeing, may lack the cognitive resources to direct the motivational force of his sorrow towards the alleviation of Omar's anguish (or the continuation of his pleasure). This is the case even when we can take Avon's sorrow to include Omar's situation as its object. Take a case, for instance, where Avon is feeling sorry that Omar has lost his job, but that there are several reasons why he might feel that way; he cares for Omar's wellbeing, and also cares more than most about the health of the local economy, and perhaps Omar also owes Avon money. If Avon is not aware which of his concerns is causally responsible for his sorrowful state, it will not necessarily motivate him to act in a way that promotes Omar's wellbeing in a way we would expect to be consistent with a sympathetic feeling.

### *1.2.3 Contrasting Sympathy and Empathy*

By investigating and refining the ordinary usage concept of sympathy, specifically how it compares to the way people generally talk about empathy, we have arrived at a point where we can use what we have learned about the former to inform a sustained discussion of the latter.

We have learned so far that neither sympathy nor empathy are feelings *per se*, but are rather processes that cause feelings to arise in a particular way. The process of sympathy can be distinguished from empathy by virtue of the relative positions that the sympathiser and empathiser place with their respective targets; this is reflected in the common idea that sympathy is a matter of feeling *for* someone, while empathy is a matter of feeling *with* someone.

What we take 'feeling with someone' to consist in has far reaching consequences. I discuss the different ways we can consider to affective states to 'match' in the following sections. Now, however, it is important to bring up what will turn out to be the most problematic feature of the most popular views of empathy; that is, what constraints we put on the matching affective states that suffices to mark them out as empathic.

Even once we accept that successful affective matching was required of a state in order for it to count as empathic<sup>32</sup>, it doesn't take much to show why empathy can't solely consist in that. There must also be some constraints regarding the cause of the matching state. If Avon and Omar are both confronted with a bear, the fact that they are both afraid at that moment cannot by itself mean that they empathise with one another; their feelings have nothing to do with one another and everything to do with the bear. To take the example further, claiming that affective matching by itself, even in its strongest form, is sufficient for empathy will entail that Avon and Omar, confronted with different bears some hundreds of miles apart from one another, will even then be empathising with each other, which is clearly not the case.

The majority of empathy theorists acknowledge the lesson of the 'bear' case, and embrace it by building into their theories some constraint on the source of the affective states in question. Hence Vignemont and Singer, in their definition, mandate that for a state to count as empathic it must be "(iii) elicited by the observation or imagination of another person's affective state".<sup>33</sup> Hoffman suggests the narrower constraint that a state counts as empathic only if it is generated by the "attended perception of the object's state", that is, that we must both perceive and attend to the emotional state of the target of our empathy in order for it to count.<sup>34</sup> There is much to commend Vignemont and Singer's theory. As we will see later it is not by itself satisfactory, but does clearly point us in the direction of one that is. Hoffman's proposal, on the other hand, certainly seems too restrictive, particularly if we want to allow that we can empathically share affective states with people who are not present by imagining. So, what is the correct constraint we should place on the source of an affective state in order for it to count as empathic?

---

<sup>32</sup> On my view, such a feeling would strictly count as a part of a proto-empathy, but most others consider a successful affective match the ultimate goal of empathy.

<sup>33</sup> Vignemont & Singer (2006)

<sup>34</sup> Hoffman (2000)



### 1.3 Theories of Empathy: Higher and Lower

At the heart of many disagreements about the appropriate causal constraint we should take to be definitive of a feeling felt empathically lies the fact that there are several psychological, neurological, behavioural and physiological phenomena that appear to share some empathy-like features in that they are capable of manifesting affective states that match those of another. *Prima facie* they could all count as causal constraints that make an affective match count as empathy. This leads naturally to the questions of whether any of these phenomena counts as empathy proper to the exclusion of the others, whether all these phenomena are empathies of different sorts, whether they share a common feature that is itself empathy or whether empathy is a disjunctive concept ranging over them all. Indeed, these options are not all mutually exclusive, which serves to further muddy the problem.

To begin with I will outline what I take to be the phenomena claimed to be candidates for either solely, collectively, disjunctively or otherwise falling under the description of empathy; these are:

- (a) Emotional Contagion
- (b) Low-Level Simulation
- (c) High-Level Simulation
- (d) Perspective Shifting
- (e) In-His-Shoes Imagining, and
- (f) Prosocial Behaviours<sup>35</sup>

---

<sup>35</sup> See section 1.1 for summaries of these.

These six phenomena are our candidates for empathy. One unified conception of empathy, according to which each of these phenomena are instances of empathy, is defended by Stephanie Preston and Francis de Waal. In an oft cited paper they have explicitly argued that ‘empathy’ ought to be considered a broad ‘umbrella’ concept. They offer what they call a *Perception-Action Model*, ranging over all cases in which, and here Preston and de Waal borrow from Hoffman’s definition of empathy, “the attended perception of the object’s state generates a state in the subject that is more applicable to the object’s state or situation than to the subject’s own prior state or situation.”<sup>36</sup> Preston and De Waal’s major contribution to the debate is to propose that all cases that lead to a Hoffman-type empathic state in a subject can be distinguished by their shared neurological base in a process that can be described under a single model as follows:

**The Perception-Action Model (PAM):**

*[Empathic states are the results of processes in which] attended perception of the object’s state automatically activates the subject’s representations of the state, situation, and object, and that activation of these representations automatically primes or generates the associated autonomic and somatic responses, unless inhibited.*<sup>37</sup>

*1.3.1 The Perception-Action Model*

According to Preston and de Waal, the Perception-Action Model (PAM) provides a suitable grounding for a unified conception of empathy, in that our list of empathy-like phenomena *a* to *f* all share a distinguishing characteristic by being describable under the same

---

<sup>36</sup> Hoffman (2000)

<sup>37</sup> Preston and de Waal (2002)

model. Each phenomenon on our list, on Preston and de Waal's view, can be considered a species under the genus 'empathy'.

In opposition to de Waal's view that the differences between the various empathy-like phenomena have been "emphasised to the point of distraction",<sup>38</sup> Amy Coplan has argued that attempting to find a unified base for the various empathy-like phenomena is a fool's errand. For Amy Coplan, and those sympathetic to her, empathy does not properly name a genus of activity, different processes describable under the same model, but rather a species of activity, a single process. Specifically, in Coplan's case empathy is to be identified with phenomenon (e): perspective shifting. Coplan argues that, of the listed phenomena, only perspective shifting is capable of being completely mapped onto any adequate conceptual framework we might propose for empathy.

Coplan claims that there are strong conceptual distinctions between our candidate phenomena that make it impossible to find a satisfactory definition of empathy that ranges over all processes (a) to (f), and that chasing such a definition will only cause problems for further research.<sup>39</sup> Others press different conceptual distinctions in addition to those mentioned by Coplan, but the two sorts of difference most often called upon in the literature are the two that Coplan leans on most heavily. The first is based on the observation that phenomena (a) to (c) are automatic, involuntary processes while (d) and (e) are voluntary and effortful activities. The second is that empathy requires a clear self-other distinction. In order for Avon to empathise with Omar, he must have an awareness that the purported empathic state originates in Omar and not in himself. This is significant for two reasons: Avon must firstly be aware that the state did not originate in his own psychology, and secondly he must be

---

<sup>38</sup> Ibid.

<sup>39</sup> Coplan (2009) and (2011)

aware that the psychology that did produce it may have significant differences to his own. Coplan argues that the only phenomenon in the above list that can meet these conceptual requirements is perspective shifting, and so that is this species of activity that Coplan identifies with empathy.

#### 1.4 Theories of Empathy: Broad and Narrow

I take it that Coplan is obviously correct on at least one point; the various empathy-like phenomena listed above are conceptually distinct. Emotional contagion and low-level simulation are both involuntary, automatic processes. This distinguishes them sharply from high-level simulation, in-his-shoes imagining and perspective shifting, all of which are mental undertakings; they are mental events that occur as a result of, or are constituted by, a volitional mental activity. Preston and de Waal admit that such phenomena as these last three (which they collectively label 'cognitive empathy') are "partially distinct from the more automatic...forms of empathy"<sup>40</sup>, but nonetheless propose that the Perception-Action Model can account for them. Their reasons are vague, and suggestive more than decisive, but the argument seems to turn on the idea that the automatic and reflexive systems of Emotional Contagion and Low-Level Simulation promote development of areas of the prefrontal cortex that are also essential to these varieties of 'cognitive empathy'.<sup>41</sup> As a hypothesised conceptual connection this is, I think, a non-starter, since although it is indeed highly likely that some of the systems that are automatically activated in cases of emotional contagion are also activated in cases of,

---

<sup>40</sup> Preston and de Waal (2002)

<sup>41</sup> *ibid.* p. 19-20

say, high-level simulation, that by no means shows that the processes are of the same kind.<sup>42</sup> The system responsible for my blink reflex will undoubtedly share many features with the system I engage when I might bat my eyelids in a display of coyness, but the fact that one process is a reflex and the other is a volitional action remains a vital point of difference between them. The developmental parallel, interesting though it is, does nothing to address the issue of how ‘cognitive empathy’ can fall under the Perception-Action Model when the Perception-Action Model explicitly only covers automatic processes, which the varieties of ‘cognitive empathy’ are emphatically not.<sup>43</sup>

The claim that prosocial behaviours can be conceptually collected under the same label as either cognitive or non-cognitive forms of empathy is even less persuasive, since while helping behaviour might well show strong correlation with empathic events, there is no reason to think that the impulse to carry out such behaviour (much less the behaviour itself) can only be caused by those events. Furthermore, even if there were no other reasonable cause to be suggested for prosocial behavior, for reasons familiar from the standard refutations of behaviourism it would be implausible to suggest that a certain pattern of behaviour was definitive of empathy in a way that it would

---

<sup>42</sup> The peer commentary following Preston and de Waal’s article suggests that many cognitive scientists and other empathy researchers share my scepticism of their attempt to overcome the conceptual challenge. Some offer solutions, but none are very promising. See, for example, Rochat (2002), Parr (2002), Hoffman (2002), Gordon (2002), Davies (2002) James et al. (2002) and also Ainslie and Montrosso (2002).

<sup>43</sup> Some studies have shown that in fact Emotional Contagion and Low-Level Simulation are not always automatic in the sense of being purely stimulus driven. Singer et al. (2006), Lamm et al. (2013) and Vignemont & Singer (2006) discuss findings that Mirroring processes can be partially inhibited when, for instance, the subject believed the pain their target was undergoing was deserved or necessary. This kind of inhibition does not seem amenable to being described as ‘volitional’, however. Even if a case were made that it could, a difference still remains between an activity being embarked upon volitionally and being capable of being inhibited volitionally.

not be to claim that of a particular mental process. Empathy must name a possible *cause* of prosocial behaviour, not the behaviour itself.

Because of the sharp conceptual distinctions cutting three ways across phenomena (a) to (f) – (a), and (b) are automatic, subconscious processes, (c), (d) and (e) are volitional and conscious mental actions, and (f) names a pattern of behaviour often resulting from (a) to (e) which are mental processes, and is therefore a different kind of thing altogether – it seems that Coplan is correct in her assertion that the attempt to bring all six phenomena under a single concept is futile. However, it might be objected, indeed it has been objected by some neuroscientists,<sup>44</sup> that this particular difference I have pointed to is not sufficient to show that there is no sense in which cognitive and non-cognitive forms of ‘empathy’ (ignoring the question of helping behavior for the sake of charity) cannot be brought under the same concept.

#### *1.4.1 The ‘Two Routes’ View*

Alvin Goldman has suggested that the automatic and cognitive brands of empathy discussed above actually name two distinct ‘routes’ to empathy.<sup>45</sup> There is a ‘mirroring’ route that makes use of mirror neurons, is automatic, sub-personal and often occurs unconsciously. There is also a ‘reconstructive’ route to empathy (equivalent to what Preston and De Waal called ‘cognitive’ empathy), where empathy results from an effortful imaginative process to reconstruct, in the subject, the affective state of the target.

This kind of approach to defining empathy has many benefits. The problem of providing a unified account of empathy that captures both the automatic and cognitive routes, as Coplan pointed out, seems insurmountable. It is, however, plausible on the face of it that empathy can consist in either of these two processes, and furthermore

---

<sup>44</sup> De Waal (2002), Iacoboni (2011)

<sup>45</sup> Goldman (2006) and (2011)

that these processes seem to provide at least the cornerstone of a sufficient causal constraint to counter cases like Avon and Omar's bear. The thought seems to be that, although the processes by which they come about might be distinct, there is enough in common between the states that result from these different processes that they ought to be considered of the same kind, viz. empathic. So much seems to reflect a growing consensus among empathy theorists, but as it stands I believe this view faces some important difficulties.

The problems facing the Two Routes View partly stem from the fact that nearly all current theories of empathy foreground affective matching, in that they often take an affective match between a subject and object, coupled with an appropriate causal constraint, to be necessary and sufficient for empathy. I believe that the dominance of affective matching in the debate around the nature of empathy is unwarranted, resulting in problems for the Two Routes view on its own terms, and the neglect of the centrality of the epistemic function of empathy as a way of offering a distinctive kind of knowledge about another person. The next section will offer some considerations of the place of affective matching in empathy research.

### 1.5 Affective Matching and the History of Empathy

Two of the definitions of empathy I quoted from the popular press suggested that empathy involves feeling what another person feels (E<sub>1</sub>, E<sub>3</sub>). Another has it that empathy merely involves understanding another's feelings, without necessarily sharing them (E<sub>2</sub>). One even suggests that empathy does not involve sharing the feelings of another at all (E<sub>4</sub>). This disagreement, echoed in academic debates about the nature of empathy, comes down to whether and to what degree empathy involves affective matching; that is, having an affective state

that is isomorphic with (the same as or similar to) the affective state of another.<sup>46</sup>

The recent history of empathy research has mostly focused around conceptions of empathy that foreground affective matching, such as those proposed by Hoffman, Iacobani and Vignemont and Singer. I will argue later that this has been a mistake, but to explain why the mistake was made in the first place it is instructive to briefly rehearse the short history of empathy as a philosophical concept.

Although the term ‘empathy’ is a relatively new addition to the English language, appearing for the first time as a translation of the German *Einfühlung* (“feel into”) in 1909,<sup>47</sup> the use of the term to refer principally or exclusively to the sharing of another’s affective state is an even more recent development. Before the resurgence of interest in empathy research in the mid-1980s, empathy was conceived far more broadly as a means by which one could ‘feel one’s way into’ another object, be it a person, sculpture or an animal. One of the great so-called ‘empathists’, Rudolf Lotze, went so far as to claim that “No form is so unyielding that our imagination cannot project its life into it”,<sup>48</sup> a sentiment echoed by much of the work of the empathists of the 19<sup>th</sup> and early 20<sup>th</sup> centuries such as Theodor Lipps and Vernon Lee. ‘Feeling into’, in the sense used by the empathists, referred only to a method of understanding the object to be empathised with. Once we had ‘felt our way into’ whatever object we had set our empathic sights on, the empathists believed that we could avail ourselves of a peculiar kind of understanding of what it was like to be the object in question. *Einfühlung* is not a matter of reproducing the feelings of another object, most candidate objects obviously not being the kinds of things that have feelings of their own. Rather, it is a matter of using

---

<sup>46</sup> Precisely what kinds of similarities between states amount to ‘isomorphism’ is something I talk about in depth in later sections.

<sup>47</sup> Titchener (1909)

<sup>48</sup> Lotze (1856)



one's own feelings to understand the *being* of the target object, be it human, animal, mineral or vegetable. The idea of feeling one's way into a thing seems to have been deliberately vague on the accounts of these philosophers, *Einfühlung* principally serving as a broad, romantically minded foil to the contemporary scientific attitude of prioritising empirical observation as a route to understanding the world, and not subjected to much in the way of prolonged analysis or description.

The views of the early empathists are at various stages confusing and implausible, and I will not seek to defend them, but the point to be taken from this brief historical detour is that the concept of empathy as a route to understanding is by no means historically limited to the understanding of the affective states of other people (much less merely the *sharing* of those states). Possibly the last great empathy theorist of the early 20<sup>th</sup> century, Edith Stein, herself saw empathy as “the experience of foreign consciousness in general”,<sup>49</sup> a view considerably narrower than that of Lotze and Lipps, but certainly encompassing more than the modern definitions cited above. Empathy in the early to mid-twentieth century was serving a much more focused role than it had previously; principally as the groundwork for a positive component of the phenomenological tradition's rejection of an ‘inference from analogy’ as the foundation of our perception of other people as minded, conscious subjects.<sup>50</sup> While it may be true that empathy has always been taken to involve feelings (on the empathiser's part at least), I think it is a mistake to take for granted that feeling the same as the object with which one empathises has always been thought of as the *sine qua non* of any empathic episode. The early empathists thought that empathy was a way for our feelings to inform our understanding of an object, some phenomenologists thought that our feelings gave us a way of understanding others as

---

<sup>49</sup> Stein (1917)

<sup>50</sup> *Ibid.*

conscious subjects. Neither of these views of empathy require affective matching in the way we currently understand it, and yet affective matching has become the dominant criterion of empathy across almost all current theories.

I believe there are two principal reasons why affective matching has become so central to contemporary empathy research. The first is that the conceptual framework of this research has been somewhat contaminated by the continued closeness of empathy and sympathy both in the minds of ordinary people, and in the academic tradition that owes so much to the work of Hume and Smith. In general terms, the common conflation of sympathy and empathy (of which more later) encourages a conception of empathy that is fundamentally or distinctively affective in nature, since sympathy is distinctly affective and is generally better understood than empathy. In terms of the legacy of Hume and Smith, they originally credited the concept of sympathy with playing two distinct roles; first, with encouraging us to share the affective state of other people and, second, with motivating ourselves to alleviate the suffering of other people. On most modern analyses, the motivational role is restricted to sympathy, whereas empathy is credited with our frequently sharing the affective states of other people. This analysis is fine as far as it goes, but by describing empathy principally in terms of a partial role it played in Humean Sentimentalism we foreground its sentimental qualities in a way which is not necessarily warranted. Empathy, for a modern reading of Hume, needs to allow people to understand that another person is suffering in order that we can then sympathise with them, and so be motivated to help. Broadly speaking, it isn't necessary that the understanding that empathy grants us is primarily based on affective matching. However, empathy is generally understood to primarily consist in affective matching anyway because, in part at least, our modern conception of empathy is descended from Hume's conception of sympathy. Affective matching *was* a crucial component of Humean

sympathy, because the one emotional state was required to both inform the subject that the object of their sympathy was suffering *and* motivate them to alleviate it. The sharing of a negative affect achieved these two aims, but since we no longer think empathy and sympathy ought to be conflated, empathy no longer needs to consist in affective matching in order to play its role in Humean moral motivation.<sup>51</sup>

The second reason I believe affective matching has become so ingrained in contemporary empathy research is that since affective responses are relatively easy to observe they have become the focus of much scientific work on empathy. It is relatively easy to empirically observe whether or not one person is feeling the same way as another person, but very difficult to determine with the same level of clarity whether one person is identifying with another person in more psychologically rich ways (sharing their evaluative dispositions, for instance). The ease of measuring brute affective states might easily have led to other possible facets of empathy being largely ignored. Furthermore, the discovery of mirror neurons, and related advances in the understanding of how our brains might be wired to automatically simulate the affective states of others, have attracted claims that these are the processes that ‘really’ constitute empathy, which the consequence that empathy can be effectively reduced to affective matching and nothing more.<sup>52</sup> Many empathy theorists seem deeply impressed by these claims, to the extent that theories of empathy are engineered from the ground up to incorporate them in a way that is both at odds with much of the historical literature on empathy (as noted above), but also becomes unwieldy in combination with their own views, as I shall make clear in the following sections.

Furthermore, it is a mistake to think that the fact that empirical observation of affective states is easy *relative* to observation of more

---

<sup>51</sup> Davis (1980)

<sup>52</sup> Preston and de Waal (2002)

psychologically rich states means that empirical observation is a simple matter on its own terms. As a specific example, it is typical of research from the sciences published on empathy or its cognates (mindreading, ‘emotional intelligence’ and the like) to measure the ability of people to empathise using the ‘Mayer-Salovey-Caruso Emotional Intelligence Test’.<sup>53</sup> The test is based on identifying emotions expressed in photographs. The test does not claim to tell whether its subjects are correct in the identifying of emotions expressed in photographs, merely whether subjects’ answers agreed with answers given by acknowledged experts. If people were systematically unable to detect a particular kind of emotion from a particular facial expression, this test would be unable to detect that fact.<sup>54</sup> Neither of these oversights is insurmountable, and indeed some work on empathy include neither,<sup>55</sup> but the point remains that affective matching has become definitive of empathy in a way that I believe is unfortunate.

The preceding few paragraphs gives me the first tentative basis on which to introduce one of the major claims of this thesis; that empathy is not, at root, a matter of affective matching. I do not disagree that a successful empathising will involve affective matching, I merely dispute the idea that affective matching is empathy’s ultimate purpose or central feature.

Recall that according to the Complex Theory, empathy is the making of a demonstrative ascription to a target of the form ‘[the target] feels like *this*’. Importantly, the ‘feeling’ referred to in the empathic ascription isn’t intended to stand for merely the affective states or emotions of the target, but rather the indeterminately wide-ranging phenomenal character of what it’s like to be the target. The emotions

---

<sup>53</sup> Mayer et al., (2004)

<sup>54</sup> As we will see later, the suspicion of those kinds of systematic failings are a part of what motivates my own view.

<sup>55</sup> Ickes (1997), Miller (2011)

and affective states of my target will certainly form a part of my demonstrative ascription, and will also inform my imaginings about what it might be like to be a person with those emotional states. However, the feeling I will eventually ascribe to you will (probably) be richer than a mere set of affective states. The feeling I ascribe to you when I empathise will include the feeling of what it's like to be a person like you with those emotions at that time, not merely the feeling of the emotions themselves.

The view of empathy I am proposing is psychologically and narratively very rich. That is, I am proposing that empathy is capable of delivering us an understanding of what it's like to be another person that takes into account not just the emotional and affective states they may be experiencing, but also those states' relationship to one another and to the object of empathy, as well as to their situation, personal history and character. As such, the ability to affectively-match one's target when attempting to empathise remains of crucial importance on my view; the type, intensity and object of the emotional states of one's target provide much rich material from which to construct an experience to ascribe to them. However, although I agree that affective-matching is a necessary part of empathy, I do not think that it is sufficient for empathy, nor do I think it is sufficient *grounds* on which to make the demonstrative ascription that I take empathy to consist in. Affective matching, on my view, is one of several *proto-empathies*; that is, imaginative projects which need to be undertaken in order to provide sufficient grounding for an empathic ascription. The next section will outline why I think that affective matching, even with appropriate causal constraints, cannot provide sufficient grounds for empathy by itself.

## 1.6 Regularity, Reliability and Usefulness

As I outlined in the previous section, it seems that historically affective matching has not always been considered necessary for empathy. After all, one cannot match affects with an object that has no affective states, and since those early empathists thought that one could empathise with that kind of object, it can't be the case that they considered affective matching necessary for empathy. My disagreement with contemporary empathy theorists such as Iacoboni, Hoffman and Vignemont and Singer is not, however, about whether or not empathy involves affective matching. Rather it is my contention that what they have identified as empathy is actually a variety of proto-empathy, and as such something that is necessary to ground an empathic ascription (or, if we prefer, a necessary part of the empathic process), but insufficient for empathy as a means of understanding another person. In what follows I lay out some desiderata for a theory of empathy, that it makes empathy regular, reliable and useful, and show how different conceptions of what constitute affective matching open to the Two Routes View struggle to meet them.

Among the desiderata for a good theory of empathy must be that it allows that empathy can successfully occur with some regularity. That is, although we need a theory that grants empathy sufficient causal constraints that we can evade the 'bear' challenge we must still make it plausible that a process occurring within those constraints could regularly lead to successful empathy with another person. One way of making it more plausible that any given process will result in an affective match between an empathiser and their target is to adopt a conception of affective matching that requires only a very minimal standard of similarity for two states to count as isomorphic.

There is some disagreement among various theorists as to precisely what isomorphism in empathic affective matching must amount to. Some propose that a congruence of affect between empathiser and

target, nothing stronger than some qualitative similarity or identical valence between the two, is all that we need claim is involved in affective matching.<sup>56</sup> Qualitative similarity is described in very loose terms by those who suggest it as a criterion for affective matching, so it is difficult to assess what precisely it amounts to. I think it is not unreasonable to suggest that it might mean something like ‘involving some of the same perceptible affective properties’, such that Avon’s state counts as affectively matching Omar’s if Avon is feeling anger and Omar frustration, because they both might involve feelings of tension. ‘Identical Valence’ as the criterion for affective matching is only slightly less vague; two affective states have identical valence if they are both ‘positive’ emotions (such as joy and pride) or are both ‘negative’ emotions (such as anger and fear). Again, the identical valence criterion for affective matching allows for some odd cases; your shame and my grief would count as matching affective states on this view, and it would be a stretch to suggest that we could be empathising with one another on the basis of our having two such different emotions, identical valence notwithstanding. Others place stricter demands on what will count as affective matching, requiring that the feeling of the empathiser and the target are of the same kind, although accepting that the matching state can vary in degree or intensity from that of the target.<sup>57</sup>

I think it is highly unlikely that a mere identical valence will suffice for the kind of affective matching that we need to play a role in our conception of empathy. For one thing, evidence both from cognitive neuroscience<sup>58</sup> and from our own experience tells us that empathy is capable of a great deal more precision than that. For another, the idea that all empathy can reliably lead us to understand is the valence of another person’s affective state leaves us with a dramatically etiolated

---

<sup>56</sup> Hoffman (2000), Preston and deWaal (2002)

<sup>57</sup> Coplan (2011)

<sup>58</sup> Debes (2010)

notion of empathy. A further desideratum for a good theory of empathy, in addition to allowing empathy to regularly occur, is that it must also posit the understanding that empathy grants us as in some way useful. If empathy were not useful, it would be extremely difficult to account for how so much evolutionary effort has been expended in developing the subtle and complex neurological systems that underpin it. I think it is obvious that an empathy that could only reliably replicate the valence of another's affective state could not be described as especially useful. If Omar is feeling apprehensive and Avon, upon empathising with Omar (either automatically or cognitively), simulates anger, thus understanding Omar to be angry (an emotion of identical valence), Avon is liable to make some serious social blunders on the basis of that false belief. For our current purposes I can afford to be pluralistic insofar as what ought to count as isomorphism in an affective match. I am attracted to Coplan's criterion that two people's states affectively match only when they are the same kind, although they can vary in degree and intensity.<sup>59</sup> That may turn out to be too restrictive, particularly if, as I suspect, that would necessitate the empathiser being aware of what type of feeling they were having. In any case, I think it is clear that in order to preserve the usefulness of empathy, it ought to consist in something more than identical valence. Once we do assume affective matching to consist in something stronger than emotions of identical valence, Goldman's 'two routes' view becomes problematic. Unfortunately for the Two Routes view, it turns out that when one route is taken, affective matching is ruled out if it consists in something stronger than identical valence. Even worse, if the other route is taken affective matching is guaranteed, but useful understanding of the target is ruled out. In what follows I will spell this criticism out in detail.

---

<sup>59</sup> Coplan (2011)



Goldman's view is explicitly based on Vignemont and Singer's definition of empathy. That is:

*“There is empathy if: (i) one is in an affective state; (ii) this state is isomorphic to another person's affective state; (iii) this state is elicited by the observation or imagination of another person's affective state; (iv) one knows that the other person is the source of one's own affective state.”*<sup>60</sup>

Their view suits Goldman well as a starting point, firstly because it is in itself an initially plausible (if ultimately flawed) account of empathy as affective matching. Secondly, because the disjunction built into the condition of their definition governing appropriate sources of empathic states gives conceptual space for either route to be taken in any given case of empathy. However, their conception of empathy ultimately fails because the additional constraints necessary to ensure that sufficient understanding is granted in each 'route' to empathy makes it highly unlikely that either one could be successfully navigated.

#### *1.6.2 Automatic 'Empathy' is not Useful*

Whatever else empathy is, it is clear that one of its functions is as a method of coming to understand the inner life of another person.<sup>61</sup> In order for it to serve this purpose it is not sufficient that we merely feel the same way as another person, nor even that we also feel the same way as another person because that person feels that way. There must be some intentional attitude linking the empathiser's state with that of her target; if I am going to form beliefs about another's inner life, it must be based on information that I have access to. Vignemont and Singer prescribe a knowledge condition; we must *know* that we feel the same way as our target because they feel that way. More precisely,

---

<sup>60</sup> Vignemont and Singer (2006)

<sup>61</sup> This phrase is Heinz Kohut's (1984), and is particularly representative of how much of the psychoanalytic community have conceived of empathy as relating to their practice. See also: Schafer (1959), Davis (1990)

we must know that the target is feeling the same way as we are, that our feeling is suitably caused by our target's feeling that way, and that the feeling in question was originally a product of the target's psychology, and not of ours. I do not think we need to go as far as prescribing knowledge; a justified belief will do the work just as well. In either case, these stipulations are warranted by the regularity, usefulness and reliability criteria that I referred to earlier; if we want empathy to be useful to us (specifically, useful in forming beliefs about the inner lives of others), then it must happen, we must know that it is happening and it must provide us with some useful, and one would assume accurate, information.

The knowledge condition, which is captured in part (iv) of Vignemont and Singer's definition, has very different effects on the overall picture of empathy offered depending on which route offered by the disjunct in condition (iii) applies in each case. If one was to attempt to empathise following the reconstructive route it seems that one could not fail to know the identity of the target of one's empathy. It does not seem possible that Avon could attempt to imaginatively reconstruct in himself Omar's affective state without knowing that it was Omar whose state he was reconstructing. The same is true if we replace the knowledge condition with a justified belief; if Avon is attempting to reconstruct himself in Omar's affective state, then he will always have a justified belief that it is Omar's affective state that he is reconstructing. It seems, therefore, that the knowledge condition is somewhat redundant if the reconstructive route is being taken in any given case.

The condition is necessary, however, if the automatic route to empathy is taken. The Perception-Action Model (which I take to be a good description of the automatic, if not the reconstructive, varieties of empathy) does not entail that the empathiser be aware that her state is causally related to that of another person. PAM does specify that it must be *attended* perception of the object's state that triggers

the automatic simulation of that state. It is unclear precisely what is meant by 'attended' perception here. It could plausibly mean that one has to be aware of one's perception of, say, the empathic target's facial expression or that one has to be aware of what kind of expression it is that one is perceiving in the target. I'm not convinced that either interpretation allows certain documented cases of emotional contagion that Preston and de Waal would certainly want to include under PAM.<sup>62</sup> These cases include such instances as a person talking to someone who smiles, and whose smiling triggers an automatic mimicking of that precise facial expression in the subject, causing the subject to feel the same kind of emotion as their interlocutor. In such cases it does not seem that the subject cognises their perception in any sense; they do not need to recognise that they have perceived a smile in order for it to have the effect noted. It seems to be generally accepted that there is no requirement that the subject consciously register the expression being perceived in order for some automatic mirroring process to occur.<sup>63</sup>

However, even if the perception of a target's state does have to be attended to in such a strong sense, there is no reason to think that the conscious realisation that one's target is in a state identical to one's own will lead one to the belief that the one state is causally related to the other. Imagine that Avon and Omar are once again both facing a bear, but that Avon is generally not afraid of dangerous wild beasts. If Avon sees that Omar is afraid, and then himself becomes afraid as a result of an automatic mirroring process, he would be entirely justified in putting his fear down to a change in his attitude towards bears and not a sub-personal, automatic mimicking of Omar's fear. If Avon

---

<sup>62</sup>Goldman (2011)

<sup>63</sup> There have been some attempts to argue that mirroring processes do give rise to understanding of the sort we are concerned with, but these have been uniformly unpersuasive. (Vignemont (2009) cf. Debes (2010))

knew something about the psychology of social cognition he might make the causal connection between his state and Omar's, but failing that it may well be more reasonable for him to believe that his fear is a straightforward affective response to the bear he sees than that his fear has its source in Omar.

The condition that an empathiser must know that her affective state has its source in the target of their empathy, as suggested by Vignemont and Singer, successfully blocks some cases that we would not want to call empathy. If Avon contracts Omar's fear via emotional contagion, but believes the state originates in himself, we may resist calling Avon's fear empathic. However, that condition also seems to block other cases that we may want to call empathic. If Avon contracts Omar's fear via emotional contagion, but believes that it originates in Lester, who is also present, the knowledge condition seems to prevent that being counted as empathy. Is there really no empathy in such a case? Is it not rather that Avon is mistaken about the target of his empathy?

It is precisely these kinds of worry Alvin Goldman (2011) had in mind when he wrote, of Vignemont and Singer's definition:

*"I think it is fair to require a receiver to have some sort of intentional attitude directed toward the target by which the resonating state is linked to him. Otherwise, it doesn't seem like a case of empathy. I suspect that condition (iv) is too strong an intentional condition of this kind, but I don't have a wholly suitable replacement for it."*

Replacing the knowledge condition with one requiring a justified belief would evade this problem, since Avon, although mistaken, may be justified in believing that his fear originated in Lester rather than Omar. I am sceptical, however, of the possibility that a replacement for the knowledge condition can be found that avoids the worst of the above problems. A justified belief condition, rather than a knowledge condition, will still be too strong to allow the vast majority of instances of automatic mirroring to count as empathy. This is simply

because the nature of the empathic project we are trying to capture requires the kind of intentional state that can justify an ascription of the feeling one has to another. There are, as I will show, other ways of justifying such ascriptions than by requiring knowledge. However, I think it is highly implausible that any such justification can be found by proposing some intentional state that is unconsciously brought about along with the mirrored affect, and that is what Goldman's intentional state would have to be in order to fulfil its proposed function.<sup>64</sup>

### *1.6.3 Reconstructive Empathy is not Reliable*

The reconstructive route to empathy guarantees that condition (iv) of Vignemont and Singer's definition is met, but the automatic route does not. Interestingly, the reverse holds of condition (ii). No matter how strictly or otherwise we define 'isomorphism' it will always be the case that the automatic route is more or less guaranteed to result in an affective match between subject and target, whereas the reconstructive route to empathy is liable to misfire. Automatic mirroring processes, often activated by 'microexpressions' in the target (which are by their nature nearly impossible both to consciously mimic and consciously perceive), are not susceptible either to fraudulent emotional expression or to inept interpretation of emotional expressions in the same way as their reconstructive counterparts.<sup>65</sup> Reconstructive empathy relies on the empathiser being able to consciously garner what information they can about their target's emotion in order to simulate their mental state as accurately as possible. The reliability of the reconstruction therefore

---

<sup>64</sup> I am concerned here with securing an *internal* justification for an empathic ascription. We might be able to secure an external justification for such an ascription by appealing to reliable mechanisms that regularly provide affective matches (such as emotional contagion), but such mechanisms will fail to reliably provide an intentional link from the empathiser to their target. Such a link is necessary for any useful, conscious understanding of the target to result from the empathising.

<sup>65</sup> Iacobani (2011)

depends on the reliability of the methods by which we come acquire that information.

The reconstructive route to empathy can misfire in a number of ways. Avon, attempting to empathise with Omar, may be mistaken about several features of Omar and his situation that could lead him to imagine having an affective state that is not isomorphic with Omar's. Avon might simply not be good at consciously reading people's body language or other physical expressions of emotion and may thereby mistake Omar's experience of guilt for anger, leading him to imagine the wrong type of emotion in himself in his effort to empathise. Avon would not even have to be unusually bad at picking up on physical cues to make this kind of mistake. Facial expressions are one of the major sources of information concerning the mental states of those who bear them. Even though people are generally accurate in their reporting of the kind of facial expression they perceive in another (a smile, a frown, a gasp), they are far less reliable when it comes to determining the emotion state that is being expressed in that way. A widely reported study conducted at MIT showed that people often smile as an involuntary expression of frustration. Their trial involved one group of participants who were video recorded carrying out a frustrating task, and later video recorded doing their best to deliberately express frustration with no frustrating stimulus. Another group of participants was asked to judge what emotion was being expressed in each case. It was discovered not only that people who perceived frustration-smiles on video stills of the participants performed no better than chance at identifying the emotion expressed (from a list of three options), but also that when asked to act frustrated only 10% of participants smiled in the course of their pretence. A final interesting piece of data from the experiment is that people tended to be significantly better at identifying pretended

emotions than genuine ones.<sup>66</sup> This suggests that people are in general fairly unreliable at determining precisely what emotions are associated with what facial expressions, especially in cases where they are genuine expressions of some emotion. The fact that people seem to be significantly worse at identifying genuine emotions than feigned ones certainly seems to threaten the idea of reconstructive empathy being a reliable route to understanding. If anything, it suggests that we are more likely to identify the emotions people wish to be perceived as experiencing, and not what they actually do experience.

---

<sup>66</sup> Hoque et al. (2012), see also, Brackett et al (2004), Brackett and Mayer (2003)

## 2. THE COMPLEX THEORY OF EMPATHY

### 2.1 Introduction: What we need from a New Theory

It would seem to follow that the reconstructive route to empathy, unlike the automatic, cannot be reliably embarked upon solely on the basis of perceived facial expressions. The issue here is one of justification: if the end goal of empathy is to come to some understanding of the feelings of another person, we need some way to justify the ascription of those feelings to that other person. We need justification of two different sorts. The first, which reconstructive empathy always provides, but automatic empathy rarely if ever does, is some reason for thinking that the feeling being felt is not an affective response to the situation in which we find ourselves, brought about by our usual affective response mechanisms. The second sort of justification, which automatic empathy does provide, but which we might say is impotent absent the first sort of justification, is some good reason for thinking that the state we are experiencing is isomorphic with the state being experienced by the target.

I have given some evidence that people generally perform poorly at identifying the type of emotion indicated by any given facial expression, however that is certainly not the only source of information we can call on in order to determine the type of affective state one would have to conjure in oneself to bring about an affective match between ourselves and our target. Despite our propensity to mistakenly identify the particular emotions associated with certain types of facial expression, we may nonetheless be justified in taking our judgement in a given case as accurate if we have some additional support for it. Recalling once again Avon, Omar and the bear, we might think that, despite not being generally reliable in identifying emotion types by their expression, the fact that Omar is faced with a bear, a plausibly fear-inspiring experience, lends credence to Avon's



belief that the fear state he imagines in himself is isomorphic with Omar's.

It seems to follow from Avon's belief about the object of Omar's fear that the content of his imagining will change to incorporate it.

Consider the kind of imaginative project that Avon is engaging in when he attempts to imaginatively reconstruct Omar's state. It is not, primarily, a matter of propositional imagining, of Avon 'supposing' that he is in a state isomorphic with Omar's. Rather, it is an example of 'sensory imagining', that is, of imagining from the inside a sensory experience.<sup>67</sup> Given that, we should understand Avon's imaginative project to reconstruct in himself Omar's affective state as an effort to imagine having an experience of being in that state. Avon, in this case, is trying to imagine having Omar's experience of being afraid. I contend that for Avon to successfully complete such an imaginative project would entail that he not only imagine feeling afraid, but feeling afraid-of-a-bear, where the object of the fear is a part of the imagined experience. It might be claimed here that I am requiring too much to be included in the content of the imagined experience; why could it not be the case that one *experientially* imagines only the feeling of fear, while *propositionally* imagining, of that fear, that it was caused by facing off to a bear?

There is a general, though not unanimous, consensus among philosophers that emotions are intentional states; that is, that emotions are directed towards objects or states of affairs.<sup>68</sup> If we agree, it seems that we must require that if Avon is to successfully imagine having an experience relevantly similar to the experience Omar is having that Avon must not just imagine 'being afraid', but 'being afraid of' the object or state of affairs towards which the fear is directed. I am uncertain about the possibility of imagining the

---

<sup>67</sup> see Peacocke (1985)

<sup>68</sup> An influential account of the intentionality of emotional states is given in Goldie (2000)

experience of being afraid without also imagining the experience of whatever the state is directed towards. My intuitions have it that *if* it is possible to merely imagine experiencing being in the affective state one would have if one were afraid, without imagining some object towards which the fear is directed, then one is simply not imagining the experience of being afraid; that fear is intentional is, I think, one of its essential features.

It should be noted that I do not claim that Avon must necessarily *visualise* the bear that is the object of his imagined fear. In order to empathise with Omar, Avon *does* need to experientially imagine fearing a bear, but it is not clear to me that this essentially involves the sensory imagination of the bear that is the object of the fear. I can, after all, imagine being angry with my father without visualising, or sensuously imagining of any sort, that I perceive him doing something to anger me. I am still, in such a case, experientially imagining being angry at my father though, and not merely experientially imagining being angry and propositionally imagining that the imagined anger is directed at my father. There is clear introspective evidence to be found on this point, for if I imagine the experience of being angry at my father for forgetting my birthday, and then the experience of being angry at my roommate for the same offence, that most minimal change in content noticeably changes the emotion as I imagine feeling it. It is not the case that each time I am simply imagining being angry and then tacking on an intentional object via a propositional imagining, but the feeling of each is itself partly determined by its object. The intentional object of the emotion can, therefore, be safely considered a part of what is experientially imagined, even if the perception of the object is not a part of that imagining.

Even if we do not want to take such states as fear to be essentially intentional, for purely practical purposes it is plausible that we should expect the empathiser to include in the content of her imagining some details of the situation to which the imagined state is a response. In

the interests once again of making empathy useful, it seems that any advantageous understanding of another's feeling must include some details of how it came about. How much more useful than simply knowing that Omar is afraid is knowing the kinds of situation that make Omar react that way? If indeed the point of empathy is to come to some sort of understanding about Omar himself, rather than one of his many temporary affective states, then it follows that what would be interesting about this particular empathic experience would not be that Omar feels fear, but that he feels fear in this type of scenario. A simple experiential imagining of only the affective component of Omar's fear divorced from any object, even if it were supposed it could be so imagined, could not be described as revealing anything whatever about Omar besides his current affective state. All Avon could learn from such an imagining is what it is like to feel as if one is afraid, something Avon may well need to know already in order to simulate the fear.

Interestingly, while it is (arguably) impossible for Avon to imagine feeling Omar's fear without also imagining that it is fear-of-a-bear, if Avon is experientially imagining being in the situation to which Omar is responding it does not appear to be necessary that he thereby imagines being afraid. So, while Avon cannot experientially imagine being afraid without thereby experientially imagining being afraid-of-something, his experientially imagining being in a bear-facing situation does not entail that Avon imagines any affective or emotional response to that situation whatever.

This tells us something interesting about experiential imagining in general; that imagining an experience of an F does not necessarily involve imagining a *full* experience of an F. By this I mean that imagining an experience of a given object, or a given scenario, does not entail imagining that experience in all sense modalities; imagining seeing an F does not necessarily involve imagining hearing an F, smelling an F and so on. Given this, why should we think that

imagining an experience of an F entails imagining one's affective response when confronted with an F? Although the imagined experience will be a unified one, it seems that the sensory content of any given imagining is subject to the imaginative engagement of those sense modalities that would, in the case of a real experience, be responsible for the perceptual content of that experience. In short, if Avon were to imagine himself in Omar's situation, that of facing a bear, he does not thereby necessarily imagine an affective response to that situation. So, even if Avon knows for a certainty that Omar, when confronted with the bear, is afraid of it, Avon is at liberty to experientially imagine being in Omar's situation without also imagining that fear. If, however, Avon is just imagining Omar's fear, it seems that the bear-facing situation must make some appearance in the content of the imagining, if only as the intentional object of the imagined fear.

This is important because it means that there are at least two distinct imaginative processes that need to be engaged in order to complete Avon's imaginative project. Avon must engage what we might call his somatosensory imagination in order to reconstruct the feeling that he takes Omar to be experiencing, and he must also engage some variety of imagination that will provide an imagined object towards which the fear is directed. This will, in the most successful cases of empathy, be one or more varieties of sensory imagination (though not in all or even most cases of empathy in general). That is because, I think it is reasonable to assume, the more sensuously replete the imagined situation is the more fine-grained and intense the imagined affective response to that experience is able to be.

These two distinct kinds of imaginative process I will call *feeling-with* and *doing-with*. Feeling-with is imagining being in the same affective state as another, while doing-with is imagining being in the same situation as another. Feeling-with and Doing-with are both kinds of *proto-empathy*; they are the imaginative projects we must undertake

in order to properly ground a demonstrative ascription to our target. By understanding empathy to consist in more than one kind of imaginative process we can begin to see how reconstructive empathy can be a reliable method of coming to understand a target's affective state. Let us modify the case of Avon and Omar a little so that Omar faces the bear alone, then runs away and meets Avon a short time later. Suppose Omar tells Avon he has just seen a bear, and Avon perceives that Omar seems more than a little shaken. Where Avon has not seen the bear himself, it is perhaps less certain that Omar has seen a bear, and certainly less certain that Omar's fear was in response to the bear. However, that is still probably a reasonable set of beliefs for Avon to adopt in this situation. Now, obviously Avon does not need to empathise with Omar in order to make the judgement that Omar's apparent state, fear, is an affective response to his reported situation, facing a bear. Avon may, however, want to attempt empathy with Omar in order to understand *what it is like* to be afraid of a bear. Remember that our problem in this section was how to ensure that the affective state brought about in reconstructive empathy can reliably track the affective state of the target, but here we must be a bit more specific about what that means.

## 2.2 Making Reconstructive Empathy Reliable

Without empathising, but with the information about Omar's situation and physical state, we could justifiably ascribe to Omar the state of fear. However, when Avon empathises with Omar he ascribes not the *type* of state he experiences to Omar, but the *quality* of that experience. At the conclusion of Avon's empathic project, he will not necessarily be in a position to say that Omar feels angry, but rather that Omar feels *like this*. The goal of empathising is to be able to make this kind of demonstrative ascription; to arrive at an

understanding of what it is like to be the target in their current state.

How is that kind of ascription justified by an empathic imagining?

One might think that the justification lay in some tacitly accepted principle of analogy; whereby if I imagine feeling as you do in your situation, the quality of my feeling can be judged to be reliably similar by virtue of the fact that we operate with relevantly similar psychological mechanisms. This can't be the whole story. It is a common thought among empathy theorists that reconstructive empathy requires what is known as a 'self-other distinction'.<sup>69</sup> That is, if Avon is to gain some kind of understanding of Omar by empathising with him, he find some way of quarantining his own psychology from the imagined stimulus and response of Omar's situation. If we recall once again the variation of Avon's story where he is not typically afraid of bears, we can see why this must be the case. Avon is not generally afraid of bears or bear-facing situations, so if he imagines from the inside the experience of a bear and the experience of being afraid of a bear there would be something incongruous, possibly even incoherent, about this imagining. Just as Elizabeth Anscombe and others have observed the difficulty inherent in imagining desiring a saucer of mud if there is nothing about mud that I find desirable; by the same token it seems difficult to conceive of imagining being afraid of bears if one did not find anything about bears fearful.<sup>70</sup>

It is important to consider the end goal of such imaginings. The empathic imagining is intended to produce understanding, to solve a mental puzzle that will lead us to knowledge of the experiences of others. The imagination of the saucer of mud, likewise, may be intended to lead me to an understanding of what it might be like to eat it. Both kinds of imagining plausibly involve measures of aesthetic sensitivity, but in the case of empathy that aesthetic sensitivity is

---

<sup>69</sup> Coplan (2009, 2011), Goldie (1999) and (2000)

<sup>70</sup> Anscombe (1958), Goldie (2006)

tuned to determining how ‘right’ or ‘wrong’, ‘pleasing’ or ‘not pleasing’, ‘congruous’ and ‘incongruous’ are a given combination of a feeling-with, doing-with and being-with. Although there are ways that we can appeal on a cognitive basis to the fittingness of our proto-empathic imaginings (such as whether a fear of bears is consistent with being unafraid of a bear), we can also judge the level of harmony between the various elements of an imagined experience by exercising our aesthetic sensibilities.

For these reasons, I am tempted to understand empathic incongruity as an aesthetic sensitivity to the consistency of a set of proto-empathic imaginings. I believe that, conceived this way, empathic incongruity will turn out to be familiar to us from at least two excellent sources. The first of these is John Dewey, for whom aesthetic sensibility plays a crucial role in unifying ‘experience’ (of sounds, images, feelings etc.) into ‘experiences’; that raw experience wrought into some narrative form through which we can understand it as a unified event. Dewey’s experiences are typified by “internal integration and fulfilment reached through ordered and organised movement”.<sup>71</sup> Indeed, for Dewey this aesthetic character, exemplified by ‘integration’ and ‘organisation’, pervades all experience, and has a special place in intellectual inquiry. It is the aesthetic sense of tension that exists when there is a problem to be solved, that motivates the drives and directions one takes towards the solution, and it is the aesthetic sense of resolution that brings that problem solving experience to a close. It is the aesthetic, for Dewey, that drives the sense of progression through any inquiry or the solving of any puzzle, which will certainly include the empathic puzzles that are our concern here. If, in the end, we were to say that empathic ascriptions are justified by virtue of their aesthetic qualities, we would be keeping them in good company.

---

<sup>71</sup> Dewey (1934)

The second source is Immanuel Kant, who writes in his discussion of humour in the third Critique:

*“In everything that is to excite a lively convulsive laugh there must be something absurd (in which the understanding, therefore, can find no satisfaction). Laughter is an affection arising from the sudden transformation of a strained expectation into nothing. This transformation, which is certainly not enjoyable to the understanding, yet indirectly gives it very active enjoyment for a moment. Therefore, its cause must consist in the influence of the representation upon the body, and the reflex effect of this upon the mind.”<sup>72</sup>*

In this passage Kant is endorsing something like the still popular incongruity theory of humour, according to which something is funny when it creates and then resolves a tension in understanding.

Although he does not use the term, I think the notion of the incongruous as ‘that in which the understanding can find no satisfaction’ is an illuminating one to apply to the case at hand.

Compare again imaginative resistance, where our imagination strains against our will, with a feeling felt-with alongside an incongruous character trait in the imaginer, such that the empathiser would not usually feel that feeling. In the latter case, the imagining seems to get off the ground easily enough, but there is something in tension, something ‘unsatisfying’ about holding it in mind. If empathy were a joke this tension, this unsatisfying element in the experience, would collapse in on itself, leaving only laughter. Empathy is, perhaps obviously, more profitably compared to a conventional puzzle than a joke. Empathy is, as we have described it, a route to understanding, or even a mode of inquiry. We have a question to be answered (‘what is it like to be Omar at this moment’, for instance) a set of tools and a method to follow. We have puzzle pieces (the imagined experiences), a puzzle box (the empathiser’s imagination), and there are only a very

---

<sup>72</sup> Kant (1911 [1790], First Part, sec. 54)



few correct ways to arrange the pieces in the box. Our sense of empathic incongruity is analogous to feeling the pressure as one slides a piece in the wrong slot, or seeing the inelegant pattern that results; the tension it detects is the disharmony of ill combined proto-empathies. This disharmony, ‘unsatisfying to the understanding’ as it is, motivates us to alter our imaginings until a satisfyingly harmonious experience is the result.

This kind of incongruity in imagining, or rather the lack of it, is what serve to justify the demonstrative ascription of the empathiser’s experience to their target. If Avon were generally afraid of bears, then the attempt to imagine experiencing being afraid in response to a bear would not be accompanied by that sense of incongruity. This is because there would not be any relevant difference between Avon and Omar’s character (which I take to be something like the mental dispositions that, among other things, constrain their reactions to any given situation<sup>73</sup>). This in turn means that the quality of Avon’s imagined experience of the bear is likely to closely resemble that of Omar’s. This justificatory power of the congruity of proto-empathies can itself be justified simply by considering how difficult it would be to have congruous proto-empathic imaginings that could not be truthfully ascribed to their target. Here is where an inference from analogy comes into play: while we cannot take for granted that Avon and Omar’s psychologies function in analogous ways when it comes to predicting whether they will react in the same way to the same situation, we can take for granted that people, in general, do not have such incongruous inner lives.

Now, in the case of Avon, Omar and the bear it could be wondered whether the possible understanding Avon will gain from this project will be worth his effort. If Avon understands that Omar is feeling ‘like this’, where ‘like this’ refers to Avon’s imagined fear towards the

---

<sup>73</sup> Following Wollheim (1984) and (2000), Goldie (2006) and (2012)

imagined bear, and it is no different to how Avon himself would react in that situation, what has Avon learned? He has learned that Omar has responded in much the same way as Avon himself would in the same situation, which could be useful to know, but probably not very surprising to learn. Perhaps this is not the most interesting of empathic projects, but it is hardly a typical case; one would generally not need to embark on a difficult imaginative project in order to understand what it would be like to be afraid of a wild bear. However, if we take, for a final time, the scenario in which Avon is *not* typically afraid of bears, what must Avon accomplish in order to rid his imagining of the incongruity that blocks his justifiably ascribing that state to Omar?

### 2.3 Making Automatic Empathy Useful

The sub-personal, automatic route to affective matching fails to meet the standards of usefulness in a theory of empathy in two ways. First, there seems to be no available practical reason why anyone would think that a feeling that they themselves felt did not originate in them, but rather in someone else. Secondly, even if one did believe that one's affective state did not originate in oneself, how could the demonstrative ascription of that state to another ever be justified?<sup>74</sup> Such concerns are supposed to be blocked by the fourth condition in Vignemont and Singer's definition of empathy, that one must know that one's affective state originated in the target of the empathising, but the possibility of that condition being met in any given case seems, *prima facie*, a very small one.

---

<sup>74</sup> Since automatic empathy is not supposed to involve any reconstructive imagining it cannot straightforwardly avail itself of the lack of incongruity justification described above.

It could be argued that I have set the bar for usefulness too high. Perhaps empathy need not be useful in the sense that it adds to our understanding of the inner lives of others. It is doubtless plausible that there are many evolutionary advantages to the tendency to spontaneous sharing of affect with our fellows, even if those affective states are never thought of as having originated in anything other than the 'usual' way. Even so, it is hard to swallow the idea that simply having a feeling that was caused (in a particular way) by the feelings of another, with no intentional state to link the feeling to that other, counts as empathy at all.

I can, however, hypothesise two scenarios that could plausibly lead to somebody having a reason to question the source of their affective state. As far as I know there is no empirical evidence that speaks to the plausibility of these explanations. Perhaps neither of them ever actually occurs. However, if we are to preserve the notion that the mirroring route is indeed a form of empathy, then the following suggestions seem like they would be worth considering.

**Scenario 1:** O is angry, S becomes angry via a mirroring process. S's anger is incongruous with regards to certain other features of S. S begins to reason counterfactually about her own psychology, asking herself questions such as "what kind of person would I have to be to respond in this way?" S reaches the conclusion that she would have to be very much like the sort of person O is and, knowing what sort of person O is, ascribes the state to O.

**Scenario 2:** O is angry, S becomes angry via a mirroring process. S's anger is incongruous with regards to certain other features of S. This incongruity leads S to think in general terms about her anger, and in the course of this wonders if anyone present might be feeling the same way. S imagines herself as each other person present in order to satisfy her curiosity and, upon finding that her state could plausibly be occurring in the psychology she imagines of herself when she is imagining herself to be O, ascribes the state to O.

In both suggested cases I am once again employing the notion of ‘congruity’ of affect with situation and character, though this time its epistemic role is to motivate the whole empathic project, and not just a belief in the verisimilitude of the imagined experience. The ‘features of S’ that I mentioned could be features of either S’s character or situation that would make her mirrored anger appear as an unexpected, perhaps jarring, feature of her experience. Finding oneself experiencing an unexpected feeling does seem to prompt one to, and give one a reason for, thinking about it in general terms (“Why do I feel like this?” “Why would one feel as I do?” etc.). Although it is also plausible that in both cases many people would just dismiss their feelings as irrational, or as the product of some underlying mood that they had not previously attended to. However, someone who is not given reason to be curious about the feelings of others is someone who does not have a reason to empathise, and scenarios such as the ones above seem to provide such a reason. Someone who is not made curious about their mirrored state, because it is absolutely congruous with their own character and situation, is not empathising and nor should we suppose they would try. Not only would they have no reason to doubt that the affective state in question was generated by their own psychologies in the usual way, but it would likely be of benefit to them to consider it as such. If those around you are angry, afraid, joyous, etc., and there is nothing in your situation or character that would conflict with your finding yourself with the same feeling, then it is probably appropriate for you to have that state at that time. That state would neither prompt nor form a part of an empathic process, since it would never be considered anything other than a product of your own psychology.

I think that scenarios such as these two could plausibly occur with some regularity. If one undergoes  $n$  mirroring episodes per day, then I am suggesting that only and all those that feature incongruity will be candidates for empathy, while all those congruent with one’s current

situation will just be taken as one's own affective responses. Since I can see no way that congruent feelings stemming from mirroring processes could be candidates for empathy in any case, I believe I have provided the largest possible set that could serve some role in an empathic imaginative project.

One important thing to note, however, is that both of the scenarios I put forward involve the use of proto-empathies in order to justify the concluding ascription. Indeed, this is true of any scenario that could both plausibly occur and involve demonstrative ascription (as opposed to the mirrored feeling prompting a purely cognitive ascription; 'Omar feels that he is afraid' rather than 'Omar feels like *this*'). Any scenario in which the bare bones of an empathic understanding are gained from a mirrored response (that is, an understanding based on the sensory experience of the mirrored feeling) will necessarily involve some experiential imagining against which to test the feeling's congruity. Since the only reason for using the mirrored feeling as a prompt to embark on an empathic project lies in its incongruity with the empathiser's self or situation, it will not become any less incongruous without bringing into imagination some other beings-with and doings-with that may make a better fit. Indeed, no matter how we come by the state that, at the end of an empathic project, we will demonstrably ascribe to our target, it seems the only way to justify that ascription will be to engage in some variety of proto-empathic imagining.

In the light of these thoughts I suggest that automatic empathy should be considered as distinct from empathy proper. Even if we can find a reason to suspect that a feeling given to us by emotional contagion did not originate in ourselves, it seems that the only way to empathically ascribe it to another is to engage in proto-empathic imaginings. Since proto-empathic imaginings are the preserve, are in fact definitive of, the reconstructive route to empathy, it seems reasonable to conclude that when automatic empathy is to behave like empathy, it must

become reconstructive empathy.

#### 2.4 Congruity and proto-empathy

The difference between Avon imagining being afraid of bears when he imagines being confronted with one, and imagining that he isn't (or just importing details of his real character, that he isn't disposed to fear bears, into the imagining) is the difference between in-his-shoes imagining and perspective shifting. Richard Wollheim has marked the distinction between the two in the following way: If I am centrally imagining (imagining from the inside) that I am Sultan Mahomet II entering Constantinople then I might imagine that I am doing as the Sultan did on that day, or I might imagine that I was him doing as the Sultan did that day. In the former case I am imagining 'being in the Sultan's shoes'; I am not imagining that I am the Sultan, and it is therefore coherent that I could imagine coming face to face with the Sultan himself in such a case. By contrast, if I am perspective shifting and thereby imagine not only doing what the Sultan is doing, but being him while he is doing it, it is incoherent that I could imagine coming face to face with the Sultan, because it is incoherent to imagine the experience of coming face to face with oneself. When in-his-shoes imagining we imagine experiencing what another experienced, but when perspective shifting we imagine having the experiences of another person.<sup>75</sup>

Empathy involves the latter kind of imagining. Only then can Avon imagine experiencing Omar's situation as Omar would experience it, some understanding of which is supposed to be the aim of the empathic enterprise. In imagining Omar's experience Avon must adopt, in imagination, all those mental dispositions of Omar's that in

---

<sup>75</sup> Wollheim (1984) Goldie (2006)

some way constrain his experience of the bear; the most obvious of these being the disposition to fear bears. This kind of imagining is the final kind of proto-empathy I will discuss here; it is called *Being-With*. If Avon's feeling of fear during his empathising with Omar (Avon's feeling fear-with Omar) is congruent both with his imagining of himself in Omar's situation (Avon's doing bear-confrontation-with Omar) and his imaginative adoption of the mental dispositions that govern Omar's response to the bear (Avon's being afraid-of-bears-with Omar), then he will be justified in demonstrably ascribing his imagined experience to Omar.

It may seem to some that the requirement that the empathiser adopt in imagination all mental dispositions governing the experience of the target at issue is too strong. Indeed, Peter Goldie has very eloquently articulated this worry, in a discussion of the work of Richard Wollheim. The story of Avon, Omar and the bear is a very simple example of an empathic project. One that in all likelihood nobody would embark upon, as I have already mentioned. Goldie discusses a more complicated case, that of empathising with Salieri, particularly regarding his envy of Mozart. Goldie also stresses the importance of adopting what Wollheim (1974) calls the *repertoire* of the target of one's empathy in imagination. That is, the set of mental dispositions that will constrain his actions in the imagined scenario. Goldie claims, however, that although the repertoire of the target needs to be adopted as a proto-empathic imagining, the practical barriers to succeeding at being-with one's target in a sufficiently robust way make it very unlikely that one will succeed in any given case. His pessimism is motivated by a similar view of the intentionality of emotional states to that which I argued for in section 4.4, that emotions are essentially directed; imagining that one is experiencing Salieri's envy is to imagine being envious-of-Mozart, where the intentional object affects both the content and qualities of the imagining.

Where Goldie becomes wary of the prospects of anyone successfully being-with Salieri is when he notes that the project of empathising with Salieri's envy of Mozart is exponentially more difficult than that of, say, Omar's fear of the bear. This is because:

It is Mozart, in all his particularities, with all his idiosyncrasies, that Salieri envies, not just a successful person in the same field of endeavor who happens to be Mozart but who might just as well have been someone else.<sup>76</sup>

So while we might substitute, in imagination, any old bear for imagining the actual one Omar faced and come out with much the same result, the same cannot be done in the case of empathising with Salieri. We cannot substitute Mozart in our imagination for any old composer, nor even for any young rival upstart from our own life towards whom we are disposed to feel envy. The crux of the issue is that the dispositions that make up Salieri's repertoire are *also* essentially directed, not just the emotional experiences that they effect when combined with particular stimuli. Salieri's disposition to envy Mozart is every bit as directed at Mozart as is the experience of envy that Salieri undergoes when he hears of Mozart's latest triumph.

While I am sympathetic to this worry, I do not think it warrants the pessimism that Goldie seems to take from it. It is true that, since I never knew Mozart, much less have experience of being his bitter rival, I cannot be in a perfect position to imagine having the kinds of dispositions that would constrain Salieri's various emotional experiences of envy. However, I do not think there is any problem with admitting that, while still claiming that my attempted being-with Salieri, my adopting his repertoire to the best of my ability, so long as the resulting imagined experience is congruent and harmonious, will be so far from Salieri's actual experience as to be useless. Indeed, it is reasonable to expect even successful instances of empathy, by which I

---

<sup>76</sup> Goldie (1999)



mean empathic projects that produce some useful understanding of the target, to vary in exactly how accurate the imagined experience is as a representation of the target's own experience. A related worry, also voiced by Goldie as well as Coplan and others, is that most instances of attempted empathy will turn out to be cases of 'pseudo-empathy', where an empathiser is successfully feeling-with and doing-with, but insufficiently quarantines their own character from their being-with, so that the response imagined will be their own and not the target's. Again, I do not believe this poses anything like a fatal problem with the idea of empathy being reliable and regularly occurring. For one thing, if I am right about incongruity playing a justificatory role, any instance where one's own dispositions would be incongruous with the imagined experience will show themselves as such. Flawless empathy, the precisely accurate ascription of the imagined experience to the target is, obviously, very improbable, for the reasons Goldie suggests. However, there can certainly be *successful* cases of empathy that don't quite manage to meet that extremely high level of accuracy; cases where the understanding of the other person gained is still accurate enough to be useful, and still furnishes some genuine, and accurate, new understanding of the inner life of the target.

### 3. CONCLUSION

So, we have arrived at the following conclusions: firstly, that the automatic route to empathy requires some input from the reconstructive route in order to be properly attributed to the target of that empathy. Secondly, in order to reliably achieve an affective match, the reconstructive route to empathy requires imagining from the inside not just the affective state of one's target, but also their situation and relevant aspects of their character. Finally, that congruity of proto-empathic imaginings can serve as a justification for an empathic ascription of the imagined experience to the empathy's target. Vignemont and Singer's definition, although it has been widely endorsed and is *prima facie* a strong one, appears to be insufficient as a conception of empathy that is at once useful and reliable. In order to be useful we must introduce an element of reconstruction into the automatic route to empathy, and in order to be reliable the reconstructive route must involve imagining far more than simply the affective state of the target. Both the processes of automatic and reconstructive empathy have epistemic deficits that block their being useful as a tool for understanding without introducing constraints that make the other nearly impossible to undertake. I contend, therefore, that we should simply avoid understanding automatic mirroring as a kind of empathy.

Empathy, as I have presented it, is a method for coming to understand the inner life of another person. We have seen that it consists in a set of experiential imaginings (proto-empathies) that, by virtue of their congruence in imagining, justify a demonstrative ascription of the form '*the experience that I am imagining having is like the experience that O is having*', or more succinctly, 'the target feels like *this*'.

I want to finish by returning to my claim that, at bottom, no good reason to think of empathy, properly understood, as being solely or principally concerned with affective matching. That is, empathy as we

have come to describe it, comes with no reason to see feeling-with as primary in any sense in relation to other forms of proto-empathy.

This is because the question of how my target feels, or rather, what it is like to feel as my target does, is by no means the only kind of question we can answer using empathy. Consider the following examples:

**Scenario 3:**

Avon seeks to understand how Omar feels.

Avon imagines being Omar in Omar's situation (being-with, doing-with).

In imagination, Avon emotionally reacts to Omar's situation in the way Omar would (feeling-with).

Avon's imagined reaction is not incongruous with respect to his imagined character or situation.

Avon ascribes that emotion to Omar.

**Scenario 4:**

Avon seeks to understand why Omar feels as he does.

Avon has good reason to believe that Omar feels a certain way (verbal report, an automatic response which Avon has already figured out is not his own, etc.)

Avon also knows the situation that prompted this response.

Avon imagines feeling the way Omar does in the situation that Omar is in.

Avon imaginatively reconstructs the necessary character traits that would make Omar's reaction in that situation congruent.

Avon ascribes those character traits to Omar.

**Scenario 5:**

Avon seeks to understand why Omar feels as he does.

Avon has good reason to believe that Omar feels a certain way (verbal report, an automatic response which Avon has already figured out is not his own, etc.)

Avon has good reason to believe he can reliably simulate Omar's character (they have been friends a long time, he has empathised with Omar as in example (3) many times previously etc.)

Avon imagines being Omar and feeling the way Omar does.

Avon imaginatively reconstructs a situation that is congruent with his imaginings of Omar's state and his character.

Avon ascribes Omar the property of having been in that situation.

These scenarios all seem to be consist in imaginative projects we might reasonably expect people to carry out. All use proto-empathies; in each case one has a good reason to believe that they can accurately replicate two proto-empathies and, on the basis of that, is able to determine the congruence or otherwise of their imagined answer to their initial question. Using the same kinds of imaginative projects, and in the end making the same sort of empathic ascription ('Omar's experience is like *this*'), there seems no reason to call scenario 3, where one seeks to discover the feeling of the target, an empathic project, but scenarios 3 and 4, where one already knows or understands how the target feels, but wishes to understand other features of the target's inner life as something other than empathic.

Empathy is not, according to the Complex Theory, all about coming to share the feelings of another. It is all about coming to understand the inner life of another. To do this it is not sufficient that we simply mimic their affective state, nor only adopt their characters or situations as our own in imagination. To arrive at an empathic understanding of another we must figure out not just the states they are in, but also how these states are related to one another. Avon *understands* what it is like to be Omar, in fear for his life, only when

he has some experience of how Avon's encounter with the bear, filtered through his repertoire, results in the fear that he feels. Empathy is more than mimicry, it is a method not just of understanding what people are going through, but also, and more importantly, why they are going through it.

### 3.1 The Complex Theory of Empathy: Applications and Implications

#### 3.1.1 The Moral Value of Complex Empathy

The theory of empathy we adopt will have implications for the role it can play in our understanding of ethics, moral psychology, aesthetics, and the various other issues in which empathy has been suggested to play an important part. Here I want to offer some brief thoughts about how the theory of empathy I proposed in the previous chapters, what I have called the 'Complex Theory of Empathy', fares in comparison to some of the other theories I have discussed when it comes to the moral and other values of empathy that are so commonly attributed to it.

I agree with Helen Battaly<sup>77</sup> that empathy, under any of the conceptions I have discussed (The Perceptual Action Model, the Two-Routes View, and each of the individual psychological processes that were candidates for empathy in section 1.1), cannot constitute a moral nor an intellectual virtue. Empathy, on the understanding of these views, lacks a number of the conditions sufficient for virtue. Virtues are habits or dispositions, indicating what one *would* do under certain circumstances.<sup>78</sup> Empathy on any of the views discussed is clearly not a disposition, it is an action (unconscious, in the case of the Perceptual

---

<sup>77</sup> Battaly (2011)

<sup>78</sup> Alson (1993)

Action Model), a kind of action which we might well say we are disposed to engage in, but not a disposition in itself. If there is anything virtuous about empathy, it must be that empathy is the kind of action we might expect a virtuous person to frequently undertake.

The Complex Theory, however, like any of the other conceptions of empathy as a voluntary, imaginative activity, is not at liberty to say that any empathising will always be a morally good thing. Empathy can, of course, be used for nefarious purposes. If I wanted to know how to undermine your confidence, scare you, or otherwise make you miserable, empathising with you may well give me some understanding of how to make that happen. On the other hand, it is possible that the more demanding conceptions of empathy, such as the Complex Theory, are less susceptible to being used for morally wrong ends than other, more simple kinds. This is because the effort that it would take for me to empathise with you in a way rich enough, accurate enough and deep enough that I would be able to determine your emotional frailties might expose me to enough of your inner life that I would find it difficult to go through with it.

One moral value of empathy on the Complex Theory, however, which I believe it can make use of more than any other theory of empathy I have discussed, is as a way of contributing to the development of the virtue of open mindedness.<sup>79</sup> Empathy with others, according to the Complex Theory, may avail us of the experience of imagining different congruent sets of proto-empathic imaginings, which I would affirm are a way of developing open-mindedness in a very interesting way. Someone who empathises a lot isn't open minded because they necessarily understand how a lot of different people *feel*. However, they might become more open minded in virtue of learning that several different proto-empathic combinations can be experienced as congruent, and therefore that there are several different ways of

---

<sup>79</sup> Arpaly (2011) defends a conception of open-mindedness as a moral virtue.

developing one's character, forming one's beliefs and relating to one's feelings.

### 3.1.2 The Complex Theory, The 'Folk' and Empathy Research

I have now described the Complex Theory of Empathy, by way of describing how one can appropriately ascribe a collection of mental states to a target by engaging in congruent proto-empathic imaginings. I call this view the 'Complex Theory' not because it makes empathy relatively cognitively demanding (although it does), but rather because it takes empathy to involve a complex of imaginative states justifying an ascription. The Complex Theory meets the desiderata of making empathy reliable, useful, and able to occur regularly.

I want to finish this section by revisiting the questions I gleaned from the disagreements between the 'folk' theories of empathy I outlined at the start. I said that a good theory of empathy should be able to address the following questions that were raised by both the 'folk' and current 'refined' theories of empathy:

**Q1.** Does empathy entail compassion, pity or sympathy for the target of the empathy?

**Q2.** Does empathy require that the empathiser feels the same as their target? If so, the same in what sense?

**Q3.** What kind of understanding does empathy provide, and what capacities does that understanding grant the empathiser? Must it, for instance, allow the empathiser to respond appropriately, or to lend moral support to the target?

**Q4.** How is empathy distinct from cognitive mind-reading (deploying a theory of mind to determine a target's mental state)?

On the first question, we can say that no, according to the Complex Theory of empathy, empathy entails neither compassion, nor pity, nor sympathy for the target of empathy. I distinguished empathy from sympathy in section 1.2, and the Complex Theory of empathy requires no attitude from the empathiser to their target, other than a curiosity about their inner life.

The second question relates to one of the key claims of this thesis, which is that affective matching has been over-privileged as it relates to theories of empathy. Empathy, on the Complex Theory, *does* require affective matching, but it is quite clear that affective matching is merely a means to understanding, and not the final aim of empathy.

As to the third question, the understanding granted to an empathiser on the Complex Theory does not necessarily imbue them with any particular capacities. As I mention in the previous section, there is nothing in the Complex Theory to prevent someone using empathy for immoral ends. However, we can say that the understanding granted by empathy on the Complex Theory is of a very useful kind; a successful empathiser will understand the inner life of their target to a significant degree, allowing them to socially (or anti-socially) interact with them in a competent way.

The Complex Theory is also very well equipped to address the final question. Cognitive mind reading, wherein one person deploys a Theory of Mind to understand the thoughts or actions of another, can be clearly differentiated from empathy on the Complex Theory by virtue of the Complex Theory's focus on demonstrative ascription of feelings (broadly construed). Whereas someone who employs cognitive mind reading will always be able to cognise the mental states they ascribe to their target (they will be able to say that Avon is afraid



because he believes that there is a bear), the qualitative nature of the empathic ascription means that an empathiser will not necessarily be in a position to do the same. They will, instead, be in a position to use the phenomenal character of their imagined mental states, created as a part of an empathic project, as a sample with which to understand the feeling of their target.

PART II:  
EMPATHY AND NARRATIVE ENGAGEMENT

*You think your pain and your heartbreak are unprecedented in the  
history of the world, but then you read.*

- James Baldwin

## 4. PERSPECTIVES AND NARRATIVES

### 4.1 Introduction: Proto-Empathies and Narrative Engagement

In the previous two chapters I proposed a theory of empathy, according to which empathy is an imaginative process comprised of three proto-empathic imaginative projects; feeling-with, doing-with and being-with, undertaken with the aim of making a demonstrative ascription to a target of the form *[the target] feels like this*. The rest of this thesis will show that this view of empathy gives us the resources to develop a powerful theory of how audiences engage with narratives. I will argue that engaging with narratives is an essentially empathic activity, and that understanding narratives in this way gives us a host of theoretical resources with which to explain our emotional responses to narratives as well as their effects on our moral and cognitive development.

To show that narrative engagement is an empathic activity, on my terms, is to show that it involves proto-empathic imaginings, undertaken with the aim of formulating a demonstrative ascription to a target. That is, that narrative engagement involves imaginatively occupying the thoughts, feelings and character of another person, with the aim of making an ascription of the form *[the target] feels like this*. In the case of narratives, I will argue that the target of empathic engagement is the implied narrator of that narrative.<sup>80</sup>

It is tempting to think that the proto-empathic imaginings of doing-with (imagining being in the same situation as another) and feeling-with (imagining being in the same affective state as another) would be

---

<sup>80</sup> I intend that what I say here should go for any kind of narrative representation, but I will initially confine myself to talking about literary narratives.

easily incorporated into any account of narrative engagement. After all, doing-with is one way of representing a state of affairs to yourself, and being told a story largely consists in having a series of events represented to you. Similarly, feeling-with is a way of undergoing an affective response differently from how one ordinarily would, and engaging with stories frequently finds us in affective states that seem to differ from our usual pattern of responses.<sup>81</sup> However, merely engaging with representations of events and undergoing affective responses different from those we would normally expect to have are not sufficient for doing-with and feeling-with.

In the first place, doing-with and feeling-with are both kinds of imaginings, and it is not immediately obvious that all engagement with narrative representations is imaginative. However, if we have a suitably broad notion of imaginative activity encompassing such things as supposition and entertaining thoughts for the purposes of examination or enquiry, then narrative engagement can be said to be imaginative in at least that sense.<sup>82</sup> Engaging with a story of any kind, fictional or not, involves at a minimum the entertaining of thoughts concerning the occurrence of the events described, and if this can be described as imaginative then they meet the imaginative criterion for doing-with. Similarly, since feeling-with isn't the imagining of any affective state, but an affective state brought about by a particular (empathic) imaginative means, then provided we can show that an affective response to a narrative has been brought about through some empathic activity, that response counts as a feeling-with.

---

<sup>81</sup> This feature of narrative engagement gives rise to the 'problem of disparate response', which I examine in detail in section 4.5.

<sup>82</sup> Friend (2012) and Matravers (2014) have persuasively argued that fiction cannot be defined by reference to the imagination (contra Walton (1990) and Currie (1990)). I am interested in *all* narratives, however, not just fictions. Indeed, the fact that all narratives plausibly involve imagining to some degree is a part of the reason Friend and Matravers are sceptical that imaginative involvement can be the *sine qua non* of fictional narratives as opposed to non-fictional ones.

The crucial feature of feeling-with and doing-with that marks them out as proto-empathic imaginings, rather than discrete imaginative projects, is that they are undertaken with the aim of giving a demonstrative, qualitative ascription to a target (that the target *feels like this*). This is what distinguishes doing-with from more simple ways of imagining events. As such, for engagement with narratives to count as an instance of empathy, we have to be persuaded that the imaginings I have mentioned (imagining a series of events and having affective responses brought about by those imaginings) are undertaken with the aim of making a demonstrative ascription to a target. This is even less obvious than the claim that engaging with narratives is essentially imaginative, but both can be persuasively argued for. The latter, as I have suggested, merely falls out of the stipulation that we should consider supposing or entertaining the truth of various kinds of thoughts for the purposes of examination, enquiry or pleasure as ways of imagining. The former claim, however, that engagement with narratives aims at a demonstrative ascription of a mental state to a target, requires a more robust argument.

#### 4.2 Being-With and the Self-Other Distinction

One way to argue that narrative engagement involves an empathic demonstrative ascription is to show that narrative engagement involves the third variety of proto-empathic imagining, *being-with*. Being-with is the act of imagining having the character of another, and identifying it as such, with the aim of making a demonstrative ascription. Feeling-with and doing-with are both imaginative activities that aim at a demonstrative ascription, however they are both kinds of imagining that can be undertaken in an etiolated, non-empathic form. I can, and often do, imagine series of events without thereby intending to ascribe the experience of those events to another.

Similarly, I can also easily undertake an imaginative project that results in an affective response that is outside the realm of my usual response pattern without ascribing the resulting feeling to another. In both of these cases I am imagining myself being in states other than those in which I naturally find myself. Being-with, on the other hand, is a kind of imaginative project that isn't amenable to being undertaken in a similar, non-empathic form. If I am imagining having a character different from my own, then in several crucial respects, I am imagining myself being a person who isn't me; in other words, imagining being in the possession of a different character entails the demonstrative ascription of the imagined experience of having that character to a real or hypothetical figure who 'actually' possesses it.

To see why this is true we must consider the experience of being-with imagining, compared to the experience of doing-with and feeling-with. The experiential content of being-with imaginings is response dependent in a way that the contents of feeling-with and doing-with imaginings are not. While I can imagine having a different character to that which I possess ordinarily, there is no phenomenal character to that imagining unless I carry it out as a part of a wider imaginative project that gives my imagined character traits a situation in which to manifest themselves. Imagining myself as a kind of person who has a character more generous, more conservative or more snobbish than my own will, by itself, yield no experiential content unless I imagine a situation in which my imagined generosity, conservatism or snobbery can manifest themselves through different ways of understanding or responding to that situation.

This means that if I imagine having, say, a more generous character than I in fact possess (perhaps as generous as some particular person, perhaps not), it can only be experienced if I imagine having a more generous character in a particular situation. Let us say that I am imagining having a more generous character while offering to buy the first round at the pub. What can we say about my attitude to this

imagining? I am likely to be surprised at how happy I am to make the offer of buying the first round. Perhaps my usual character is so *ungenerous* that I am even surprised that the scenario of my buying the first round seems congruent with my character. These are both responses it would be reasonable to expect me to have to this imagining, but importantly both of these responses require me to maintain a *self-other* distinction (recognising that some of my imagined feelings stem not from my ordinary psychological processes, but rather from a set of traits adopted in imagination). I am surprised, so I recognise that my imagined experience is not what I would expect of my usual experience. I am, of course, in one important sense imagining myself in the pub. However, I am not under any illusion that my real self would respond so graciously to buying the first round; I know that a self-other distinction is appropriate. Indeed, if I am imagining myself as having a different character I must maintain a self-other distinction, otherwise my imagining will collapse into my simply imagining being in a situation with a character that is *in fact* different to my real one, but which I don't *recognise* as being different. In such a case I can't be said any longer to be engaging in a being-with imagining, since I am not imagining having a different character. I am instead mistakenly attributing to myself in imagination a more generous character than I, in fact, possess.

To see the importance of a *self-other* distinction, we need to recall some of my earlier discussion of empathy between *in-his-shoes* imagining and *perspective shifting*. *In-his-shoes* imagining is a form of what Goldie and Coplan termed 'pseudo-empathy', whereby I imagine being in the same situation as my target (doing-with), but due to the fact that I don't imagine having the same character as my target (being-with), the resulting imagined experience does not mirror the experience of the target, and so the empathic project fails. If I am attempting to empathise with Avon, who is confronted with a bear, but I am not afraid of bears and I fail to imaginatively modify my bear-

facing dispositions, then the demonstrative ascription I make to Avon will not be accurate. *Perspective Shifting* is required for successful empathy, where the empathiser engages in a being-with imagining, thereby modifying the relevant features of their character and enabling the empathiser to respond to an imagined scenario as their target would, rather than as they themselves would. A second crucial feature of perspective shifting, as well as the imaginative modifications of the relevant mental dispositions in the empathiser, is that the empathiser maintains a clear *self-other* distinction; meaning that the empathiser is able to distinguish between the states she is experiencing as a part of her empathic project from those which she is experiencing as a part of her standard pattern of psychological activity. Because any being-with imagining can only have experiential content when it is coupled with imagining an event that my adopted character can respond to, and any such imagining also requires me to maintain a self-other distinction, any being-with imagining entails making a demonstrative ascription to a target. I must maintain a self-other distinction in order to prevent a being-with imagining collapsing into a different kind of imagining altogether, wherein I am simply imagining myself as I am in a situation, but am mistaken about my character. In order to maintain this self-other distinction, I must postulate an *other*. Furthermore, I must postulate that other with the aim of ascribing to them my experience in an empathic fashion. In a being-with imagining I know that the experience *feels like this*, but I also know that the experience does not *feel like this* to me, by virtue of my self-other distinction I know that it *feels like this* to someone else. A brief sidebar is necessary at this point to address a possible concern about the way I am talking about making a self-other distinction between two versions of myself. It isn't difficult to see how there might be something a little troubling in the idea that I can imagine being a more generous version of myself who I can nonetheless consider an 'other'. The worry is simple to answer, however, by



clarifying that the 'self' and 'other' distinguished by the 'self-other' distinction doesn't strictly refer to any metaphysical self, of the sort that philosophers might worry would not survive a brain transplant. Rather, the self of the self-other distinction is merely the empathising agent as they are at the moment of their empathising, including and especially their ordinary psychological mechanisms and dispositions. All we need to capture in the self-other distinction is the idea that an agent engaging on an empathic project should recognise that their responses during that project may be the result of the imagined adoption of a different set of psychological dispositions, and not of their ordinary psychologies. Those new dispositions can just as easily belong to an alternate version of oneself (past, future or purely hypothetical) as to any other real or hypothetical figure. Indeed, it is highly desirable for an account of empathy to be able to accommodate such cases, in order that we can use empathy as a way of describing how we are able to feel alienated from our past, future or hypothetical selves, while at the same time recognising that those past, future or hypothetical selves are nonetheless continuous with the self who feels alienated.<sup>83</sup>

These considerations help us to clarify the project of the bulk of this chapter: to show that narrative engagement is an empathic activity by way of showing that it involves being-with. Since being-with is a kind of imaginative activity that essentially involves the postulation of an other, and cannot be undertaken in an etiolated form without the postulation of an other (unlike feeling-with and doing-with), if it turns out that narrative engagement routinely involves being-with imaginings, then by virtue of that we can confidently say that narrative engagement is an empathic activity with some hypothetical or actual figure who holds the imagined character as the target. The target of the ascription needn't be an *actual* person (it is completely coherent,

---

<sup>83</sup> Some philosophers who have relied on similar thoughts include Velleman (1996), Schechtman (2005) and Goldie (2012)

for instance, to empathise with a hypothetical version of yourself with significantly different character traits). However, to imaginatively adopt a character is to ascribe that character to a real or hypothetical persona in a way that to imagine a series of events does not require one to ascribe the experience of those events to an actual or hypothetical agent.

The second half of this thesis will proceed in three main parts. First, I will discuss the phenomenon of perspectives, and how appealing to them can provide powerful explanations for familiar features of our reading stories. Next, I will clarify and expand upon the notion of a 'narrative perspective', its relation to a hypothetical author, and how it can serve as a target for empathy in narrative engagement. Finally, I will conclude with some thoughts on how understanding engaging with narratives as an empathic activity gives us promising new avenues within which to explore the philosophy of the narrative arts.

#### 4.3 Perspectives

A good story can make us laugh, cry, or think about things that we have never laughed, cried or thought about before. One way a good story can become a great story is when it is capable of making us laugh, cry and think about things that we ordinarily *could* not laugh, cry or think about. That is, stories can make us laugh at things we would ordinarily cry at, cry at things we might usually find inconsequential, and think about people and places in ways we ordinarily never could; not just because those people and places are perhaps the fictional inventions of some creative mind, but because we are not the sorts of people who would ordinarily think of those things in those ways.

Here I want to describe the abilities stories have in affording us different ways of thinking about the world; the way they can afford us

the opportunity to engage with different perspectives and ‘ways of seeing’. I will also investigate the ways that stories can accomplish these various feats and argue that engaging with the perspectives that stories provide are instances of empathising, with an implied storyteller as the target of that empathy.

Throughout my discussion of the powers of stories to help us experience different ways of seeing, I will develop an account of what I will call the ‘narrative perspective’. Briefly, the narrative perspective of any given story is the general psychological outlook which it would be reasonable to ascribe to a flesh and blood teller of that story; it is the narrative perspective of a story that makes the story itself gloomy, cynical, optimistic or romantic, or that gives the story the capacity to frown on or applaud different characters. I will argue here that not only *can* we talk of stories having perspectives in this way (distinct from the perspectives of their flesh and blood authors, audiences or fictional narrators) but that we must do so in order to properly explain some common kinds of response to familiar kinds of stories. Before I move on to discuss narrative perspective, however, it will be helpful to get a clearer idea of what perspectives in general are, and how they feature in our experience of the world at large.

#### *4.3.1 Perspectives and Ways of Seeing*

Perspectives are something we refer to fairly freely in folk-psychological explanations of things like moral and aesthetic disagreement, or in accounting for why people are disposed to attend to different features of social, political or other kinds of events. We might diagnose a disagreement over hate speech legislation as being due to the fact that one person has a broadly liberal perspective on such things, and the other has a more conservative perspective: Fiona has a politically conservative perspective, and so disapproves of hate speech legislation, whereas Siobhan’s is more liberal, and so is more approving. Of course, there may be reasons other than their respective political perspectives why Fiona or Siobhan might approve

or disapprove of hate speech legislation. If one of them was a political philosopher, they may have come to their conclusion in a much different way. Perspectives are not the only way we form beliefs, and certainly not the only thing that determines our responses to situations, but they are what is responsible for how a situation, event or object *strikes* us; perspectives govern at least our immediate responses, and if those responses are not reflected on they will usually remain a principal determinant of them.

It is this kind of perspective I want to give an analysis of, and particularly the perspectives of this kind that play an important role in the way audiences respond to stories; they may not govern our *final* responses to a story after serious critical reflection, but they do determine how the contents of a story strike us. It is important at the outset to distinguish this sense of perspective from another common meaning; a perspective simply as a spatiotemporal location in a scene (or the appearance of the objects in that scene for an observer in that location). The perspectives that I'm interested in are not visual phenomena, although they are sometimes referred to as if they were, not just as perspectives but also as 'points of view', 'outlooks' and 'ways of seeing'. The common visual terminology is a little misleading, but also illuminating in an important way. Take the following example:

If two people are standing at opposite sides of a statue they will each see it in a different way (they will attend to different physical features of the statue, they might interpret its meaning in different ways due to that fact, and so on). Similarly, if one has a patient and thoughtful character, but the other is impatient and dismissive they will likely respond to the statue in very different ways then too (the former might think it provocative while the latter dismisses it as childish). If we take the 'seeing' in 'ways of seeing' (or the visual association with the term 'perspective') to pick out not just the visual experience of a subject but rather the way the world seems to them in a more general sense, we can start to get an idea of how perspectives of these kind can

influence us. The statue-viewers each have different ways of seeing the statue in at least two important senses, both of which can be a significant influence on their response to the statue; they have different ways of seeing the statue in the sense that is due to their having different spatial locations relative to it, and they also have different ways of seeing the statue in the sense that they are each psychologically primed to respond to statues of this kind in different ways.

These latter, psychological perspectives are what I want to show are central to a good understanding of narrative engagement.

Perspectives are a very broad phenomenon, governing a great deal of our psychological lives. Perspectives are psychological dispositions peculiar to a subject that affect the way the world appears (in a suitably broad sense) to them. They are one reason that Fiona responds one way to the hate speech proposal, seeing it perhaps as an attack on individual liberty, whereas Siobhan sees it as a means of overcoming ingrained prejudice. They are also a reason one person sees a statue as provocative and the other as childish. These *ways of seeing* are the effects of perspectives on our experience; they are what is responsible for making some features of a scene strike us as more central or prominent, make us prefer some kinds of explanations over others, and leading us more broadly to organise our thoughts about subjects and events in particular ways.<sup>84</sup>

#### *4.3.2 Perspectives and Character*

Features of our characters, as the statue case example shows, can help to determine our perspectives (and, by virtue of that, our ways of seeing). Thinking again in general terms, a person of optimistic character is likely to see things in ways that a pessimist will not. An optimist takes a development as an opportunity what a pessimist takes

---

<sup>84</sup> I owe the conception of perspectives as involving organising our thoughts about a subject to an excellent, though unpublished manuscript by Elisabeth Camp.

as a crisis; they see the same proverbial glass as half full that the pessimist sees as half empty. Importantly, however, a person's perspective is not identifiable with their character or personality. When philosophers speak of character, usually in the context of virtue ethics, they mean something like traits that dispose one to consistently act in a certain way (someone of a generous character can be expected to behave generously in various situations, for instance). This notion of 'character' is very broad, and a person's perspective may form a part of what we call their character (in that it is one thing that might lead someone to consistently act in a certain way). However, it is still helpful to prise apart the notion of perspective from the broader idea of character because while one's character plausibly consists in all one's stable psychological dispositions, a perspective consists in a narrower class of psychological traits which govern the way the world appears to the subject (their *way of seeing*), and thereby indirectly govern the way the subject acts in the world.

Let me use the example of generosity further to illustrate those parts of our character that I think ought to be thought of as constituting our perspective. To ascribe someone a character trait (like generosity) might be to simply make a claim about their behavioural dispositions (perhaps that they will tend to offer to buy the first round at the pub). It is often the case however that calling someone generous picks out more than this including, for instance, that they tend not just to buy the first round at the pub, but that they are happy to do so. Neither the behavioural disposition nor the disposition to hold a certain attitude to that behaviour are elements of a person's perspective, though both of them are parts of a person's character. To get at our generous person's perspective we should look at what features of their character might *cause* our generous person to be happy to buy the first round. The cause of their happiness is surely something to do with the way the idea of buying the first round presents itself to them, not as an onerous financial burden but perhaps as an obvious and easy

way to make their fellow drinkers comfortable. A generous perspective is here manifesting itself in this *way of seeing* the opportunity to buy the first round of drinks. Doubtless perspectives of this sort form part of what it is to have a generous character, but it is an importantly distinct part. If the possession of a generous character can be seen as possession of a collection of various dispositions that tend to result in generous behaviour and attitudes (which I think it must be), then a generous perspective is the subset of those dispositions which is responsible for the world appearing to the generous person as providing frequent occasions on which it is appropriate to behave in a generous fashion. Similarly, a person of a pessimistic character will have, as part of that, a pessimistic perspective consisting in dispositions responsible for making the world appear such that pessimism is often appropriate to it; for every character kind there will be a corresponding perspective kind responsible for the way the world appears to a person with that sort of character.

Now we have something of an idea of how I want to characterise perspectives and their function in the broader psychology of personality. Next I want to say a little more about how they fulfil this function. I have said that a perspective of any given kind is responsible for making the world appear a certain way to its holder, but what exactly does it mean for the world to ‘appear a certain way’? And how do perspectives achieve this?

#### *4.3.3 Generating Ways of Seeing*

The role of perspectives is to generate ‘ways of seeing’. Recalling the statue example, a ‘way of seeing’ in my sense (*a perspectival way of seeing*) is analogous in some ways to a more strictly visual ‘way of seeing’ resulting from the position of a viewer relative to an observed object or scene. For our statue viewers, their differing perspectival ways of seeing resulted in one thinking the statue challenging and the other thinking it childish, why might this be? To answer, let me first

suggest how exactly a statue might appear challenging to one viewer and childish to another. Suppose that the statue in question is Henry Moore's *Horse* (1984).

Why might people have two radically different responses to the same statue? The first viewer, in finding it childish, might plausibly have attended more to features like the Horse's very simple facial features, its slightly bulbous outline, apparently broken tail and rough, unfinished surface. The second viewer, finding the statue challenging, might attend rather to the way the graceful curves clash with the sharp edges of the broken tail and legs, and the fact that there is little that is distinctively equine about the figure other than the oddly simple head and grossly exaggerated withers and flank, perhaps emphasising the similarity between the muscled figure of a horse and other familiar Moorean subjects (such as nude human figures). Having given prominence in their thoughts to these different features of the statue the viewers would certainly appraise the statue in different ways; the first viewer thinking, for instance, that the simplicity of the head is a sign of absence of care or skill in the sculptor, the second that it is a sign that the sculptor thinks the head is unimportant other than as a signifier that the figure represents a horse.

The different responses of our viewers of the statue are caused, in this case, by certain features of the statue striking one or the other viewer as more prominent or significant, and by our viewers tending to seek different sorts of interpretative explanations for those features that they do find prominent. We might say that they have different ways of seeing or characterising the statue; they do not respond differently because they have different visual access to various features of the statue, rather they have different ways of organising their thoughts about the features of the statue to which they share visual access. Their perspective on the statue governs not just what they notice, but



how they are disposed to characterise the statue and its sculptor, and by virtue of that, the response they are likely to have.

Generalising from this example we can start to see the usefulness of this characterisation of perspectives as mechanisms for generating ways of seeing objects and scenes we encounter in the world. Recall once again the case of the generous person who sees the opportunity to buy the first round of drinks as an easy way to make her companions comfortable. Our generous person here has a distinctively generous way of organising her thoughts about the opportunity to buy the first round of drinks; the fact that doing so will make her drinking companions comfortable is a more prominent feature of the scene to her than the fact that she will likely bear a greater financial burden than some of those others present. We can characterise all the other examples of perspectives I have mentioned in similar ways too; the pessimist sees an event in such a way as to emphasise its negative consequences for her, and will tend to give explanations for those events that perhaps rely on ascribing cynical motives to others; the liberal will tend to explain perceived injustices in terms of group rights rather than individual rights and will give prominence in her thoughts to features of events that emphasise wrongdoing by members of corporate elites, say, rather than members of disenfranchised groups.

We have arrived now at an understanding of perspectives as collections of psychological dispositions that generate ways of seeing. Several philosophers have offered theories that suggest perspectives can be used to explain some interesting features of our experience of literature, especially fiction.<sup>85</sup> Next I will discuss two of these

---

<sup>85</sup> Currie (2010) talks about what he calls ‘frameworks’ which are similar to my conception of ‘perspectives’ in many ways. Harold (2007) and Liao (2013) make use of a notion of ‘outlooks’ which are also similar to what I call perspectives, although they deal strictly with moral evaluation. Gendler (2000), Walton (1994) and Moran (1994) all appeal to something like perspectives in their discussions of imaginative resistance.

interesting features, the idea being to give some *prima facie* reasons for supposing that integrating perspectives into our theories of literary engagement will be illuminating and useful. These features are firstly the fact that we sometimes find it difficult to engage with morally deviant stories (the problem formerly known as ‘the puzzle of imaginative resistance’); second, that we sometimes respond differently to fictions than we would to the same events in real life, sometimes called the problem of disparate response.

#### 4.4 Perspectives and the Philosophy of Literature: Imaginative Resistance as Transmission Failure

The basic issue brought out by the so-called puzzle of imaginative resistance is the fact that we seem to find it difficult to engage with stories of particular kinds, as illustrated by the now infamous one-line fiction:

**G:** “In killing her baby Giselda did the right thing; after all, it was a girl”<sup>86</sup>

There are several issues of philosophical interest that might apply to understanding why we find it difficult to ‘go along’ with a story like (G). For instance, it might tell us something about the attitudes we have about our moral beliefs, or about the limits of our imaginative capacities, or about the kinds of things that can be made true in stories. It is this last point that I want to focus on here, what has been misleadingly termed ‘the fictionality puzzle’.<sup>87</sup>

---

<sup>86</sup> Walton and Tanner (1994)

<sup>87</sup> Originally by Weatherston (2004). It is misleading because, as we will see shortly, the effect is not limited to fictional stories.

Derek Matravers has recently proposed a promising, but incomplete, solution to the fictionality puzzle, putting forward a view that I will refer to as the *Transmission Failure View* (TF).<sup>88</sup> According to (TF), (G) and other stories that exhibit similar features (specifically, those that invite an audience to engage with a story that features the presentation of morally perverse truths), is just one way in which the following principle of narrative engagement can fail:

**M:** For some proposition P and some narrative N, P is true-in-N if P is asserted in N

Matravers claims that M can fail in (at least) the following ways:

- ‘Determinancy failure’ (when the author or narrator has made a mistake or is not asserting P, or is asserting P ironically, and so on).
- ‘Coherence failure’ (P is incoherent when taken with other propositions asserted in N such that understanding N in such a way that N is not incoherent is not worth the effort).
- ‘Transmission failure’ (when P contains a judgement of the sort that an audience will not accept on the authority of the storyteller).<sup>89</sup>

Of these three, (G) is an example of transmission failure. According to (TF) it is not in the power of the storyteller to make it true in (G) that Giselda did the right thing by killing her baby simply by asserting it. This is because the kind of judgement contained in (G), a moral judgement, is not the kind of judgement that can be made true in the narrative by virtue of (M).

Because Matravers rejects the account of fiction as prescriptions to imagine (familiar to us from Walton and Currie<sup>90</sup>) it is not open to

---

<sup>88</sup> Matravers (2014)

<sup>89</sup> Matravers (2014) p. 136

<sup>90</sup> Walton (1990), Currie (1990). Matravers follows Friend (2012) in his rejection of defining fiction in terms of prescriptions to imagine.

him to claim that storytellers cannot make us engage with (G) because of some necessary limitation on our ability to imagine morally deviant worlds. This is just as well, since content-based principles regarding the limits of our moral imagination have proved difficult to defend.<sup>91</sup> This is partly because there seem to be plenty of counter-examples to the notion that we are unable to imagine morally deviant worlds as a matter of necessity. Many works of fiction featuring morally questionable anti-heroes, such as *V for Vendetta*, *The Great Gatsby* and *Crime and Punishment*, seem to require imaginative flexibility regarding moral matters of the sort that we wouldn't expect to be capable if we were unable to imagine morally deviant worlds.

Philosophers have tended to prefer explanations of our apparent inability to imaginatively engage with (G) in terms of a related, more general inability to imagine conceptually incoherent sets of propositions, sets of propositions that the imaginer doesn't understand, or similar limitations of our imaginative abilities.<sup>92</sup> These kinds of explanations for imaginative failure when it comes to cases like (G) focus on contingent features of the examples (that there is not sufficient context given for the judgement in (G) to be comprehensible for instance), rather than claiming that imaginative engagement with (G) will fail simply in virtue of (G) containing a morally deviant proposition.

Matravers, by contrast, has an easier time of it when it comes to defending content based principles to account for the failure of storytellers to make morally deviant propositions true in a story (although his easier time here results in a problem down the line, which I hope to help him solve). Because, for Matravers, fiction isn't defined by reference to imagination, there is no pressure to avail himself of a solution to the fictionality puzzle that makes reference to

---

<sup>91</sup> Gendler (2000), cf. Stock (2005)

<sup>92</sup> Camp (2004), Matravers (2003), Stock (2005)

the imagination either. Instead, Matravers analyses transmission failures such as (G) in terms of the failure of certain kinds of testimony, in this case *moral* testimony.

As we saw above, the reason (G) fails to engage the reader (according to (TF)) is because a storyteller doesn't have it within her power to make it true in the narrative that Giselda did the right thing. The analogy with moral testimony is clear; the storyteller, in telling (G) to an audience, is presenting the story as befitting a particular moral judgement. However, this is precisely the kind of situation in which moral testimony is often said to be problematic. Consider the following example from Alison Hills' discussion of the subject:

*“Eleanor has always enjoyed eating meat, but has recently realized that it raises some moral issues. Rather than thinking further about these, however, she asks a friend who tells her that eating meat is wrong. Eleanor knows that her friend is normally trustworthy and reliable, so she believes her and accepts that eating meat is wrong.”<sup>93</sup>*

It seems open to us to question Eleanor's acceptance of her friend's judgement on the issue of vegetarianism. Not because the view is mistaken, but because moral judgements are not obviously the kinds of things that we accept based simply on another's say-so. A story similar to (G) but borrowing the moral question at issue from Hills might help to bring out the analogy:

**G\*:** In raising her child on a strict vegetarian diet Giselda did the right thing; after all, animals need not suffer for her family to be fed.

Somebody who was unsympathetic to vegetarianism might (if moral testimony *is* in fact problematic in the way Hills and others have suggested) experience transmission failure in (G\*) for just the same reasons as more or less everybody (one would hope) experiences transmission failure in (G).

---

<sup>93</sup> Hills (2009) p.94

It is interesting to note that, according to (TF), an audience that experiences Transmission Failure in engaging with (G\*) will do regardless of whether that story is presented as fiction or non-fiction. M is a way of making things true in any kind of story, and it is a way of making things true that depends on testimony. When being told a story by a storyteller, an audience takes it on the authority of the teller that any assertions made are true according to the story, excepting only those cases where the testimony fails; which ought to include not only those exceptions laid out in (1) -(3), but also any other instance of assertion which, in general, testimony is not sufficient to warrant a belief in the hearer. M is not sufficient to make the respective judgements acceptable to an audience of either (G) or (G\*) because M is a way of making things true in narratives that depends on testimony, and (G) and (G\*) are prime candidates for testimonial failure.

#### *4.4.1 Transmission Failure and Mildly Deviant Stories*

The Transmission Failure view is certainly appealing, particularly as a contribution to the larger project of explaining interesting features of our engagement with stories (including our inability to engage with stories like G) without relying on increasingly unpopular views about fictional stories being characterised by prescriptions to imagine. Given that transmission failure appears to occur across the gamut of narrative forms, fictional and non-fictional alike, it is far more satisfactory for an explanation to not rely on features that define one or the other (or neither, as the case may be).

Matravers' view does, however, come up against a potentially serious problem, which he anticipates in this way:

*“[This] solution works only if the reader of a fiction engages with a representation of non-actual events using their actual moral beliefs. However, this discounts a possibility. Why can they not assume a non-actual persona (one who does not find the claims made in the narrative*

*morally problematic) and get on with enjoying whatever fiction it is that they are reading?*<sup>94</sup>

Matravers has explained transmission failure by analysing it as an instance of testimonial failure. However, there are some important dissimilarities between instances of straight-forward testimony and testimony given as a part of a storytelling, specifically in the norms that seem to govern how we should receive that testimony. When it comes to a great many stories, particularly fictions, much of the value in engaging with them is due to virtues of the work other than that they might accord with our existing beliefs. We might be more interested in being entertained with suspenseful drama or dazzled with fine prose than in having the story function as a straight-forward source of testimony, moral or otherwise. To that end, just as we might entertain the possibility of time travel in order to get at the literary value in *Slaughterhouse Five*, why would we not entertain the appropriateness of deviant moral judgements on the same grounds?

Obviously we must beware of making too much of the analogy between entertaining the possibility of time travel and entertaining the rightness of infanticide. For one thing, a story that failed to make the possibility of time travel true in the narrative would be an instance of either determinacy or coherence failure, and not transmission failure; the possibility of time travel is something that can be made true in the narrative by (M) provided it maintains both coherence and determinacy, the rightness of infanticide by contrast fails simply by virtue of its subject matter.<sup>95</sup> The point remains, however, that we tend to be flexible about shelving our real-world attitudes about things like time travel for the sake of engaging with some worthwhile story,

---

<sup>94</sup> Matravers (2014) p. 142-3

<sup>95</sup> In (G) at least. Matravers, as I will shortly argue, seems committed to the view that the rightness of infanticide can't be made true in any narrative, and this I have reason to doubt.

so Matravers owes us an explanation for why that doesn't routinely happen with stories like (G).

I believe the Transmission Failure view can overcome this challenge, however the first step in showing how is to recognise that the problem goes deeper than Matravers anticipates. (TF) as laid out by Matravers seems to preclude *any* story from making it true in the narrative by M that any moral judgement is appropriate. Transmission failure in (G) is not due to *how* deviant the moral testimony is in (G), only to the fact that there is moral testimony at all. This is a problem because there are many examples of less severely morally deviant stories in which no transmission failure appears to occur. Earlier I mentioned that it has been difficult to show that imagining moral deviant worlds *tout court* is impossible, because we seem to be able to imagine without difficulty fictional worlds wherein morally dubious actions are praiseworthy and heroic. Although Matravers is not committed to the notion that engaging with morally deviant fiction involves imagining morally deviant worlds, such stories still provide a problem for (TF) if the moral praiseworthiness of what we would otherwise consider immoral acts is attempted to be made true in the story by M. (TF) in its current form seems committed to the implausibly strong claim that M cannot be used to present a morally bad thing as otherwise in a story without succumbing to transmission failure.

It is, of course, open to defenders of (TF) as it stands to claim that any instance of an audience successfully engaging with a mildly morally deviant story is an instance of a storyteller making use of something besides M. However, it would be difficult to spell out precisely what method besides M would be useful to a storyteller who wished to engage an audience in a morally deviant story. In the most general terms, to make something true in a story a storyteller only has the resources of her own assertions about what is the case in the story, any relevant conventions of storytelling and any facts she might reliably expect an audience to *import* to the story from their stock of beliefs



about the world.<sup>96</sup> The first of these is captured by M, and we already know that (TF) predicts the failure of any attempt to use M to present a morally bad thing as morally praiseworthy. Nor can a storyteller make a morally perverse story by the use of import, since she could only expect an audience to reliably import morally *unperverse* truths into any story she told. The most promising option is, I think, that there might be certain storytelling conventions that permit an audience to engage with a morally perverse story. Shen-Yi Liao and James Harold have both proposed views along similar lines; specifically, that the *genre* a story is presented as belonging to can alter the way we respond to the moral assertions made as part of that story.<sup>97</sup> Taking their view as a starting point it might be possible to sketch a theory according to which conventions of genre or form signal to an audience that morally perverse truths are to be entertained in a given case. However, I think it is doubtful that any such project could succeed. I have particular worries about the view that Liao and Harold endorse which I will turn to at a later stage. However, for our current purposes it is sufficient to note that the sheer breadth of narrative genres and forms that feature mildly deviant stories makes it very unlikely that genre or form are what make audiences able to engage with them all.

#### *4.4.2 Perspectives and Moral Testimony*

The challenge for (TF) is, therefore, more serious than Matravers first claimed. Not only does he have to explain why we cannot entertain alternative moral beliefs for the sake of engaging with a good story, he also has to explain how transmission failure does not occur in cases where (TF) predicts that it should. That a solution to the first part of

---

<sup>96</sup> Walton (1990), Gendler (2000)

<sup>97</sup> See Liao (2013), Harold (2007). It should be noted that both authors are principally concerned with how we can be morally educated by morally perverse stories, rather than how we can engage with morally perverse stories at all. This is a topic I will devote more time to later in the thesis, the discussion of which will include a detailed evaluation (and rejection) of Liao's view.

the challenge will also provide an answer to the second is promising, since both are to do with explaining the scope of transmission failure; why we don't prevent it the way we seem to prevent other kinds of failures of (M), and why it seems to not occur in some common kinds of cases.

Matravers offers himself the start of a solution to the first part of the challenge by way of some thoughts on moral testimony by Alison Hills and Rob Hopkins. Hopkins has argued that pessimism about moral testimony comes in two kindred forms; the view that moral testimony fails because it does not make moral knowledge *available* or, that moral testimony fails because it cannot make any available moral knowledge *useable* by the recipient.<sup>98</sup> Hopkins argues that the second route to pessimism is more promising, in part because for that knowledge to be warranted for use in one's operations as a moral agent (as opposed to merely knowledge about moral matters), the belief in question must meet the following requirement:

**The Requirement:** having the right to a moral belief requires one to grasp the *moral* grounds for it.<sup>99</sup>

Matravers does not say as much, but I take it that this thought from Hopkins is supposed to apply to the case at hand in something like the following way: Moral testimony operating according to (M), that is, testimony to the effect that some moral fact obtains in a narrative, operates according to the same norms as moral testimony simpliciter. In (G) we have access to no *moral* grounds on which we could legitimately deploy the belief that Giselda did the right thing (although we may have reasons of getting at literary or other kinds of value for entertaining such thoughts). Because we have no *moral* grounds for adopting or entertaining a belief that Giselda did the right

---

<sup>98</sup> Hopkins (2007). Hills (2009) offers the similar thought that competent use of moral knowledge requires *understanding* the moral grounds for that knowledge.

<sup>99</sup> Hopkins (2007) (emphasis in original)

thing, we cannot use the belief that Giselda did the right thing as a part of our moral engagement with the story.

This seems fine as far as it goes. While it is not obviously true that the norms that govern real-world moral testimony should apply to moral testimony given as a part of storytelling I can't think of any particular reason why they would not. This solution only struggles when we start to think about how we should generalise from its explanation of (G). I said above that it appears that Matravers is committed to the claim that any and all moral testimony given as a part of a storytelling is bound to succumb to transmission failure. I am sure Matravers would want to deny that this causes his view problems in accounting for the kind of mildly deviant stories I have mentioned, but he does not directly address the issue. The examples he deals with are all at least fairly extreme; besides (G), he illuminates the Hopkins requirement by affirming that "we cannot understand how something could be a moral reason for the denigration of non-white ethnic groups or the mass extermination of Jews."<sup>100</sup> By far the least morally deviant story he tests his theory on is Peter Fleming's anecdote detailing how he shot an alligator simply to ease his frustrations, presenting that as a perfectly appropriate thing to do. The fact that Matravers mentions, indeed focuses, on the less morally outrageous Fleming example indicates to me that he does not think the *severity* of the moral deviance in each case has much to do with why that case succumbs to transmission failure. Besides, Matravers hasn't furnished himself with the resources necessary to discriminate these cases on such grounds. However, it is still telling that even the least morally deviant example he discusses expresses a moral belief that would be shared by very few modern readers (certainly very few readers of Matravers). Focusing on these examples is unhelpful because it discourages the raising of a very important question that needs

---

<sup>100</sup> Matravers (2014)

answering in order for Matravers' view to evade the challenge; what, if anything, could count as moral grounds for accepting moral testimony given as a part of a storytelling?

This brings us on to the second challenge. If Matravers lacks the resources to discriminate cases of moral testimony in storytelling on the basis of the severity of the deviance of that testimony, how can he account for cases where we appear happy to adopt minor changes to our moral outlooks in order to engage with a seemingly worthwhile story? If we say that no story can provide moral grounds for the acceptance of any moral testimony given as a part of the telling of that story, then these cases simply cannot be accounted for. There may well be cases where people already possess the moral grounds required to engage with what, it turns out, are morally deviant stories, but these are not the cases we are concerned with. Furthermore, it doesn't seem plausible to insist that wherever there does not appear to be transmission failure when mildly deviant stories engage their audience that is just because the audience already possess the moral grounds for those deviant beliefs; after all, people will generally not accept that they have moral reasons for believing what they think are morally deviant assertions.

A satisfactory solution to both parts of the challenge can be found by incorporating perspectives into the Transmission Failure view. Recall that Matravers' challenge to himself was to explain why we don't simply 'adopt' beliefs in line with morally deviant stories in order to engage with them. My additional challenge was that (TF) needs to explain how transmission failure is sometimes overcome or avoided in order to explain the fact that we seem to be able to engage with stories that are mildly morally deviant. The answers to both hang on whether and how morally deviant stories can give us Hopkinsian moral grounds for engaging with them. In brief, my proposed solution is that adopting a perspective can provide moral reasons for adopting moral beliefs on the basis of testimony. The first part of the challenge

is evaded because, it turns out, we *do* habitually adopt morally deviant beliefs in order to engage with morally deviant stories (or, rather, we adopt a perspective according to which those beliefs are appropriate). The second part of the challenge is partly answered in the same way, but my solution can also explain why transmission failure will occur for some morally deviant stories and not others; transmission failure will occur when the story fails to provide the reader with a perspective that can ground that story's moral content *or* when the invited perspective of the story is so different from the audience's own that the expected value of engaging with the story is not worth the trouble of adopting that perspective.

It is fairly easy to show that perspectives, as I have described them, can provide the kinds of moral grounds the Hopkins requirement demands. In defence of the idea that someone may legitimately say that something 'feels [morally] wrong' even though they can offer no further moral reasons for thinking so, Hopkins gives the following examples of what may count as moral grounds:

*"In allowing the legitimacy of moral beliefs based on nothing more than 'feeling' we are not, in fact, allowing beliefs in the absence of a grasp of the moral reasons for them. Rather, what the case shows is how broad a notion of a moral reason The Requirement should deploy. In allowing belief based on 'feeling' we leave room for such things as seeing the proposed action in a certain light, or finding that it conflicts with our settled moral dispositions to act and (in a stricter sense) to feel in certain ways. Someone reacting in such ways as these to the claim that some action is morally acceptable does indeed grasp what is wrong with it, for all that she is unable to articulate that wrongness."<sup>101</sup>*

Perspectives are certainly able to contribute towards 'seeing actions in a certain light'. Recall that perspectives are standing dispositions to generate ways of seeing (to pick out certain types of features of scenes

---

<sup>101</sup> Hopkins (2007)

as more prominent or central, etc.); perspectives generate particular ways of seeing events and objects. Given that these ways of seeing involve favouring certain kinds of causal explanations and, crucially, evaluations, it is easy to see how they can provide the moral grounds that Hopkins describes. A generous perspective, to recall my earlier example, will for instance provide moral grounds for thinking that buying the first round is a morally appropriate action for the reasons that it will dispose one to characterise an evening at the pub as an opportunity to increase the wellbeing of your friends, to show that you value their company, to rid anybody else of the burden of having to go to the bar, and so on. These reasons are all moral reasons (and good moral reasons at that), and the fact that the generous perspective can furnish somebody with the disposition to characterise scenes in ways that make those moral reasons prominent gives us a simple way to show how adopting alien perspectives can provide moral grounds for adopting alien moral beliefs.

Although I have shown that adopting an alien perspective is *sufficient* for granting moral grounds for adopting alien moral beliefs, it might be objected here that the former is not *necessary* to the latter.

Perhaps there are other ways that stories enable their audiences to meet The Requirement, thereby evading Transmission Failure. The most obvious alternative might be termed the *information view*, according to which all the necessary work to provide moral grounds for the reader is achieved by storyteller making other simple assertions in the story; that the storyteller makes the kinds of things true in the story that can serve as moral grounds for accepting perverse moral truths that are also asserted in the story.<sup>102</sup> This objection can be met simply by observing that information on its own often underdetermines the audience's response to a scene or story; indeed,

---

<sup>102</sup> Keiren (2003) defends a version of this view.

even altering the order such information is revealed can drastically alter an audience's response to the events depicted.

Daphne du Maurier's *Rebecca* is a fine example of a story in which a principal character holds a dark secret. What is important for our purposes is to note that the point at which Maxim's secret (that he murdered his wife in a jealous rage) is revealed is just as significant a factor in determining our view of him by the end of the story as what the secret itself consists in. If we had been told of Maxim's murderous deeds in a prologue, having not yet established his clearly tortured conscience, an audience would be far less kindly disposed towards him than we end up being when we finally discover the truth. The place in the story at which we are told Maxim's secret strongly determines the way we respond to learning that secret. The information given in a narrative is certainly among those features of a story that determines what perspective we are being invited to adopt on the events that transpire. However, since the *order* that information is revealed doesn't itself constitute information in the sense we're interested in, and the order the information is presented clearly determines the audience's response in at least some cases, the information view can't sufficiently explain the audience's response.

A defender of the information view might at this point object that the order in which events are narrated or information revealed should, in fact, count as information presented to the audience. The idea being, perhaps, that the order in which information is revealed to the audience is itself information that an audience gains as they progress through the story (call it 'ordering-information'). While it is undoubtedly true that an audience does gain ordering-information as they read a novel, and learn at what points certain crucial pieces of information are received, the ordering-information does no work in explaining the audience response to *Rebecca* as the story unfolds. This is so because ordering-information is information concerning *external* features of a story. That is, features of a story we speak of when we are

not engaged with it *as* a story, but rather considering it as an artefact. It is an *external* feature of *Rebecca* that it was written by Daphne du Maurier, is around 150,000 words long, that it is a gothic mystery, and so on. *Internal* features of *Rebecca* include things like the nervousness of the protagonist, the fact that Manderley has latticed windows and a gravel drive, and that Manderley's housekeeper is named Mrs. Danvers. Information narrated to an audience relates only to *internal* features of a story. Ordering-information, on the other hand, is an external feature of a story because what order an author chooses to reveal information to us is information about the way the story is constructed; not information revealed to us about narrated events by their being narrated, but information revealed to us about the narration itself. A response to an external feature of *Rebecca*, such as ordering-information, is a response to *Rebecca* as an artefact and might include such thoughts as "Oh, what an interesting literary character he turned out to be!", but never "Oh, that poor man!". Our surprisingly soft feeling towards Maxim require that we think of him as a man, and not as a fictional character, which in turn requires us to take an *internal* perspective on him when we respond. In short, ordering-information cannot determine our responses to events in stories as we engage with them, because ordering-information is not narrated information, but rather information about the narrative *qua* artefact.

I will gloss over the question of how exactly a storyteller invites an audience to adopt a particular perspective; the methods seem to vary from case to case. Some of the more obvious ways might include making use of metaphor in such a way that it leads to consistent characterisations of some kinds of characters or events in the story in ways consistent with a particular perspective.<sup>103</sup> If a story kept referring to the actions of the Sheriff protagonist in leonine terms and

---

<sup>103</sup> Camp (2006)



the actions of the Outlaw with, say, serpentine imagery, it might encourage the audience to adopt a perspective according to which lawmen are generally to be trusted, praised and so-on, and those they suspect of wrongdoing as being more than likely good-for-nothing varmints.

What is more interesting to us now are the mechanisms by which a story can *fail* to get us to take up a perspective suited to engaging with that story. I mentioned two: when the story simply does not provide us with an opportunity to take up a perspective appropriate to engaging with the story, or when the perspective offered to us is so far removed from our own that we don't feel engaging with the story is worth the effort of adopting the perspective. Stories can fail in the first way by at least two different means; they can offer us the *wrong* perspective or they can offer us no alternative perspective at all. In both cases transmission failure occurs for the same reason; we are engaging with the stories with our existing perspectives and they do not allow us to characterise the narrative in a way that will furnish us with grounds for engaging with its moral content. Stories can fail to engage us the second way by inviting a perspective so far removed from our own that the sheer psychological effort needed to think ourselves into that perspective, and make it coherent enough to be able to provide Hopkinsian moral grounds, does not strike us as worthwhile. It is perhaps tricky to describe precisely what makes one perspective 'far removed' from another, but I think we can characterise the distance in terms of the specific mental states (especially beliefs, evaluations and feelings) in which perspectives partly manifest themselves. If I find that an invited perspective on a story would involve me, for instance, responding positively to gender-selective infanticide, I am likely to disregard the project as a waste of psychological resources; the kinds of perspective shifting hurdles I

would have to jump through to arrive at such a response are simply not worth the expected value of engaging with a story like (G).<sup>104</sup>

We have arrived at a solution to the challenges facing the Transmission Failure view. We can now, as we could not before, explain why transmission failure occurs in cases like (G) but not in what I have called ‘mildly deviant stories’ such as *Crime and Punishment* and *V for Vendetta*. (G) will, I think, inevitably fail for both of the reasons I mentioned; it does not give the reader an opportunity to adopt a perspective that would furnish her with grounds to share the story’s moral content, and any perspective that *could* provide such grounds is so far removed from that of any likely reader that hardly anyone would even try to adopt it.

#### 4.5 Perspectives and the Philosophy of Literature: Disparate Response

Transmission Failure is an example of the limits of authorial authority when it comes to making things true in stories, but it is important to recognise that storytellers can exercise a significant amount of power when it comes to determining how we respond to their tales. Like the first objection to (TF), the problem of disparate response is motivated by the datum that we sometimes respond to events described in stories in ways other than we would if we encountered those events in different circumstances. In this section I will examine several ways of conceptualising the problem of Disparate Response and conclude that any satisfactory explanation of the phenomenon will have to make use of perspectives.

---

<sup>104</sup> Note that I can remain neutral on whether it is possible or not to *imagine* that (G) is true. It may be possible with sufficient context for the imagining built around (G) (as Stock (2005) suggests), but even if it is the distance of the invited perspective from that of any likely reader will almost invariably result in transmission failure.

There are countless examples of stories in the literary canon that appear to successfully induce audiences to experience emotional, aesthetic, humorous and moral responses to their represented events in ways that seem surprising, given how we would expect audiences to ordinarily react to those kinds of events. Indeed, for some of the most (in)famous examples such as *Lolita*, inducing the audience to depart from their ordinary emotional and moral standards in engaging with the story seems to be one of the central goals of the author. Entire genres of story, such as dark comedy, violent slapstick and almost anything featuring a morally questionable protagonist, depend for their success on getting the audience to laugh at what they would ordinarily find horrifying, hope for the success of plans they would ordinarily wish to be thwarted and cheer for characters whom they would ordinarily despise. The problem of Disparate Response is why it should be the case that we would respond to representations of events in stories in ways that apparently depart from how we ordinarily would respond to those events.<sup>105</sup>

#### 4.5.1 *The Problems with the Problem of Disparate Response*

The Problem of Disparate Response is simple to grasp at a surface level, and I would hazard that the effect is familiar to anyone who habitually reflects on their responses to stories. However, when we try to analyse Disparate Response it becomes clear that it isn't obvious what we should think of as the problematic feature of this phenomenon, or how we should understand the finer details of the phenomenon itself.

In particular, it is difficult to nail down precisely which set of 'ordinary' responses ought to serve as the comparator to the 'disparate' responses brought about by engaging with stories. For instance, we might think that my support for Charles Ryder's conversion to

---

<sup>105</sup> Currie (1995, 2010), Goldie (2003), Kieran (2003) all formulate the problem of Disparate Response in approximately this way.

Catholicism in *Brideshead Revisited* constitutes a fine example of Disparate Response, but by virtue of disparity with what expected response of mine does this support for Ryder's conversion qualify as 'disparate'? Going forward I will refer to responses that have supposedly been brought about as a result of the phenomenon of Disparate Response as 'disparate responses', and those responses from which disparate responses are different as 'ordinary responses'. Is my support of Ryder's conversion a disparate response because:

- I would respond differently to a non-fictional story relating the same events?
- I would respond differently if I witnessed an exact real-world version of those events first-hand?
- I would respond differently if I witnessed an approximately similar version of those events first hand?
- If I myself were to write about these events I would invite a different response from my readers?
- My response, on reflection, surprises me because it seems to run counter to my more general beliefs about the value of religious conversion?

To begin to get a grip on which set of ordinary responses should be the comparison we should note that, as in the case of Transmission Failure, the phenomenon of Disparate Response is not confined to *fictional* stories; two different, non-fictional accounts of the same real-world events can often provoke responses in their readers that are different both from one another and from our other candidate sets of ordinary responses.<sup>106</sup> To take an example, Volker Skierka's biography of Fidel Castro is broadly sympathetic, whereas Leicester Coltman's is positively damning. While both are reporting on the same subject matter, the life and work of Fidel Castro, a reader's response to the events described is likely to differ wildly depending on *how* they are reported. A reader of Skierka's book is likely to respond one way, a reader of Coltman's book quite another. More to the point for our

---

<sup>106</sup> Alward (2006)

current discussion, neither of these seems necessarily connected to any candidate set of ordinary responses.

Disparate Response, therefore, doesn't appear to be associated with fictional stories so much as stories *per se*. As such, it would be difficult to suggest that my response to *Brideshead Revisited* counts as disparate because I would respond to a non-fictional report of the same events in a different way; firstly because it is not at all obvious that I *would* respond in a different way (if the report were the same in all possible respects while being non-fictional), and secondly because Disparate Response is a feature of engaging with non-fictional stories too, and we cannot therefore use our responses to non-fictional stories as a comparison class to determine what responses count as disparate for other non-fictional stories.

Perhaps, then, what marks our responses to certain stories out as disparate is not that they are different to how we would expect to respond to other *reports* of similar events, whether fictional or not, but how we would respond if we came across similar events in real life. Thinking again about my response to Charles Ryder's conversion to Catholicism, perhaps my support for his conversion counts as disparate not from my expected response to a non-fictional *report* of such a conversion, but from how I would respond to real-world experience of those events.

However, using real-world experience as the benchmark of our ordinary responses is also problematic for the purposes of identifying disparate responses. When considering a case like *Brideshead Revisited* there is an initial complication in that the novel is narrated in the first person. What would count as a real-world experience of the events of *Brideshead*, therefore, might just be a real-world experience of seeing, hearing or reading somebody like Ryder recounting certain events of his (non-fictional) life. This would amount to nothing more than an experience of being told a non-fictional story, however, and as we have already seen, non-fictional

stories are just as capable of eliciting Disparate Response as fictional ones.

To obtain an adequate comparison class in order to usefully speak of Disparate Responses, we need to divorce the events of a story from any narrating of those events. Because the phenomenon we are attempting to describe occurs in response to all kinds of story, to determine what our ordinary response to any narrated events might be, we must consider them in a non-narrated form. This is, unsurprisingly, a rather difficult enterprise. Part of what it is to narrate a series of events is to selectively include and disregard various elements of those events, to mark some out as especially central or worthy of attention, and to present those events as following on from one another in a coherent fashion.<sup>107</sup> This means that, from the outset, it is impossible to gather a narrative-neutral collection of facts from a story to evaluate and respond to in your ordinary way; the collection of facts available in the narrative have already been selected to fit their existing narrative structure. This is of course especially true of events narrated in fictional stories; because fictional events often have no existence beyond their presentation in a story, there is no narrative-neutral fact to be evaluated or responded to.<sup>108</sup>

Even if we ignore the complications of first-person narratives and deciding what set of facts we need to hypothesise responding to, locating a sensible set of ordinary responses is far from a simple matter. It is hard to make sense of the idea that we can hypothesise experiencing the events of *Brideshead Revisited* first-hand while still remaining ourselves in all those ways necessary to ascertain what our ordinary response to those events would be. My ordinary responses

---

<sup>107</sup> Goldie (2012)

<sup>108</sup> There are doubtless cases where some audiences *do* have access to the facts of a story in a narrative-neutral way. For any story about a real life event within living memory, there will be those whose ordinary responses to the events of the story will have been determined by their actual responses to those events. This kind of case is hardly typical, however, and as such wouldn't be a promising starting point to establish what 'ordinary' responses might be.

are doubtless partly determined by things that have happened to me, things I have been told, taught and learned to expect. When I am hypothesising what my ordinary responses to the events of *Brideshead* would be I would presumably have to maintain those elements of my dispositions to respond which are informed by these experiences; I want to be able to say that my ordinary response to, say, Julia's distress at being unable to marry Rex in a Catholic ceremony might be informed by my modern attitudes about the insignificance of such things. However, if I want to maintain these kinds of responses in my ordinary repertoire, then wouldn't my ordinary response to the events of *Brideshead* also include other responses informed by my real-world experiences, such as being surprised to find myself witnessing events that occurred some fifty years before my birth?

It is tempting here to just stipulate that only responses to the details of the events and not the occurrence of the events *per se* should count as 'ordinary'. This would rule out my being surprised that I was witnessing events taking place before I was born but include any modern influences on my responses to the things taking place. However, I am not convinced that there is a principled way to make this distinction, or if there is that it would even be a useful one to have. In the first place it is often the occurrence of an event itself that is the subject of both ordinary and disparate responses (such as when we learn about the event of Kurt's suicide, to take another example from *Brideshead*). We would want to include such responses in our 'ordinary' class, but the proposal on the table rules that out. On the point of principle, it is not likely that we could have any response to any event *sans* detail, if details of events like Kurt's arrest and suicide include things like the fact of Kurt's death. Even if we wanted to say that we can be surprised at the occurrence of events *per se*, a very difficult question is then raised as to how we can be surprised at the *event* of Kurt's death without thereby being surprised at the *fact* of Kurt's death.

#### 4.5.2 *The Expectation View*

The lesson to take from all this, I think, is that there is unlikely to be any satisfactory way to determine our ordinary responses to stories along these lines. Hypothesising coming into contact with the events of a story won't reliably provide what we have in mind when we talk of 'ordinary' responses, either because we can't extricate the events from the narrative in which they feature, or because it will often be impossible to sensibly place ourselves within those events and respond in what we would think of as an ordinary fashion.

It doesn't seem possible to adequately delimit a class of responses along these lines that can serve as a comparison to determine what responses count as 'disparate'. However, this isn't to say that Disparate Response is not a genuine phenomenon, nor that it is not especially prominent when it comes to our experience of engaging with stories. Rather, I think we just need to make do with a more straightforward description of the phenomenon of disparate response: a response is not 'disparate' from what our *actual* responses to a narrated event would be if it were not narrated, but rather from what our *considered expectations* are of how we would respond to a narration of those events.

Specifically, a response is disparate if it runs counter to my expectations of how I would respond to that story in virtue of my considered beliefs about my dispositions to respond to those kinds of events. My encouraging response to Ryder's conversion is disparate not because I would respond differently if I knew Ryder in real life, but because I would not expect, if I reflected on what I know about my own views on the value of religious conversion, that I would respond encouragingly to such a conversion in a story.

This view of the problem of disparate response has several virtues when compared to what we may call the *actualist* position (the view that a response to a story is disparate by virtue of a comparison with how we would respond to an actual version of those events). The



most obvious advantages of the view I propose (let us call it the *expectation* view) are that it is not subject to the problems outlined above that face the actualist view; there is no need to delimit a class of responses that involves regarding events in a story in a non-narrated form, nor to answer difficult questions about what parts of myself are responding to what parts of the story.

One interesting advantage of my proposal is that it gives us a way to refrain from having to decide whether a response counts as disparate when it deals with events that are completely removed from the experience of the audience of a story. If someone had never given much thought as to whether they would be supportive or condemning in the face of a situation such as the illicit affair of Charles Ryder and Julia Flyte, then not much is gained by claiming that their response is disparate. The actualist, however, is committed to claiming that such a response may well count as disparate, if it happened to differ from how the reader would respond to an actual event of the same kind. The expectation view is at liberty to avoid terming a response to an unfamiliar situation as either disparate or ordinary; it is simply new. Furthermore, the expectation view also permits us to say, quite sensibly, that our response to an event in a story might determine our expectations of how we might respond to a similar event in real life. The actualist view might struggle to follow the same thought, since our responses to stories are apparently so often ‘disparate’, and therefore unreliable as a way of setting our considered expectations of how we might respond in the future.

The expectation view allows us a simple and plausible way to describe the relationship between the *framing effects* famously demonstrated by the psychological experiments of Tversky and Kahneman and the phenomenon of disparate response.<sup>109</sup> Participants in these experiments were asked whether they would give their approval to one

---

<sup>109</sup> Tversky and Kahneman (1981)

of two policies which, in both cases, would save 200 lives at the risk of the loss of 400 lives. Participants were more likely to give their assent to a policy which was described in terms of lives saved and less likely to give their assent to the same policy described in terms of lives at risk. This experiment, and others like it, show that the way a narrative is *framed* (in this case, whether it a scenario is described in terms of lives saved or lives risked) has the effect of manipulating an audience's response to it in predictable ways.

Framing effects are certainly related to both the problem of disparate response and perspectives in general. A shift in perspective from one scenario to the other, by emphasising lives saved in one case and lives lost in the other, gives us a good way to explain why the framing effect carries the force it does in determining participants' responses. As for disparate response, it has been suggested that the phenomenon of disparate response constitutes a special subclass of framing effects.<sup>110</sup> This characterisation works far better if we take the expectation view of disparate response rather than the actualist view. On the expectation view we can simply say that disparate responses are those brought about by framing effects, and which run counter to a subject's expectations. It is not open to the actualist, however, to make an analogous claim (that disparate responses are those brought about by framing effects, and which are different from how the subject would ordinarily respond). Firstly, because framing effects have only been observed as a feature of stories (broadly construed); participants in experiments are experiencing the framing effects of two different narratives concerning the same information. Since, as I have previously shown, the actualist cannot use a response to other stories as a benchmark for what counts as an ordinary response, the actualist cannot claim that the response to any of the stories exhibiting framing effects counts as ordinary. From this it follows that the actualist also

---

<sup>110</sup> Currie (2010)

cannot claim that any of these responses count as disparate, since there is no adequate comparison class to determine what counts as an ordinary response to being told any kind of story, including the minimal narratives exemplified by the vignettes frequently used in psychological experiments.

Given that my goal here is to show that perspectives are essential to describing the phenomenon of disparate response, I might appear to have made things difficult for myself by rejecting the actualist account of the phenomenon. It is relatively easy to show that perspectives are essential to a good description of disparate response in actualist terms. As the experiments concerning framing effects show, it is the perspective an audience is invited to adopt on a scenario that determines their response to it every bit as much as the scenario itself. This is a very important point, since by itself it fends off what is possibly the strongest objection to both the actualist and the expectationist views of disparate response, a return to what I have called the *information* view. It has been suggested that perspectives have no explanatory value in explaining disparate response, and that the apparently different responses that we have to stories can be explained simply by the information we are given in those stories, and how that would differ from information we might have access to that would lead us to our 'ordinary' responses.<sup>111</sup> Cases like the experiments of Tversky and Kahneman show that this cannot be the whole story. Assuming that framing effects are closely related to the phenomenon of disparate response (which the information view is not equipped to dispute), such cases prompt different responses with the same information, merely presented in a different way. The presentation, therefore, must be what is doing the work to bring about the effects in which we are interested, and this presentation is best described in terms of perspectives. It is also worth mentioning that, as with the

---

<sup>111</sup> Kieran (2003)

perspective centred approach to Transmission Failure, a perspective centred approach to disparate response gives us a way to predict that occasionally the invited (disparate) response to a story will not materialise in the audience; because adopting the required perspective to produce that response is too much effort for the expected value of engaging with the story.

An actualist can explain disparate response in terms of perspectives very simply; we have different responses to stories than we do in real life because stories encourage us to adopt different perspectives than those with which we usually operate. The expectation view must concede a certain advantage of neatness to the actualist on this point: the actualist can use perspectives to both explain *when* a response counts as disparate (when it is brought about during the adopting of a perspective that is not our ordinary one) and to explain *how* we can have different responses in different stories. The expectation view concurs with the actualist that perspectives explain how we can have responses that count as disparate, but appeals to the confounding of our considered expectations to determine *when* a response counts as disparate. The expectation view, in full, appeals to perspectives in something like the following way: we are surprised by some of our responses to stories because they run counter to how we would expect ourselves to respond to that story, given what we think our usual perspective is on those kinds of events. Returning to the case of Ryder's conversion, I might think that I would expect my response to be disapproving, because I think of myself as being the sort of person who tends to disvalue religious conversion, especially as a solution to the kind of problems Ryder is having. However, when I reflect on my actual response to the closing paragraphs of *Brideshead Revisited*, I find that my perspective on these events is different to how I would expect it to be, which in turn has led to a soft-hearted response, bordering on supportive.

And here is where I think the most significant advantage of the expectation view lies over the actualist view; in emphasising our reaction to our own responses, the expectation view opens the door for a powerful theory of the role of disparate response in accessing cognitive value in stories. Disparate responses, on the expectation view, give us an opportunity to reflectively realise our capacity to adopt different, coherent perspectives on various situations; the surprise I feel at responding positively to Ryder's conversion is cognitively valuable not just because it makes me realise that I am capable of seeing the situation in a way that merits that positive response, but because that perspective, and the response that comes with it, are experienced as consistent, coherent and natural.

## 5. EMPATHY AND THE NARRATIVE PERSPECTIVE

### 5.1 Locating the Narrative Perspective

I have shown that two thorny problems in the philosophy of literature can be both clarified and solved by appealing to perspective shifting. I take it that the pervasiveness of these phenomena (particularly Disparate Response) is sufficient motivation to embrace the idea that perspective shifting is, therefore, a central part of narrative engagement. This takes us half way to showing that narrative engagement involves being-with, and is thereby an empathic activity. What remains is only to argue that narrative engagement is undertaken with the aim of making a demonstrative ascription to a target.

The first and most important question to address is the nature of the target of the proposed empathising. Since we are looking for a target that is common across all kinds of narrative, a good place to start would be to identify the holder of the perspective that we appealed to in solving the problems of transmission failure and disparate response; what I have referred to as the *narrative perspective*.

Recall that perspectives are collections of psychological dispositions that generate ways of seeing. If we are looking for possible targets of empathy that plausibly hold the narrative perspective, we ought to start with figures who exist across all kinds of narrative and who have the capacity to possess perspectives as we have understood them. Furthermore, the narrative perspective, when held by this figure, must be able to perform the functions I have assigned to it in my discussion of Transmission Failure and Disparate Response; i.e. the narrative perspective must be able to provide moral grounds for moral beliefs (a la *The Requirement*), and be able to provoke disparate responses in audiences. There are several figures of various kinds with which we

are familiar from both philosophical and critical thought about literature who fit the bill, including authors and audiences (both actual and hypothetical), characters featuring in stories, and narrators of various kinds.<sup>112</sup>

Let us first divide the candidate figures along lines that will be useful in order to compare them for their suitability for our purpose. We can first distinguish some candidates from others by virtue of whether they are real (flesh and blood) people, figures internal to the narrative or hypothetical figures constructed from or implied by the text. Potential holders of the narrative perspective who are real, flesh and blood people include:

- Flesh and blood authors (e.g. Daphne du Maurier)
- Flesh and blood audiences (e.g. Me reading *Rebecca*)
- Flesh and blood narrators (e.g. Anna Massey reading the audiobook edition of *Rebecca*)

Flesh and blood figures of this kind, being real people, all have the capacity to hold perspectives that can provide moral grounds for moral beliefs and provoke disparate responses (if their perspective were to be adopted by another).

Other potential candidates for the holder of the narrative perspective are internal to the narrative itself, including:

- Narrators of stories related in the first person (e.g. Doctor Watson)
- ‘Effaced’ narrators (narrators internal to a story who are not made explicit in the narrative, as in most stories featuring omniscient third-person narration).
- Characters featuring in the story.

First person narrators are also uncontroversially able to hold the right kind of perspectives since, along with characters in stories, they are a

---

<sup>112</sup> Some individual figures will of course occupy more than one role. Any fiction told in the first person, for instance, will feature a fictional character who also serves as the story’s narrator.

kind of person that we can speak of as exhibiting the outlooks or ways of seeing which are the manifestations of perspectives. The same is true of effaced narrators, though as I will discuss there is some controversy over whether they exist in all cases.

The final group of candidates include hypothetical or implied figures, such as:

- An implied/hypothetical/ideal author
- An implied/hypothetical/ideal narrator
- An implied/hypothetical/ideal audience

Whether figures such as implied authors, narrators and audiences exist is, unsurprisingly, more controversial than whether the above flesh and blood or narratively internal figures exist. However, I will be arguing that it is among these that the holder of the narrative perspective will be found, specifically by an implied narrator, who shares some features with the traditional implied author figure so often invoked in literary criticism.<sup>113</sup>

Given the relatively controversial status of implied or hypothetical narrative figures, I will begin by showing how the other, less controversial candidates are unsuited to my purposes, before then going on to explain how an implied narrator works well as a target for empathic engagement in stories, without necessarily committing me to the problematic features of such implied figures that have caused so many to doubt their existence.

#### *5.1.1 Flesh and Blood Authors, Audiences and Narrators*

The Flesh and Blood candidates are fairly easily dismissed as holders of the narrative perspective. Taking the simplest case first, to say that Flesh and Blood audiences hold the narrative perspective would not answer our question. We accept that audiences possess an imagined analogue of the narrative perspective, and that this is part of what it is

---

<sup>113</sup> Booth (2005)



to engage with something as a narrative. What we are after, however, is an explanation of how they come to possess the narrative perspective in the first place. Suggesting that a flesh and blood audience holds the narrative perspective is therefore true, but by the lights of the debate in question, only trivially so.

A similarly simple refutation is available against the idea that the holder of the narrative perspective, the figure with whom we empathise in engagement with narratives, is the Flesh and Blood Author of the story in question. It is far from uncommon to see a Flesh and Blood Author deliberately create a story featuring a narrative perspective far removed from their own.

Wayne Booth reports that Saul Bellow once described to him the process of revising his novel *Herzog* as “wiping out those parts of my self that I don’t like”.<sup>114</sup> It is clear that many authors think, along with Bellow, that the narrative perspective on a work is at least partially under their control, that they have the power to dictate the authorial voice that an audience takes as the originator of the narrative perspective with which they engage. As Bellow erases from the text evidence of those attitudes of his he would prefer not to be expressed through the work he simultaneously constructs a narrative perspective lacking in those attitudes for the audience to engage with.

Of course not every author’s process of revision will be similar to Saul Bellow’s. It could well be that many authors revise their works in order to put as much of themselves into a story as possible, or to get clear on precisely what their attitudes are to the events related in the story. However, just because a revision of a story may be made with the objective of ensuring a faithful presentation of the Flesh and Blood Author’s actual perspective on the narrated events, that is no reason to think that any particular story consists in such a faithful presentation. I would hazard a guess that most, and certainly a large proportion, of

---

<sup>114</sup> Booth (2005)

authors have always put other, more artistic considerations before the strictly accurate rendering of their own perspective on a story. Even though as a matter of fact the attitudes or perspective expressed in any given work may be very close to those of the Flesh and Blood Author, it is still the case that much, if not most, of the time there will be significant differences between them. Hence, we are liable to be mistaken if we ascribe to the Flesh and Blood Author the attitudes or perspective expressed by their work.

Ordinary readers, however, often do take the perspective of a story to be that of the work's Flesh and Blood Author. There is something very reasonable about this commonplace habit, and something very misleading about it. What is reasonable about this, or rather the reasonable attitude it illuminates, is that when we read a story we are aware that it was written by a Flesh and Blood Author for some purpose. An important part of our engagement with any work of literature is often taken to be getting to know the author through their work. We speak of the 'world' of Austen, or Dickens, or Orwell; when we read their works we are aware that we are not just reading a story plucked from thin air, or randomly inscribed on a rock by the erosion of the sea, but a story written by a particular author according to a particular cultural tradition. While I do not wish to suggest that getting at Orwell's thoughts on the Russian Bolshevik revolution is the only reason we read *Animal Farm*, it is certainly among the reasons we might read it. Similarly, it is plausible that a typical feature of our engagement with any work of literature involves the belief that, in reading their work, we are discovering something of the Flesh and Blood Author's way of seeing the events described. Indeed, we can find some evidence of this in the habit of audiences to mistakenly attribute views they have taken from authored works to their Flesh and Blood Author – such as when some readers of *Animal Farm* and *Nineteen Eighty-Four* assume that his books betray broadly anti-

socialist feelings on the part of Orwell himself, when Orwell was famously sympathetic to socialism throughout his life.

However, because we have limited access to the Flesh and Blood Author's actual perspective on events such as those portrayed in any given story, and because we have to allow for the possibility that their perspective could differ wildly from that implied by the text, we are forced to conclude that the author whose thoughts and feelings we are discovering cannot be identified with the Flesh and Blood Author of the work.

Many of these points go double for the Flesh and Blood narrator of a story. Let me be clear that in speaking of Flesh and Blood Narrators of stories I am including those who read aloud stories that others have authored, those who tell stories that they author or make up themselves, and those who fall anywhere between, such as those who recite folklore or bed-time stories, or narrative-minded stand-up comedians and orators. Narrators who do not author their own material will obviously not always share the perspective of the narrative in question, and just as authors are capable of constructing a narrative perspective significantly different to their own, it is a simple matter for narrators to tell a story pretending to hold a perspective that they do not actually possess.

Neither Flesh and Blood Authors, Audiences nor Narrators appear to reliably share the narrative perspective. This means that they cannot serve as our postulated target for the empathic acquisition of the narrative perspective when we engage with stories. The same basic fact about all these characters, that they are Flesh and Blood and therefore in no relevant sense have a psychology determined by the story, appears to underlie their unsuitability as reliable holders of the narrative perspective. With that in mind, then, I now turn to figures internal to stories, whose features are determined by the storyteller, and in some cases the story itself, and therefore might conceivably fare better.

### 5.1.2 Figures internal to the narrative: Characters

Flesh and blood figures, then, seem obviously unsuitable to be considered as the holders of the narrative perspective. This is not surprising, mostly because Flesh and Blood figures have determined features that are not under the control of authors or storytellers, and the narrative perspective of any given story seems to be, in principle, under the control of the author.

Several philosophers have raised the question of how useful empathy is as a way of describing how we engage with stories. Amy Coplan, for instance, has enlisted a great deal of evidence from neuroscience and psychology to show that we frequently imagine ourselves experiencing a scene we read about from the point of view of the main character in the scene.<sup>15</sup> While this shows that empathy may well play an important part in how we orient ourselves in imagining narrated scenes, it by no means shows that empathy with characters is in any way central to our engagement with stories as a whole. Indeed, this last suggestion has been forcefully rejected by Noel Carroll, who argues persuasively that empathy with characters cannot be the principal mode of engagement with fictional characters.<sup>16</sup> The main force of Carroll's argument on this is due to an empirical point; if empathy with characters were our main point of engagement with fictions, we would expect to feel markedly different emotions in response to stories than we actually do in the majority of cases.

The opening scene in *Jaws*, for instance, provides us with two characters with whom we could plausibly empathise (three if we include the eponymous shark). Chrissy, swimming in the sea, is happily waiting for Tommy to join her in the water. The audience is made aware of a presence approaching Chrissy by the introduction of

---

<sup>15</sup> Coplan (2009)

<sup>16</sup> Carroll (1990); talks about 'identification' and not 'empathy', but they are equivalent for our purposes. His point also generalises to characters in any story, fictional or not, and beyond film to any other medium.

a POV camera shot moving beneath her in the sea. Tommy is exuberant at the prospect of swimming with Chrissie, and then passed out on the beach. The anxiety of the audience in watching the (at that point unknown) presence approach Chrissie beneath the surface is not shared by anyone in the scene. We even get to see from Jaws' point of view, and we certainly don't feel the same as way as Jaws does. If audiences habitually engaged with stories by empathising with their characters, then we would expect their emotion to be one exhibited by, or at least ascribable to, one of the characters in the scene.

Carrol rightly suggests that the mode of engagement we have in the majority of cases, exemplified by *Jaws*, is one of *sympathy* for the characters. We care for Chrissie's wellbeing, and so we feel anxious for her safety after we see her from Jaws' point of view. The same, it should be emphasised, goes for non-filmic stories too. Many of the characters in most stories elicit sympathetic responses, and the audience empathises with relatively very few. I don't doubt that readers can be encouraged to empathise with particular characters, especially when one of them is serving as the story's narrator.

However, it should be noted that of course, many characters in stories are targets for neither our sympathy nor empathy. Villains, bit players, anonymous crowds; we rarely sympathise or empathise with any of these. If we argued that our main way of engaging with stories was to empathise or sympathise with characters, we would be left with the question of how we select which characters we sympathise with, which we ignore, and which we are hostile towards.

How we select which characters to have which attitudes towards is a matter for the narrative perspective. It might be argued that the narrative perspective does no explanatory work here; we can say that we sympathise with Chrissie because she is attractive and happy, or that we empathise with Oliver Twist because we spend so much time with him, or that we're hostile to Gradgrind because he is described in such unpleasant terms. However, there are no general principles of

this sort which hold across all cases. Cersei Lannister is attractive and at times happy, but we do not sympathise with her; we spend a lot of time with Jason Voorhees but we don't empathise with him; and Gollum is described in very unpleasant terms, but we aren't especially hostile to him at any point. The question of what narrative techniques can encourage us to take what attitudes towards the characters is interesting, but it is not our question. We are interested in contributing to a general model of how we engage with stories and we can't extract general principles from these kinds of narrative techniques.

Furthermore, narrative techniques like describing a character in attractive or grotesque terms, or having a character appear in a majority of scenes, underdetermine our responses to those characters. Our responses to a character are often not principally determined by how they are described in the current scene, but rather by the characterisation we have formed of them from previous descriptions. It does not seem right to say that it is the previous descriptions themselves which are proximate causes of my current attitude, since I may well have no memory of the descriptions *per se*, merely the impression of the character that they gave me. Even if we denied this, there are also many cases where it would not even be right to say that my response to a character is proximately determined by the sum of the descriptions of that character, nor by the direct effect of having one description contradict or contrast to a previous one.

As an example we might think of a character, similar to *Rebecca's* Maxim, who we have grown to like over a long story in which he is the protagonist, and who is consistently described in a pleasant way. Then, in one chapter, in a moment that springs naturally from a pending frustration in his life, he does something morally reprehensible. Now, it is tempting to say that our resulting revulsion is due to the description of this character being hitherto pleasant, but now involving him in terrible misdeeds. However, this isn't quite

right. Our revulsion is caused by our kind *feelings* towards this character and how they contrast with his surprisingly evil deed, it is not caused by the contrasting *descriptions*.

So we can see that our attitudes and responses to characters owe something to our standing dispositions to characterise events and people in particular ways; to the ways of seeing the story that the narrative perspective generates. While I have shown that we can't accurately say that we will always sympathise or empathise with any particular character, there is certainly one who we might think that we empathise with more often than not (where they appear). That is the internal narrator of the story, and it is to them that I now turn.

### 5.1.3 Figures internal to the narrative: Internal Narrators and Unreliability

Characters who narrate the stories in which they appear do seem to be able to direct our attention in the way we would expect of the narrative perspective. The way a narrator thinks about the story is often the way that the audience thinks about the story. So although not all stories are narrated by one of their characters, it will still be profitable to determine what features of an *internal* narrator's perspective are shared with the narrative perspective.

Whether a figure, or a perspective on a story that a figure may hold, is internal or external depends on whether or not the figure or perspective in question is located within the narrated events. To illustrate this, let us think about the narrators of stories who might possess each of these kinds of perspective; a narrator with an *internal* perspective will be one who features in the events of the story they tell, a narrator with an *external* perspective will not feature in the events of their story.

Dr. Watson of Conan Doyle's *Sherlock Holmes* stories is a good example of an internal narrator. The stories Dr. Watson tells are of adventures that he has with Sherlock Holmes. Even in the story *The Gloria Scott* in which we learn about Holmes's first case, the events of which occurred long before Holmes and Watson had met, Watson is

still an internal narrator. This is because *The Gloria Scott* is a story Watson tells of Holmes telling Watson about his first adventure. The story of Holmes telling Watson a story is itself a story Watson is intimately involved with.

It is important to note that Watson needn't have been an internal narrator here. Had he chosen just to relate the events as Holmes related them to him he would have been an external narrator.

However, because the story he tells involves Holmes telling him the story he becomes an internal narrator. Holmes too is an internal narrator in this story, since the story he tells Watson is of events in which Holmes was a participant. Conan Doyle on the other hand is an external narrator of both the story Watson and Holmes tell; he tells a story in which Watson tells us of Holmes relating the details of his first case.

Internal, first-personal narrators such as Watson, in many cases do hold the narrative perspective of the story in which they feature. Watson's perspective on the Sherlock Holmes stories is usually the one we share, according to which Holmes is admirable but flawed, Lestrade and Scotland Yard are pleasant but often bumbling and incapable, with a moral outlook befitting the late Victorian era in which the stories are set. However, there is one aspect of Watson's perspective which is not shared by the audience; that of his own intelligence and capability. Watson presents himself as almost comically incapable, but from Watson's insight, and Holmes' own assessments, we know that he is not as foolish as he thinks himself to be.

Watson, in his underestimation of his own abilities, betrays himself to be a mild form of unreliable narrator. Unreliable narrators come in several different forms, and they often unreliable in their perspective on the events they relate as much as on the facts of the story they tell. Watson, as I said, is a mild case, but far more stark examples exist in



Stevens the butler in Ishaguru's *The Remains of the Day* and Patrick Bateman in Easton Ellis' *American Psycho*.

*American Psycho* is the story of a wealthy investment banker Patrick Bateman. Bateman, in his late 20s when the story begins, narrates his everyday activities, from his recreational life among the Wall Street elite of New York to his forays into murder by night. *American Psycho* is a fascinating book for our purposes here, since it is likely to give rise to both Transmission Failure and Disparate Response in much of its audience. This is because the response sought by the work includes some degree of sympathy for the protagonist, who is presented as something of a victim of the capitalist excesses that have made him successful. His murderous activity is itself presented merely as the taking of Bateman's ultra-capitalist dogma, for which he has been consistently rewarded, to its logical extremes. Bateman, by the end of the story, is presented as a pathetic and pitiable figure; torn between his helpless compulsion to kill, scared by his own lack of remorse and how easily he gets away with his own crimes. He is conflicted in that he is desperate to be caught and punished, though he still works hard to avoid capture, and is bemused to find himself still protected and rewarded by the society that has fomented and encouraged his psychopathic behaviour.

Significantly for our current purposes, there is no character in the work that shares this perspective on the events of the story, and certainly not Bateman himself, who narrates it; he is alternately proud and scared, but never sees himself as pitiable; he is too self-involved to think of himself as befitting any particular judgement of others. The perspective according to which Bateman is to be pitied, then, must be one *external* to the story, as there are no figures available with an internal perspective that provokes the required responses.

*American Psycho* illustrates two important points. First, that the phenomenon of Unreliable Narrators means that first-personal narrators can't always be the holders of the narrative perspective. And

second, that there is not always a character in a story who shares the narrative perspective.

To expand on the first point, unreliable narrators, as I have mentioned, come in several different forms. It is important to note that not all types of unreliable narrator are relevant to our discussion. Some unreliable narrators are unreliable on matters of fact, and others on matters of perspective. Within each of these categories of unreliability we can further divide narrators between those who are deceptive (deliberately unreliable in order to mislead the audience) and deceived (unreliable, but not deliberately). Determining whether a narrator is reliable on the facts of a story is both complex and not relevant to our current discussion. The narrator of a realistic or purportedly true story may be said to be unreliable by virtue of failing to make their story correspond to facts in the world. If a story is not purporting to be true then it is a more difficult matter to determine what counts as a narrator who is unreliable on the facts, although some clearly are (such as the narrator of Agatha Christie's *The Murder of Roger Ackroyd*). However, we are interested not in the ways narrators can fail to be reliable on facts, but rather how they can fail to be reliable in the way they present those facts, since that is the way that a narrator can fail to share the narrative perspective.

So how, exactly, does a narrator fail to be reliable in their perspective on events? At this stage that notion is in danger of seeming incoherent; we do not, in general, think of perspectives as the kinds of things that can be held reliably or not. Certainly we can say that some perspectives might be more warranted or appropriate than others, but the term 'unreliability' suggests that there is a standard of correctness to the perspective of a narrator that we would not usually apply to the perspectives we ordinarily hold. Despite this, narrators such as Stevens in *The Remains of the Day*, who consistently sees his employer's moral fall into Nazi sympathy and apologism as noble and

praiseworthy, seems a prime example of unreliable narration, and seems to be unreliable only on matters of perspective and not fact.

I suggest that the best way to make sense of unreliable narrators like Stevens and Bateman is to say that they are unreliable in the sense that their perspectives depart significantly from the narrative perspective of the story in which they feature. To say that a narrator is unreliable by virtue of their perspective is to say that another perspective is more appropriate, but importantly that more appropriate perspective must be another that is offered as a way of engaging with the story. Consider a situation where we may have a narrator who has a perspective that we don't consider appropriate, but where no alternative perspective is offered by the story. As an example, we might think of an analogue of (G) narrated by Giselda, like so:

(G)\*\* "My name is Giselda. In killing my new born daughter, I did the right thing. After all, it was a girl".

As we have discussed, such stories are liable to result in Transmission Failure. However, it does not seem right to put this failure down to the unreliability of Giselda's perspective; her perspective seems to accord with that of the story as a whole. In such cases we do not want to say that the narrator is unreliable, rather that the perspective offered is flawed for some other reason (such as that it is too far removed from our own to make engaging with the story worthwhile).

In cases where we can say of an internal narrator that they are unreliable on matters of perspective, it seems that we must postulate a further perspective with both has primacy over theirs (in that the audience will prefer it), and that differs in some significant way. The second point illustrated by *American Psycho* and others, that no other character in the story necessarily holds such a perspective, gives us another reason to think that the narrative perspective is both a significant part of our engagement with stories, and not identifiable with the perspective of any figure internal to the story.

#### 5.1.4 *The Ubiquity Thesis and Effaced Narrators*

One possible response to the second point illustrated by *American Psycho*, that there is no character internal to the story who shares the narrative perspective, is that there is one possible internal perspective that I have overlooked – that of an ‘effaced’ narrator. According to the view commonly known in the Philosophy of Literature as the ‘ubiquity thesis’, in every work of fiction there is some fictional narrator who tells it.<sup>17</sup> I am interested here in all stories, not just fictions, but for the sake of argument let us allow for the moment that the Ubiquity Thesis can also be taken to apply to non-fictional stories, such that every story that is told is embedded within another story told by a character internal to the story. That narrator is ‘effaced’ because they are not explicitly mentioned by the story.

In opposition to the ubiquity thesis, we can point to the unfamiliar imaginative activity that would be necessary if it was true; especially in cases like *American Psycho*, where we must imagine not only that we are being told a story by Patrick Bateman, but that the story of Patrick Bateman telling us a story is also being told by some further figure in the story of whom we are not otherwise aware. It has also been argued that the ubiquity thesis cannot cope with ‘mindless’ fictions: fictions that explicitly state that there is nobody in the fictional world of the story, or that the events of the story left ‘nobody alive to tell the tale’. An early defence of the ubiquity thesis from Gregory Currie points out that there are stories “that generate games of make-believe in which we are called upon to make believe that P and to make believe that

---

<sup>17</sup> Alward (2006) opposes this definition, claiming that we can consider a ‘narrative informant’ to be “neither actual nor fictional, yet nevertheless reporting on a fictional world she does not inhabit.” The only plausible reading I can construct of this claim is that the narrative informant is fictional, but does not belong to the same fiction as that of the story she relates; we are engaged in a fiction within which another fiction is constructed (Currie (2010) makes use of a similar notion of embedded stories). If this is right it does not seem like his ‘non-actual fact-teller’ is telling fact at all, but is rather fictionally fiction-making.

not-P.”<sup>118</sup> This is not uncontroversial, but even if it is true there is a difference between the claim that it is possible that we imagine contradictory things in response to some fictions and the claim that we do so as a matter of course in encountering mindless fictions; the latter still seems implausible. The example Currie uses has to do with time travel; that we may believe that an event both did and did not occur at a particular time according to a science fiction yarn. It does not seem right to draw from the fact that we might, feasibly, believe a contradiction to be true according to a fiction that we can thereby assume that we believe contradictory things are true of any mindless fiction. There is a great deal of cognitive effort required to imaginatively navigate logically troubling stories such as those involving time travel paradoxes, and this kind of effort simply does not appear to go along with engaging with the mindless fictions Currie points to. Furthermore, it could be argued a similar cognitive dissonance is required by Currie to occur as a matter of course when engaging with any fiction narrated in the first person.

When I introspect on my imaginative habits when reading first-person narrated fictions, I can find no suggestion that I imagine someone other than the explicit narrator telling the story. Currie may yet insist that I do imagine that the story is being told by somebody else in the voice of the narrator I recognise. However, because all I am aware of imagining is being told a story by the narrator, Currie’s insistence that I am also imagining a Fictional Author who tells the story of the narrator telling the story puts me in the position of being mistaken about who is telling me the story whenever I read a first-person narrated fiction. I do not suggest that Currie’s view always entails my imagining contradictory propositions, that “Watson is telling a story” and “the effaced narrator of *A Scandal in Bohemia* is telling a story”.

---

<sup>118</sup> Currie is here defending his proposal (from which he has now distanced himself) that a sentence of the form ‘in the fiction F, P’ is true if it would be reasonable for an Informed Reader to infer that the Fictional Author of F believes that proposition to be true. Currie (1990)

According to Currie the former telling is embedded in the latter, so there is no contradiction there. Rather, the problem comes from my attitude about, in this case Watson, that he is the primary teller of the story according to the work. As far as I'm concerned when I read the Holmes stories, it is Watson and only Watson who is telling me a story in the fiction (though of course I also believe that Conan Doyle is telling me a story at the real world). When I read Sherlock Holmes, Currie's view seems to accuse me of being mistaken not just about the fact that there is an effaced narrator, but also that it is this effaced narrator, and not Watson, who is my primary source of engagement in the world of the fiction.

Possibly the most significant problem with the Ubiquity Thesis as it stands, however, is that it fails to adequately account for unreliable narrators (which is one of the reasons we turned to effaced narrators as potential holders of the narrative perspective in the first place). While a defender of the Ubiquity Thesis might claim that we do not need to habitually postulate an effaced narrator, but only do so where there is either a) no narrating character or, b) we have reason to think the narrating character is unreliable, this will not do.

Bonomi and Zucchi propose exactly this in order to overcome similar difficulties in Currie's theory to the ones I outlined here. They claim that a presumption of reliability is "at work" when we engage with fiction, and that we only question the reliability of the narrator when there are "reasons intrinsic to the [story]" that we should do so.<sup>119</sup> We are justified in, say, questioning Stevens' view of Lord Darlington because Stevens shows himself to be subservient to his master's perspective on the world.

There is certainly something correct about this view; as a description of how we approach first-person narratives I think it does justice to

---

<sup>119</sup> Bonomi & Zucchi (2003)

our intuitions. However, it falls short of a cure-all for dealing with unreliable narrators. The presumption of reliability faces the problem that in order to discern what reasons “intrinsic to the [story]” are sufficient to question the reliability of the narrator, we must already have some beliefs regarding what is an appropriate perspective on the narrator’s story. In the case of Stevens’ view of Lord Darlington, we must already have some idea that the appropriate perspective on the story is such that Stevens’ is at odds with it.

In other words, for a presumption of reliability to work we need to be justified in operating under the assumption that the world of any story is somewhat like our own. We need to be able to justify the conditions at which our narrator becomes unreliable, and in order to do this we must know how much of our intuition about what it would be reasonable for a narrator to report apply to any given fiction. As an example, imagine that Orwell had chosen to write *1984* in the first person, with Winston Smith as the narrator. Smith would certainly appear to be paranoid, which would count as a reason ‘intrinsic to the text’ to suspend our trust in the narrator’s version of events. If we were to know that, in fact, Smith is to be trusted in this case we would have to have a justified belief that the world of *1984* is sufficiently unlike our own that reports of what at the actual world would be chronic paranoia can be taken as reliable fact in the world of this fiction. Of course, in this case, we could not come by that belief without believing Smith’s narration, and so we are stuck. *A fortiori*, consider the case of an impossible fiction told in the first person (in which Smith and Jones square a circle, for instance) – a presumption of reliability would be immediately suspended in such a case, leaving us to unfairly label the narrator a liar. A presumption of reliability, defeasible where the narrator exhibits certain characteristics (intrinsic to the fiction which make the narrator seem unreliable), will make false predictions whenever the world of the fiction is sufficiently

unlike our own that the narration given would be thought unlikely if it was told as fact at the actual world.

To drive the point home, let us imagine a fiction that was primarily narrated by such a character as Currie thinks we imagine in addition to any explicit narrator in the ordinary case. Suppose we are being told the story of Hamish by, according to all appearances, Hamish himself. However, at the end of the novel it is revealed that Hamish died during the events of the story and actually the narrating character is Dougal, who was adopting Hamish's persona for dramatic effect but nonetheless believes everything he reported was an accurate account of events. According to the Ubiquity Thesis we would in such a case have to postulate yet another character in the story, on the off chance that Dougal could have been mistaken about his beliefs. The possible iterations of this wisely underused narrative flourish border on the absurd, but however absurd the story itself the insistence that a reader must posit a further character, just for the sake of having a reliable perspective somewhere internal to the fiction, is even more difficult to swallow.

We might be able to forgive the Ubiquity Thesis for its clumsy handling of mindless fictions, given that they are something of a rarity. However, the insistence that a mindless fiction entails readers so often being mistaken about who is the primary source of their perspective on the story, together with its continued inability to deal with unreliable narration, leads me to conclude that the idea that the narrative perspective is always held by a character internal to the story is false.

#### *5.1.5 Hypothetical and Implied Figures*

The Ubiquity Thesis fails to provide us with a figure who necessarily holds the narrative perspective, but this is arguably due to features of the Ubiquity Thesis that we might be able to change. The Ubiquity Thesis was initially formulated in order to address the venerable question in analytical aesthetics of what makes something true in a



fiction. A strand of thought, beginning with Currie and continuing until now, has it that the best way to analyse truth in fiction is in terms of the beliefs of a reliable narrating agent internal to the fiction. If we are not altogether interested in fictional stories as much as stories *per se*, however, there does not seem to be a binding reason why we shouldn't use the Ubiquity Thesis, should it prove useful, but stipulate that the ubiquitous narrating figure need not belong to anything as difficult to manage as a character internal to the story.

A weak version of the ubiquity thesis claims that in engaging with a story, the audience postulates a 'minimal narrating agency' who is telling that story; as Seymour Chatman has it, "every narrative is by definition narrated".<sup>120</sup> If every narrative is narrated by definition, then by definition every narrative is narrated by some narrating agency. So much constitutes conventional wisdom, and those who criticise this notion tend to only do so because they think that the minimal narrating agency that we take to exist by definition is insufficient as a means of answering various philosophical questions about the nature of engagement with stories.<sup>121</sup>

I will argue here that this minimal narrating agency is the holder of the narrative perspective, that the minimal narrating agency that audiences postulate as a part of engaging with any story is psychologically rich enough to hold things like perspectives, and that such an agency serves as the target of empathy.

In the previous section, I made a great deal of the fact that characters in a story can't be the holders of the narrative perspective because their perspectives are necessarily *internal* to the story, whereas the narrative perspective is not. I want to clarify at this point that, although the narrative perspective is not necessarily internal, neither is it necessarily external.

---

<sup>120</sup> Chapman (1990), p.115

<sup>121</sup> See, for instance Levinson (1996(a)), cf. Kania (2005)

The narrative perspective will usually be external; it will usually be held by a figure who is not a part of the story they tell. We know this because, as we have seen, the narrative perspective usually cannot be identified with the perspective of any character explicitly named in a story, and the idea of implied or effaced internal narrators turns out to be very troublesome. However, there are cases where the narrative perspective is clearly an *internal* one.

I have in mind autobiographical stories of various kinds. *The Autobiography of Malcolm X*, to take a prominent example, is a case where the perspective of the narrative is clearly framed from the perspective of one who is involved in the story, and so the narrative perspective is internal. Similarly, for any story I tell about a party I attended, or the funny thing that happened to me on my way over here, the narrative perspective will be internal, since I am telling a story about a thing that happened to me. Not only are these cases most simply described by recognising the narrative perspective as internal, but some responses that audiences have to stories of this kind require that the perspective they adopt is internal also.

For instance, suppose I'm telling a story about a party I attended. At the time I thought I was being the life and soul, the very picture of bonhomie. On sober reflection I realise I was actually being insensitive and boorish through the whole night. The story I tell about my evening expresses a shamed perspective on the events of the party, particularly highlighting the pride I took in my 'wit'.

We might expect that an audience of this story would respond with the kind of interesting 'cringe' response that is so often associated with embarrassing stories, and comedy figures such as Alan Partridge and David Brent. Furthermore, the cringe will be directed at me, the storyteller, and not merely at the version of me who misbehaved at the party. Such a response, I think, is only appropriate if a couple of conditions are met. First, my behaviour at the party must be recognised as embarrassing by the audience. It is recognised as

embarrassing by the audience because when I tell the story, I tell it with the benefit of hindsight, in the knowledge that my behaviour was in fact embarrassing, and not sociable as I thought. Second, and most importantly, the audience must recognise that I have an internal perspective on the story, in that I am involved in the story that I am telling, in order that they can properly direct the response at me, the figure who was both at the party and is telling the story.

The lesson to take from this, I think, is that the narrative perspective must be capable of moving between being internal and external to a story. The narrative perspective is sometimes internal to a story, and sometimes external, and so the figure who holds the narrative perspective must not be restricted to only being able to hold one or the other.

This consideration allows us to rule out one of our final candidates for the holder of the narrative perspective, that of an implied or hypothetical audience. An audience's perspective will almost invariably be external to a story, and so is not capable of the flexibility that we now require. Furthermore, hypothetical or implied audiences are inadequate for the equally significant reason that the perspective of a hypothetical audience would be determined by the way an idealised audience would respond to an idealised author.<sup>122</sup> This is because a hypothetical audience would need to be described in terms of its ability to respond in a way appropriate to the content of the story, and the appropriateness of any response to a story would, in turn, be likely to be described in terms of the intentions of the actual or implied author of that story. If a hypothetical audience holds the narrative perspective, therefore, a hypothetical author would hold it too, making the claim that the hypothetical audience holds the narrative perspective explanatorily redundant.

---

<sup>122</sup> Byrne (1993)

Now we come to my recommendation for the figure that we should take to hold the narrative perspective, and with whom an audience empathises during their engagement with a story. That is, an implied storyteller, encompassing the ‘minimal narrating agency’ that is present by definition in all narratives. I will expand on my reasons for this choice shortly, but first I want to clarify why I think an implied storyteller, and not an implied author, is the right option.

The implied author is a concept developed in 20<sup>th</sup> century literary criticism, most notably by Wayne Booth.<sup>123</sup> Booth’s theory was largely a response to notion popularised by Flaubert, Lawrence and others that there was something at the best unappealing, and at worst downright immoral, about an author’s own judgements about subjects of a novel weaving their way through the narrative. In contrast Booth holds that such objectivity is neither especially desirable nor possible: “However impersonal he may try to be,” writes Booth, “his reader will inevitably construct a picture of the [author] who writes in this manner”.<sup>124</sup>

Although the usefulness of the implied author as a heuristic device has never been successfully impugned, the ontological commitments associated with it have often been queried. Is the implied author the kind of thing that can communicate?<sup>125</sup> If so, implied authors are going to need to carry significant ontological costs. How far should we take the apparent anthropomorphism of the concept?<sup>126</sup> If implied authors are people, are they fictions? Simply imagined? Are they created by authors or by readers? If the former, how can they be truly independent of a work’s flesh and blood author in the way that seems to be required? If the latter, does that mean there are one or more implied authors for each reader?

---

<sup>123</sup> Booth (1961), (2005)

<sup>124</sup> Booth (1961)

<sup>125</sup> Chatman (1978)

<sup>126</sup> Rimmon-Kenan (1983)

Choosing the implied narrator rather than the implied storyteller allows us to evade all of these tricky questions (provided we can show that the implied storyteller is capable of the psychological refinement necessary to hold perspectives). More than that, however, it is unclear whether an implied author has the right kind of attitudes to their (implied) work such that an audience who empathised with them would come away with the attitudes and responses that we would want to predict. An implied author, for instance, might be assumed to have an authorial concern for their characters, or similar attitudes towards the story that come along with having authored it. We would generally not want to suggest that an audience, in empathy, would subsume these traits into their perspective.

An implied narrator, on the other hand, carries with it all the relevant benefits of an implied author, in that they are somebody an audience already has in mind when they engage with a story, and they have the right kind of relationship to the story itself to allow their perspective to be internal or external as each case requires. However, an implied storyteller comes with very little of the associated ontological cost of an implied author, largely because we are already committed to a minimal narrating agency who can fulfil the role of an implied storyteller. An implied author, on the other hand, can't be easily said to exist in every story by definition.

To conclude this section, then, let me summarise the reasons why an implied narrator is the figure we can take to hold the narrative perspective in any given story. First, an implied storyteller is a figure that uncontroversially exists across all kinds of story, in that any narrative is narrated by definition. Secondly, since the perspective of an implied storyteller can either be internal or external to the story they tell, we can account for the narrative perspective of autobiographical stories. Third, implied narrators evade the problems of unreliability that prevented internal narrator-characters from necessarily sharing the narrative perspective. This last is due to the

fact that implied narrators, by virtue of encompassing the minimal narrating agency entailed by narratives, serve by definition as the primary point of engagement with any story. As such, an implied narrator also serves as the standard of reliability for any other narrating figures embedded within the story they tell, and are therefore incapable of being unreliable themselves on matters of perspective.

## 5.2 Demonstrative Ascription in Narrative Engagement

My task in this chapter has been to show that engagement with stories involves making a demonstrative ascription to a target, in order to secure the idea of empathy being a central way that audiences engage with stories. I have identified the most plausible target of empathy, the implied narrator of the story, by virtue of its having the right kind of relationship to the story to plausibly hold the narrative perspective. There is only one thing which remains to be argued for; that we routinely make demonstrative ascriptions to implied narrators in the course of engaging with stories.

I will begin discussion of this point by way of anticipating an objection to the idea that we need to ascribe the narrative perspective to any figure at all. Currie has recently argued that there is no need to attach the narrative perspective (or, in his terms, the narrative framework of a story) to a figure such as an implied narrator. Stories, according to Currie, express their perspectives in the same way as someone who shakes a fist expresses anger. If the person shaking their fist is obviously not angry, we can still say that anger is expressed (perhaps

ironically), but we do not feel the need to ascribe the expressed anger to anyone.<sup>127</sup>

According to Currie our principal mode of engagement with stories, and in particular their perspectives, is one of *imitation*, not empathy. We imitate the expressed perspective of a story, and this is why we respond to that story in a way that accords with the perspective so expressed.

However, I doubt that we can account for our responses to stories by relying solely, or even principally, on our tendency to imitate (although I don't doubt that such a tendency is prevalent in us.)

Currie acknowledges that:

“We are so prone to imitate and so little aware of what we are doing that I expect that empirical studies, if they could be undertaken with a suitable fineness of discrimination, would reveal a vast amount of imitative activity as a part of narrative engagement, very little of which is conscious but any part of which has the capacity to contribute to our conscious sense of how interesting, effective and valuable the narrative is.”<sup>128</sup>

The problem this reveals is that our imitative habits, in Currie's view, guide far more than our imitation of a narrative's perspective. There is no doubt that many things other than the narrative perspective influence our immediate responses to a story; we might think of jump scares in film or spoken stories, for instance. So relying on imitation as an explanation for our engagement with stories, without a way to discriminate between the imitation of the narrative perspective and the imitation of other things that might lead to responses for which the narrative perspective is not responsible doesn't offer us a satisfactory explanation of the role of perspective itself in engaging with a story. In preceding sections, for instance, I have discussed the

---

<sup>127</sup> Currie (2010)

<sup>128</sup> *Ibid.* p.105

ways in which our adopting the narrative perspective explains how some of our responses are in tune with ‘the story’ as a whole, while still differing from the responses we might expect to have if we were imitating the perspective of some character or other within a scene. For instance, if Falstaff is, according to the narrative perspective, a laughable figure, then that explains why we enjoy his suffering, rather than imitating his sorrow at his misfortune.

Conceiving of narrative engagement as empathising with the holder of the narrative perspective gives us a way to discriminate between those characters whom we may imitate, empathise or sympathise with, and those who we disregard. Engaging with a story isn’t just a matter of imitating whatever imaginative fodder the narrative provides for us. It involves the characterisation of scenes in particular ways, in accordance with the perspective from which the story is told.

Empathising, then, in virtue of being a way to adopt another’s particular perspective, gives us all the necessary tools to respond to characters and events in stories in ways that are consistent with a relatively stable perspective on those stories.

So why, in the end, do I think that narrative engagement is a matter of empathy? Where empathy, on my terms, is a demonstrative ascription to a target of the form ‘[the target] feels like *this*’. The evidence for this notion is, I think, fairly straightforward once we have fixed the implied narrator as the target of the empathy; it essentially consists in the fact that the empathy model of engagement with stories predicts a familiar phenomenology of stories, predicts the responses that we in fact have to stories, and accords with our experience of engaging with stories of all kinds. For instance, some evidence for the empathy view is revealed by our habit of describing stories themselves in terms that we would also apply to perspectives. We might say that a story is pessimistic or optimistic, sentimental, romantic or naïve. The best way to make sense of this habit is to suggest that we are ascribing those perspectives to the story itself, by way of ascribing them to the



narration of those stories from a particular perspective which shares those characteristics.

With this in mind, I propose the following as a theory of empathic engagement with stories:

When we engage with a story, we empathise with the implied narrative agency of that story. In doing so, as per the Complex Theory of Empathy, we engage in proto-empathic imaginings consisting in (a) doing-with (wherein we entertain the idea that the facts of the story hold for the purposes of engaging with the story), (b) being-with (wherein we adopt the narrative perspective of the story) and (c) feeling-with (wherein we emotionally respond to the story in a way congruent with (a) and (b)). All this is done with the aim of making a demonstrative ascription to the implied narrator of the story of the form ‘the narrator of this story feels like *this* about this story’. The way we can characterise ‘*this*’ will also be an appropriate characterisation of the story.

To show that the above model is an accurate one, I will conclude with two final points. First, I need to claim that we, as audiences, have the right attitude towards the implied narrator. And second, I need to illustrate that the above model will avail us of appropriate and predictable responses to the stories with which we engage.

On the first point, it is clear that when we engage with a story we are aware that we are engaging with a story authored by some person for some purpose; we are involving ourselves in the institution of storytelling in the role of audience. Furthermore, it is also clear that in engaging with a story we are non-trivially concerned with the attitudes of the storyteller (identified here as the minimal narrating agency).<sup>129</sup> We look to the storyteller to tell us the facts of the story, and we also look to them to show us how we should respond to those

---

<sup>129</sup> Although in many cases the minimal narrating agency will be subsumed by an explicit narrator.

facts. Usually we look to the storyteller by way of looking at the story they provide us with, which is the expression of their pronouncements about what is true in the story, and also the result of their perspective on it. To say that *S* is a sad story is to say that the narrator of *S* told a sad story. We can go further, in fact, and say that to respond to *S* as a sad story is to respond to the narration of *S* in a sad way. If our narrator tells a sad story, it is fair to make a demonstrative ascription such that they also think that story is sad. Because we are interested in what our storyteller thinks and feels about the story, since a narrator's expression of what they think and feel about the story constitutes the storytelling, we are liable to ascribe to them the thoughts and feelings we take their story to betray.

Once we have succeeded in empathising with our storyteller, however, we need no longer wonder so much as to what our storyteller feels about the story, since we have already adopted what we think is the appropriate perspective for the story they are telling (i.e., the one that matches theirs). Once we have successfully empathised with the implied narrator of the story, we will respond to the story (all things being equal) in the way that the narrator intends, by virtue of sharing that narrator's way of seeing the story. They will direct us to which characters and events we are to take as especially prominent, which characters we are intended to sympathise with, and which we are supposed to dislike, all by virtue of the perspective the narrator lets us know is the one we should share.

## 6. EMPATHY AND NARRATIVE ENGAGEMENT: APPLICATION AND EXPANSION

### 6.1 Empathic Engagement and Moral Persuasion

We often credit artworks, especially narrative fictions, with having some significant role in the evolution of our moral attitudes. Assume that we are not mistaken, and that reading these works did in fact change something in our moral outlooks. If that is the case, then we are confronted with the question of under what circumstances and to what degree I can hold the works themselves responsible for those changes.

Shen-Yi Liao has argued that much of our theorising on this point has been contaminated by a tendency to focus on *realistic* fictions as examples of how fictions in general can morally persuade.<sup>130</sup> Liao limits the relevant sense of realism to the realms of the psychological and the moral, so that a work counts as 'realistic' if the characters think and act in ways that could occur in real life, and the moral facts of the fictional world are identical to those that hold at the real world. What counts as realistic fiction for these purposes, then, may well include works in the genre of science-fiction, fantasy and the like, provided that psychological and moral facts in these works are presented as realistic.

Liao advocates a theory, *Persuasion Variantism*, which challenges the dominance of what he calls *Persuasion Invariantism*. Persuasion invariantism is the view that:

---

<sup>130</sup> Liao (2013) pp 272-3, see also Harold (2007)

“a fiction is responsible for getting us to really adopt a moral (or immoral) real-world outlook when it is responsible for getting us to imaginatively adopt a similar moral (or immoral) make-believe outlook.”<sup>131</sup>

With variantism just being the view that the invariantist thesis is false.

Liao defends a version of the former which he calls *Genre Variantism*.

A couple of notes on terminology before we proceed: ‘Moral outlook’ is intended to be a broad term capturing not just moral judgements and associated cognitive states, but also ways of organising our thoughts about particular topics and situations as well as our emotional and other responses to them. Bearing in mind the preceding chapters, we can think of moral outlooks as those parts of our perspective which are relevant to our moral life.<sup>132</sup> Secondly, the sense of ‘being responsible for’ here importantly has to pick out a moral as well as a causal responsibility. We’re looking for a way to hold a work morally culpable for persuading people to hold (im)moral perspectives.

A key consideration in determining whether or not a work is morally responsible for my adopting a perspective will be whether my adopting that perspective was a predictable result of my engagement with the work. Persuasion invariantism, Liao claims, is insufficiently sensitive to the diversity of fictions, and as such often incorrectly predicts what perspectives we can hold works responsible for persuading their readers to adopt. Liao claims that different genres of fiction avail themselves of different ‘modes of persuasion’, some of which do not operate according to the general principle outlined by the persuasion invariantist view. In particular, he highlights satire as a genre that tends to morally persuade not by getting the audience to imagine a moral perspective, *m*, but rather by getting the audience to imagine holding a perspective antithetical to *m*.

---

<sup>131</sup> Liao (2013) p. 273

<sup>132</sup> Ibid. p.270

In this section I will offer arguments to show that, although Liao is right to draw attention to the variety of ways works of fiction can morally persuade, these can be accommodated by persuasion invariantism. I will also give some reasons for thinking that persuasion variantism of the sort Liao advocates is not in a good position to explain how works can morally persuade, nor to provide general principles according to which we can hold works responsible for their moral effects on readers.

### 6.1.1 *The Case Against Persuasion Invariantism*

Liao endorses an interpretation of *Catch-22* posited by James Harold according to which certain views we would find horrifying in the real world are, in the world of the novel, not reprehensible.<sup>133</sup> *Catch-22* is, therefore, not realistic (in the sense picked out by Liao), since according to this interpretation the moral facts in the world of *Catch-22* differ from those at the actual world. Liao claims that “persuasion invariantism would tell us that *Catch-22* is morally corruptive because it is responsible for getting us to imaginatively adopt an immoral make-believe outlook.”<sup>134</sup> Since *Catch-22* is generally considered, if anything, a morally praiseworthy book it would indeed be a problem for persuasion invariantism if it predicted that *Catch-22* would be morally corruptive. However, it does not.

Persuasion Invariantism, as stated by Liao, merely says that for any real-world moral perspective possessed by an audience of a work, that work is responsible for it when it prescribes the audience to adopt a make-believe version of a similar perspective.<sup>135</sup> Persuasion Invariantism does *not* claim that *every* make-believe perspective prescribed by a work results in a similar real-world attitude being

---

<sup>133</sup> Harold (2007) pp 149-50

<sup>134</sup> Liao (2013) p.273

<sup>135</sup> I would add to this that if what we are trying to pin down is a work's capacity for moral *persuasion*, then we should stipulate that the work is only responsible for a real-world outlook if the audience did not possess the outlook in question before engaging with that work.

adopted. This would be clearly false: there are many characters whose moral perspectives we are prescribed to imaginatively adopt, but whose moral perspectives are perverse even according to the world of the fiction they inhabit. We might think again of Patrick Bateman in *American Psycho*, with whom we are called upon to empathise at points during the story, but whose perspective is not the same as the *narrative* perspective of *American Psycho*, the latter being the determinant of the moral perspective of Bateman's world.

Persuasion invariantism does not need to claim that whenever a work prescribes that we adopt a make-believe perspective it is thereby attempting to persuade us to adopt a similar real-world perspective. It does need to claim, however, that absent the prescription for an audience to make-believe a perspective, that work cannot be said to have attempted to persuade us to adopt a similar real-world perspective.<sup>136</sup>

The focus of our investigation should not be, as Liao takes it to be, whether it is possible for us to adopt one make-believe perspective during the course of engagement with a fiction but come out being persuaded of a quite different real-world perspective. Rather, the point of disagreement for variantists and invariantists is whether in order for a work to *recommend* a real-world perspective to its audience (attempt to persuade its audience to really adopt a particular perspective), the work must *represent* that perspective to its audience (prescribe that its audience imaginatively adopt a similar perspective).

---

<sup>136</sup>This is not to say that it is impossible for a work to cause us to adopt a moral outlook without representing it to us, nor even that it could be held responsible for doing so. A novel featuring an unintentionally irritating geography teacher might reliably turn people against geography teachers without the work ever representing that outlook. However, in such a case the work hasn't tried to persuade us that geography teachers are bad, it has merely had the accidental effect of giving people that opinion.

A counter-example to persuasion invariantism, then, would be a work that recommended a perspective without representing it.

Liao offers three different explanations of how *Catch-22* could recommend a perspective to its audience as a non-realistic fiction.<sup>137</sup> The explanations are different in terms of *how* they suggest the work prompts us to imagine the moral facts in *Catch-22* to be non-realistic but the details of these differences aren't important for our purposes. We can grant Liao the assumption that there are morally non-realistic fictions, and that *Catch-22* is plausibly among their number. The non-realism of a morally persuasive work is not in itself sufficient to show that invariantism is false, however. In such a case invariantism could still hold provided the work prescribes imagining a realistic moral perspective in *addition* to any non-realistic ones (and that it is this realistic perspective that we are persuaded to adopt). Liao might respond that assuming an additional make-believe perspective here just looks ad hoc. However, without it we are going to have a hard time predicting which perspective is being promoted by which works in a reliable enough way to secure the moral responsibility we want.

One very attractive feature of the invariantist view is that the content of a persuasive make-believe perspective will be very similar to that of the real-world perspective it is trying to persuade us to adopt. It is easy to explain why we ended up with real-world perspective *m* (as opposed to *m*<sub>1</sub> or *m*<sub>2</sub>) if we started off with an imaginary perspective whose contents were identical to *m*. This model makes it very easy to talk about works being responsible for our possessing certain perspectives. Firstly, because there is an obvious connection between the imagined perspective the work is responsible for and the real-world perspective we are claiming the work is responsible for. Secondly, since we are familiar with the idea of imaginative states

---

<sup>137</sup> Liao (2013) pp. 274-7

causing non-imaginative states with a similar attitude and content,<sup>138</sup> the mechanism of persuasion is amenable to explanation by a number of plausible theories of the imagination.<sup>139</sup>

### 6.1.2 *The Symmetry of Satire*

If invariantism is true, then we can say that a work is responsible for the fact that I hold perspective  $m$  but is not responsible for the fact I hold  $m_1$  because  $m$  is similar to an imagining prescribed by the work but  $m_1$  is not. Variantism, on the other hand, has it that an imagining prompted by a work can be responsible for a real-world perspective that has neither contents nor attitude in common with it. So that, to take one of Liao's own examples,

*"When engaging with [Catch-22], we readers are prescribed to imaginatively adopt a make-believe outlook that treats the morally absurd as normal and sensible... [Despite this] it persuades us to really adopt a real-world outlook that questions and challenges the moral absurdities that are associated with real wars, militaries, and bureaucracies."*<sup>140</sup>

The work, on this model, has used an imagined perspective to persuade us to adopt a real-world perspective with no apparent features in common. Variantism needs to explain how we can move from the imaginative perspective that treats the morally absurd as normal, specifically the moral absurdities detailed in *Catch-22* (call this  $m$ ) to the real-world perspective that views militaries and bureaucracies as dangerously absurd (call this  $m_1$ ). In particular, variantism needs to explain how we will move from imagining  $m$  to

---

<sup>138</sup> I have in mind straightforward cases such as: I imagine that I eat fish fingers and then I imagine that I find them disgusting; from this I form the belief that eating fish fingers will disgust me.

<sup>139</sup> Particularly those that emphasise simulation. See, for instance, Goldman (2000) and Currie and Ravenscroft (2002).

<sup>140</sup> Liao (2013) p. 274



adopting  $m_1$  (rather than  $m_2$  or  $m_3$ ) in a reliable enough way to allow us to hold the work responsible for our adopting  $m_1$ .

Liao's proposed solution is that the import/export rules determined by a work's genre make it possible for an audience to be persuaded to hold a real-world perspective  $m_1$  by being prescribed to imagine having make-believe perspective  $m$ . The concept of import/export rules is indeed a helpful one here.<sup>141</sup> The idea of 'import' captures the fact that in engaging with a fiction we, as audience, do not only imagine what is explicitly presented to us in the text. We also often import imaginative counterparts of our real-world beliefs into our collection of imaginings about the fiction. The case we are discussing here is an instance of export; we are taking a moral perspective we were prescribed to adopt in engaging with *Catch-22* and applying it to the real world.

The import/export rules for a work are *symmetrical*: the domain of legitimate imports to a work is the same as the domain of legitimate exports from it. Liao rightly states that it is the symmetry of the relation of similarity that legitimises import and export of the same kind of attitudes in works that are realistic in some form or another.<sup>142</sup> So, for instance, a work that is physically realistic legitimises import of beliefs about the nature of the physical world because if world  $w$  is similar to fictional world  $f$  with respect to physics, then if we have reason to believe a physical law holds in  $w$  we have a reason to believe it holds in  $f$ . Similarly, for export, if a physical law holds in  $f$  we have a reason to believe it holds in  $w$ .

Similarity, however can't serve as the relation governing the import and export of moral perspectives from *Catch-22*. Because *Catch-22* is

---

<sup>141</sup> The terms 'Import' and 'Export' were coined by Tamar Gendler (2000). Walton (1990) includes import rules in his 'principles of generation'.

<sup>142</sup> Liao (2013) p. 280

morally unrealistic according to Liao the perspective represented by the work cannot be exported on the grounds of similarity. Liao's answer to this is to point out that similarity is not the only symmetrical relation.<sup>143</sup> If genres that are realistic in some respects can make use of the symmetry of similarity to legitimise certain import and export rules, then other genres can make use of other symmetrical relations between their fictional worlds and the actual world in order to legitimise different import and export rules. Liao thinks that satires like *Catch-22* have certain import and export rules based on the opposition relation, which is also symmetrical. So that, if in the world of *Catch-22* a moral proposition is true, we have reason to believe that its opposite is true in the actual world, and vice versa. Morally non-realistic fictions then, according to Liao, can reliably persuade us to adopt  $m_I$  by prescribing that we adopt  $m$  because  $m$  and  $m_I$  are opposites.<sup>144</sup>

The problem with appealing to opposition as the symmetrical relation governing the import/export rules of moral perspectives is that it isn't clear that moral perspectives as such are the kinds of things that have opposites. The symmetry of the similarity of perspectives is secured by the fact that for any perspective there will only be relatively few others that can count as similar to it. Opposition is supposed to also secure the same symmetry for the export of perspectives, but it is difficult to say what would even count as two perspectives being opposites. Perspectives are not propositions; the opposite of the *proposition* that militaries and bureaucracies are dangerously absurd might be its negation, but a perspective that includes a disposition to accede to that proposition isn't the kind of thing that can be negated. For instance, the opposite of a perspective that questions militaries and bureaucracies is not just a perspective that *does not* question militaries and bureaucracies. There are countless moral perspectives

---

<sup>143</sup> Ibid. p. 281

<sup>144</sup> Ibid. pp. 282-3

that do not dispose their holder to question militaries and bureaucracies but are nonetheless very distinct from each other in, for instance, their accompanying reasons for not questioning militaries and bureaucracies. Is the opposite of the perspective recommended by *Catch-22* a perspective that does not question militaries and bureaucracies out of fear, out of loyalty, or out of apathy?<sup>145</sup>

Opposition, as I understand it, is a relation that requires its relata to contradict; indeed, this is what makes it an initially attractive option here. If we are recommended to adopt the opposite perspective to that represented in *Catch-22*, then we can easily predict exactly what the adopting perspective will be. This makes it easy for us to praise or blame the work for the audience's adopting of that perspective.

Perspectives, however, seem like they can at most be contrary to one another, and not contradict. If opposition were, counterintuitively, understood as a relation among a potentially huge number of contrary relata, it would be harder to pin responsibility on any work that made use of it in the way Liao proposes. However, I simply cannot see any way to sensibly define what should count as an opposite of a moral perspective, where opposite is understood as implying contradiction.

In order to contend with persuasion invariantism Liao needs to provide a model whereby we can *reliably* get from make-believe perspective  $m$  to real-world perspective  $m_1$  and the opposition relation isn't up to that job. The mechanism needs to be reliable because in order to praise or blame a work for an effect it has on an audience that effect needs to be predictable. Opposition, although it is a symmetrical relation, doesn't in this instance provide that kind of predictability. If it even makes sense to say that perspectives can have opposites, which I doubt, it is certainly false to say that perspectives

---

<sup>145</sup> For similar reasons it also doesn't seem to be the case that we can simply say that the opposite of a particular moral outlook is its absence.

can have few enough opposites that we can fairly say that it is the work that is responsible for its adoption in its audience.

### 6.1.3 Empathic Engagement and Persuasion Invariantism

I have argued against Persuasion Variantism as espoused by Liao, but there remains the question of how Persuasion Invariantism can adequately handle cases like *Catch-22*. Although Persuasion Invariantism was formulated by Liao as a foil with which to show that current theories of how stories can morally persuade don't take adequate account of genre conventions, I will offer some thoughts here suggesting how an empathic model of engaging with stories can make Persuasion Invariantism plausible on its own terms, as well as accommodating the observations about genre that Liao points to.

According to an empathy model of engagement with stories, the narrative perspective encompasses the moral 'outlook' that a story, if it is to be morally persuasive, will both represent and recommend to us. Earlier I offered arguments to show that, in many cases, a perspective over and above that held by any internal narrating character exists in a story.<sup>146</sup> I suggest that *Catch-22* is an especially interesting instance of this phenomenon.

I disagree with the interpretation offered by Liao and Harold of *Catch-22* as a story whose perspective is one which treats the morally absurd as sensible. It strikes me as a far more plausible reading to suggest that in *Catch-22*, and indeed of all morally educative works of satire, the narrative perspective treats the absurdities of the story and characters as just that – absurdities. That is why we find *Catch-22* funny rather than impossible to engage with by virtue of what Matravers identified as coherence failure.<sup>147</sup> We do not try to reconcile the absurdities in *Catch-22*, we merely accept them as absurd and treat the story accordingly.

---

<sup>146</sup> See section 5.1.3.

<sup>147</sup> See section 4.4 above.

*Catch-22*, therefore, does not count as morally unrealistic on Liao's terms, since the narrative perspective is actually in tune with the moral outlook of the majority of its readership. I do not doubt that there can be works that are genuinely morally unrealistic on Liao's terms, but those would consist in stories whose narrative perspective was so far removed from our own that they would likely result in Transmission Failure.

I would like to conclude by noting, however, that the empathic theory of engagement with stories is very well served by Persuasion Invariantism as a way of describing how stories can morally persuade. According to the empathic theory of engagement, we always ultimately adopt the narrative perspective of a story, which in turn determines how we evaluate the moral features of the story we engage with. As Persuasion Invariantism states, a story will have morally persuaded its audience when they adopt a perspective in real life that is similar to the narrative perspective they adopted in engaging with that story.

Furthermore, I want to suggest that the empathic engagement view can incorporate the varieties of moral persuasion found in different genres of story. Although the psychological mechanism of persuasion remains the same for any story, as per Persuasion Invariantism, the empathic engagement view can happily incorporate the idea that different genre conventions are partially manifested in differences in narrative perspective. Indeed, the expectation that one is about to engage with a work in a certain genre may well prime one's empathic engagement in a way that will cohere with the expected perspective of a work in that genre. If I read a fantasy novel, for instance, I will be more ready to adopt a perspective according to which evil is uncomplicated and must be thwarted, whereas if I am reading a political thriller I will be prepared to adopt a perspective of a different kind. The empathic engagement view even allows that such genre expectations play a role in how works in different genres might each

morally persuade in a slightly different fashion. Having primed my empathic faculties for engaging with a particular genre, an author may choose to indulge or frustrate my expectations in a way which teaches me that the narrative perspective I have adopted is equipped to a greater or lesser degree to handle the moral judgements I find myself making. If I find the narrative perspective I associate with fantasy to be unable to comfortably evaluate the story I engage with (as so many find when they encounter *Game of Thrones* for the first time), that will certainly count as a genre specific method of moral persuasion. However, it is a method that is both permitted and facilitated by Persuasion Invariantism.

## 7. VIDEOGAMES, EMPATHY AND THE ROLE OF THE PLAYER

So far I have largely focused on written and spoken stories, and how our understanding of them might be improved by adopting an empathic theory of narrative engagement. While, as I have mentioned, more literary narratives are likely to provide greater opportunity to flex our empathic muscles, the 'literary' label covers only a relatively tiny proportion of the stories we are likely to engage with in the course of our lives. Although it is certainly the case that non-fictional written stories can be 'literary' in the same sense as fictional ones, stories presented in other media can't adopt that mantle quite as easily.

Nowhere is this more apparent than in the world of videogames. Although videogames offer new and intriguing ways of telling stories (new and intriguing to audiences, philosophers and critics), they have so far garnered little attention in academic circles. They are very relevant to my concerns in this thesis, however, since there are good reasons for thinking that, *prima facie*, videogames offer an excellent example of empathic engagement with stories. This is because videogames, to a greater degree than any other medium, often put the audience of the story directly in the position of the protagonist.

Before we can discuss the extent that empathic engagement is involved in playing videogames, however, we need to spend some time discussing what exactly an audience does when they engage with a videogame. As yet, there are no extant theories in analytical aesthetics that offer any analysis of this. I hope to rectify that in what follows, and conclude with some thoughts about what the nature of videogame artworks and the role of the player tells us about the prospects for a special theory of empathic engagement with interactive artworks.

## 7.1 The Extra Credits View

Surf around the innumerable games related internet forums for more than a few minutes and you are more or less guaranteed to come upon a thread on the subject of the artistic value of videogames. Among the more common of the ‘pro’ arguments concerns the supposedly unique role that players have in determining the features of their own experience of a videogame when compared to the audiences for any other kind of work. One of the most eloquent expressions of this idea is to be found in an early episode of the excellent online video series *Extra Credits* entitled *The Role of the Player*:

*“Videogames are unique...Because of their inherently interactive nature we, as developers, ship products that are by necessity incomplete. A painting on a wall is a finished work, a movie on a reel is whole and complete, a novel on a shelf is what it will always be, but a game without a player is nothing.”* (Portnow et al. 2011)

I take this to be a decent exemplar of the common view that part of what makes videogames artistically interesting is the supposed creative role of the player.<sup>148</sup> James Portnow, the writer of the series, is himself a games developer and an influential voice in the industry. With that in mind his view, while worthy of being treated simply as an expression of a commonly held ‘folk’ view, becomes even more deserving of critical scrutiny. Portnow, in the episode described as well as in some of his other writings (Portnow and Grip 2011) uses this analysis of the role of the player to advocate for a theory of design that would place player creativity at the forefront. If Portnow’s view turns out to be false, which I believe it does, designing games on the basis of its truth might turn out to be counterproductive. If artistic merit

---

<sup>148</sup> The view is not just restricted to popular commentators and critics, several theorists of interactive media have suggested similar ideas, though generally not specifically concerning videogames. See, for instance, Rokeby (1996), Ascott (2007) and Levy (1997)



cannot be derived from the creative role a player has in a videogame's production, then designing games that foreground this apparent creativity will probably not help make those videogames better artworks.

We ought not start out being too pessimistic though, and Portnow's view has many attractive qualities that may give us pause. That part of any videogame that we experience (the work's *display*<sup>149</sup>) certainly does not exist without some sort of player input. There is a very real sense, especially in certain genres, that players feel as if we are telling our own story, and indeed the narrative of any particular playing of a game is not, cannot be, fixed until it has been played. My Commander Shepard is female where yours is male; a renegade where yours is a paragon; a soldier where yours is an engineer, and takes a left turn where yours takes a right. It seems true that we experience decidedly different narratives between our two playings of the Mass Effect series, and Mass Effect is hardly atypical in this variability. The only fundamental difference between our games is that I am playing one and you the other, so it also seems true that we are each responsible to some degree for the differences in our experience. That we, as players, are responsible for the presence, absence or form of significant features of the narrative that we experience as the game does seem to make plausible the thought that players are creative partners in their own experience.

Despite Portnow's expertise in the field of game design, however, his view as stated is not quite precise enough to provide an adequate foil for philosophical analysis. The next order of business then, must be to

---

<sup>149</sup> I will use the term 'Display' in the sense coined by Dominic Lopes. A display is that part of a work that we directly apprehend (the performance of the symphony or play, the canvas of the painting, the playthrough of the videogame). Displays in this sense need not be visual. See Lopes (2010) pp. 4-5

clarify the claims made by views such as Portnow's, which I will collectively name 'The Extra Credits View'.

The Extra Credits View, under any interpretation, takes the form of a claim about the source of artistic value in videogames:

**EC:** Videogames are artistically valuable because X

Where the Extra Credits View is in need of some reconstruction is when it comes to deciding what is being claimed to stand for X. I take it that there are two interpretations of what the Extra Credits View claims to stand for X, neatly matching the descriptions of a stronger and a weaker claim respectively:

**X\*:** such works are always, in part, created by their audience.

And,

**X\*\*:** an audience's experience of such works is partially determined by their own actions.

I believe both X\* and X\*\* are plausible interpretations of the Extra Credits View. Although supporters of the Extra Credits View and its ilk do often speak of videogame players as 'creative', or 'artists' in some sense, it is far from clear that those descriptions need to be taken in the strong sense implied by X\*. There are, after all, ways of describing a player of a game as acting 'creatively' without thereby implying that anything is 'created'. Rather, they could just be using some ingenious strategy or taking a novel approach to a problem. Equally, seemingly unequivocal talk of players as 'artists' is often accompanied by descriptions of what such 'artistry' entails as crafting not 'works' but rather 'experiences'. While it makes some sense to colloquially equate the player of a videogame crafting his or her own experience with an artist crafting a work, the strength of that metaphor depends entirely on whether we take the 'experience' crafted to be a work in its own right.

Are both  $X^*$  and  $X^{**}$  suitable substitutions for  $X$  in  $EC$ ? A statement that makes a suitable  $X$  in  $EC$  will be one that is, firstly, plausible in itself, and secondly be such as to plausibly make  $EC$  true. First, let me say that I am taking for granted the notion that if we could establish that videogames, or any other art form, had unique features of any sort that gave rise to distinctive and artistically valuable experiences for its audience, that would show that the medium was of artistic value to the extent that such experiences are artistically valuable. Both  $ECX^*$  and  $ECX^{**}$  rely on the truth of this general principle, and while I do not have the space to defend it here I take it to be a reasonable assumption to make.<sup>150</sup>

To take the weaker claim first,  $X^{**}$ , I do not think that it would make a persuasive case for  $EC$  to have such a claim stand in for  $X$ . The reason is that while it is almost certainly true, it would not make  $EC$  true – or at best, if it did make  $EC$  true, it would make it trivially true. The reason why has to do with the definitions both of videogames and of interactivity. Portnow bases his argument on the premise that videogames are an inherently interactive medium, and this is certainly true. There have been several plausible definitions of videogames offered by philosophers and ludologists alike, and there are few that do not give interactivity as a necessary condition.<sup>151</sup> Interactivity itself has been persuasively defined by Dominic Lopes as *occurring in two forms, strong and weak interactivity respectively*:

*“In weakly interactive media the user’s input determines which structure is accessed or the sequence in which it is accessed...[whereas] strongly interactive artworks are those whose structural properties are partly determined by the interactor’s actions.”* (Lopes 2001 p. 68)

---

<sup>150</sup> For more on the view that the artistic value of a work rests in the value of the experiences it gives us (generally known as ‘aesthetic empiricism’) see Sharpe (2000), Budd (1995), Levinson (1996(b)).

<sup>151</sup> See, for instance, the definitions offered by Tavinor (2009), Lopes (2010), Esposito (2005)

It is clear that of the two, videogames will squarely fall into the category of strongly interactive works: players of videogames do not merely determine which parts of the work are accessed or in what order, they also partly determine significant structural features of the display of the work they engage with. As it happens, the fact that videogames are strongly interactive, or even the distinction between strong and weak interactivity, are not what causes problems for the explanatory value of *ECX\*\**, but rather the fact that videogames must be defined partly in terms of their interactivity. Portnow explicitly states, and I think he is right, that the value that he is suggesting derives from the apparently creative role of players of videogames follows directly from their status as interactive works. This means that any interactive artwork that may not meet the conditions to be properly called a videogame, but which nonetheless prescribe that its audience determines some features of their own experience of the work, will also be able to supply a similar kind of artistic value.

This presents a problem with *X\*\** as an interpretation of the Extra Credits View, since it then becomes possible to rephrase *EC* as:

**EC\*:** Interactive Artworks (including videogames) are artistically interesting because *X*.

Meaning that interpreting the *X* in *EC\** as *X\*\** will result in the following:

**EC\*X\*\*:** Interactive Artworks (including videogames) are artistically interesting because an audience's experience of such works is partially determined by their own actions.

Or, alternatively, by substituting the phrase 'interactive' with something like Lopes' definition of interactivity:

**EC\*X\*\*:** Works that prescribe that their audience's partly determine their experience of the work are artistically interesting because an audience's experience of such works is partially determined by their own actions.

This rephrasing reveals the problem with EC under the weaker interpretation. As mentioned above, either strong or weak interactivity entails that the audience (user) of an interactive work at least partially determine their own experience of it. The resulting claim becomes merely that what makes interactive artworks special is that they are interactive. This is doubtless plausible – if there is anything distinctive about the artistic powers of interactive artworks it will surely derive from their status as interactive objects, but this alone cannot establish that there is in fact anything artistically significant about the inherent interactivity of videogames; we would need to establish that there was some further feature of interactivity that entailed, or permitted, that interactive works possess some distinctive artistic value. Even having shown that videogames, and indeed most interactive artworks, must be termed *strongly* interactive, the weaker reading of EC nonetheless fails to do anything more than restate one of the necessary conditions of strong interactivity. EC then, on the weaker reading, by itself lacks much in the way of explanatory value. The stronger reading, on the other hand, makes a completely different sort of claim.<sup>152</sup> Artistic value, on this reading, is not purported to derive merely from the brute fact of the interactivity of interactive artworks. Rather, it is claimed here that the interactivity of videogames entails a fundamental ontological difference between strongly interactive and non-interactive artworks. Strongly interactive works are, it is claimed, literally incomplete until they are interacted with, and this unique feature allows us to describe players of videogames as, again literally, artists. A medium, a mass medium no less, that so completely redefined the traditional roles of artist(s) and audience(s) would be very likely to claim some distinctive artistic

---

<sup>152</sup> Incidentally, in the course of his talk Portnow makes it clear that he is concerned to present the case for the stronger reading of EC, however since I am using his view only as a foil for philosophical analysis of the general view it represents it would have been uncharitable to neglect the possibility that the weaker reading might be more defensible.

value by virtue of that fact. The next section will investigate the plausibility of this second, stronger reading.

## 7.2 The Role of the Player

The weaker reading of the Extra Credits View, it turns out, is not by itself of sufficient explanatory value to determine why videogames might be artistically valuable. We are left with the stronger interpretation of the Extra Credits View, which I will hereafter refer to as *ECV*, which can be summarised as follows:

**ECV:** Interactive artworks, including videogames, are artistically interesting because such works are always, in part, created by their audience.

As mentioned in the previous section, the stronger reading is certainly the interpretation that Portnow and his compatriots at Extra Credits would support. Even so, it is somewhat harder to see the plausibility of the stronger reading *prima facie* for a number of reasons. For one thing, when we play videogames, or engage with any interactive artwork, it really does not appear to us that we are creating art, so much as consuming it. In fact, our cultural practices around interactive art all seem to bear against the stronger reading of *EC* being the default position: we tend to give videogame players neither credit nor blame for the qualities of the games they play, videogames are bought and sold in a way that implies that the players are buying works to be enjoyed rather than tools with which to create, and gamers tend to very much see themselves as audience rather than artist. Another reason we might balk at taking *ECV* at face value is the counterintuitive consequence that each individual playthrough of a game counts as a work entirely of itself. It does not seem initially plausible that my playing of *Mass Effect 3* and yours are not instances of the same work, despite the obvious differences in the actions,

persona, sex and sexual orientation of our respective protagonists. Are there then any reasons to take *ECV* seriously at all? I believe there are, and possibly the most persuasive comes in the form of an examination of what we know of the respective roles of artists, audiences and videogame players. Are players of videogames performing a role that is more similar to that of the artist to that of an audience member? If we have a reason for believing that they are then we will have at least one good reason for accepting *ECV*.

Traditionally there are two or three roles to be filled in the course of the creation and consumption of a work of art (let us call them *art-roles*). These roles are the artist, the audience and, if the kind of artwork calls for it, the performer. It seems obvious that there is nothing to prevent the same person from fulfilling more than one of these art-roles. The writer of a play can also direct and star in it and a composer can sit in the audience for a performance of her work.

Whether the same person can fulfil multiple roles simultaneously is a more difficult question. We might think that the performer of a piece of improvised jazz, hearing that particular order of notes for the first time as she does, counts as a member of the audience for the music as well as a performer. I am sceptical of this possibility, since listening to a piece of music as a performer and listening as audience seem to require attending to different and distinct aspects of the piece. The performer listens to the notes played in a way informed by her need to decide how to play the next one, an audience may form expectations of what sounds will follow what, but that expectation seems able to be frustrated in a way that the performer's decision is typically not.

Wholly improvised works of music and theatre, it might be argued, provide cases where the same person is simultaneously artist and performer. I would be tempted to argue that in these cases there is no artist in the sense that we would normally use the word, but I do not have space to fully back that up here and the principal thoughts of this chapter do not turn on the truth of it. It would also be rash to dismiss

out of hand the notion that the role of the players of videogames might exactly consist in some mix of those of artist and audience, as this view may well be attractive to many. This scepticism about the simultaneous occupation of multiple art-roles does serve a useful purpose here, however, since it points us in the direction of a framework within which to describe the roles of performers, audiences and artists.

There are some unproblematic cases, as described above, where the same person can fill two or more artistic roles, although perhaps not simultaneously. Nonetheless, this shows that their physical or historical associations with it cannot suffice to define a person's art-role for any particular work. My having created a work does not preclude me from also performing it or engaging with it as audience, nor does my simply being within earshot of a piece of music constitute my fulfilling the role of audience. So what facts about a person do determine their art-role in a given work? I suggest, in line with several other views that it is a combination of their actions, knowledge and intentions.<sup>153</sup> Actions because being an artist, performer or audience each requires the performance of distinct kinds of actions. Artists *create*, performers *play* and audiences '*engage*'. Artists create in the sense that they cause something to exist that was did not exist before, most often by physically altering something in the world (a piece of marble, markings on a page or a canvas), though not always. *Play* in the sense that performers do it is meant as a term embracing the actions typical of stage performance (moving one's body in accordance with the instructions of the work) and musical performance (manipulating musical instruments to produce sounds in accordance with the work). *Engage* is not meant to indicate that an audience is especially attentive, rather just that they are in a position to perceive

---

<sup>153</sup> Iseminger (2004), Lopes (2010), Livingston (2005), Davies (2011)



the artwork, as an artwork, in all the ways typical of the medium in which the work is presented.

Doing the right kind of action is not sufficient to fulfil any particular art-role, the action must also be done in the right kind of way.

Suppose I was to sit at the piano and press random keys that just so happened to result in a note perfect rendition of ‘A Nightingale Sang in Berkeley Square’, and further suppose I had never heard the tune before. This would hardly count as a performance of that song, any more than a performance of ‘My Country ‘Tis of Thee’ would also count as a performance of ‘God Save The Queen’, despite the fact that the tunes are identical.<sup>154</sup> The best explanation for why these tune-makings don’t count as full-fledged performances of their sound-identical works is that the people making the tunes do not have any intention that they should count as performances of that work. So, in order to fulfil the role of a performer I must not only play the melody of ‘A Nightingale Sang in Berkeley Square’, but I must also intend to perform that song by my playing.

There is one further requirement we must make of somebody before we can say that they can fill the role of a performer, that is that they must know what would count as a performance of the relevant work. It is probable that without that knowledge it would be impossible for me to form the intention to perform, but even if the intention can be formed it would be impossible (or so unlikely as to practically amount to impossibility) to successfully carry out that performance. I suggest that these three criteria, the right action, done with the right intention, with the required knowledge to properly form that

---

<sup>154</sup> There are those who may dispute this. Philosophers who are ‘sonicists’ about the ontology of music have the view that musical works just are the sounds that make up a performance of it (Dodd 2010), and would therefore claim that works that sound the same are identical. I find the arguments against this view very persuasive (Davies 2008), but I don’t believe the truth of sonicism would count against the view that the mere making of certain sounds does not constitute a performance.

intention, provide an excellent framework within which to give necessary and sufficient conditions for art-roles.

Recall that my aim in this section is to determine how the role of the player of videogames provides a reason why we might think *ECV* initially plausible. Following our previous discussion, can we determine whether the role of the player is more like that of an artist or an audience member? We have already seen what specific actions, intentions and knowledge are required to successfully perform 'A Nightingale Sang in Berkeley Square': I must *play* the piece, with the intention that my playing constitute a performance of that piece and know what action would constitute a performance of the piece. What of the roles of the artist and audience? In the case of the artist they must *create*, that is perform some action that brings an artwork into being, with the intention that a work should be created by their action, and some knowledge of what the work will be. An audience must *engage* with the work in the sense described above, with the intention that the work be engaged with as an artwork (that is, that it must be attended to as an object of aesthetic appreciation) and no knowledge of the work is necessary, save possibly the knowledge that it is an artwork, or what kind of work it is.

Which of these roles best describes the player of videogames? Could they be classed as performers? The thought might be initially tempting, indeed it has been suggested by Aaron Smuts (2009) that no clear distinction can be drawn between the roles played by the performers of musical works and those of videogame players. If this was true it might be troubling for those, like Portnow, who wish to find some distinctive artistic value in the role carried out by videogame players. This is simply because if it turned out that players of videogames and musical performers are fulfilling the same art-role, then anything artistically valuable to be found in the playing of a videogame will not be distinctive of videogame players; musicians will have been accessing the same value for centuries.

Fortunately for supporters of *ECV*, it seems that players of videogames do not possess the required knowledge or intention to count as performers. A musician's necessary knowledge of the work to be performed, as I have suggested, essentially consists in knowledge of the rules the following of which would constitute a performance of that work. Unlike with musical works, when I play a videogame I do not have to be aware of the rules that constitute a playthrough of it before I play; those rules are encoded in the fictional world of the game in what Grant Tavinor (2009) calls *affordances for fictional action*. Affordances in videogames are presented to the player in the shape of, for instance, doors that can be opened, locations that can be explored, enemies that can be killed and health potions that can be drunk. Affordances, unlike prescriptions for musical performance, do not require any prior knowledge of themselves in order to be fulfilled – if I come across a glowing red button in *Portal*, then I can press it, but I do not need to be aware that my fictionally pressing the button is necessary for me to continue in the game; it may not be, it is just presented to me as an opportunity for the fictional action of pressing it. Because of this difference from musical performance, in terms of prior knowledge of the constitutive rules of the work, affordances in videogames can be acted upon by players-as-audience, without the intention we would require of player-as-performer. To play a videogame, then, is to follow the constitutive rules encoded in affordances for prescribed fictional action programmed as part of the videogame, and generate the videogame's displays by so doing. To perform a musical work, by contrast, is to intend to generate a display of that work by following constitutive rules the details of which you know in advance.

It could be objected here that a player could grow to know a videogame so well that their every move would be deliberate enough that their playing *could* constitute a performance. I think this is a completely reasonable consequence of my views as stated, since I see

no reason why a videogame playthrough couldn't be a performance in the right circumstances; the typical playthrough of a game, however, will not be. It may also be objected that distinguishing the roles of players and performers in this way would leave the way open for players of chess, as well as other games the rules of which need to be known in advance in order to be played, to be considered performers. I have no particular problem with characterising players of chess as performers, and anyone wishing to do so will not find it impacts on my general line of argument, but the idea might seem peculiar to some. To those people I can only say that there may yet be further differences in intention marking the performer of *Summertime* from the player of chess, not least of which might be that the primary considerations of chess players are not aesthetic, but grounded in a desire for a victorious game outcome. My distinction is only meant to support the claim that the interactivity present in videogames is distinctive of videogames and substantially different to how musicians may interact with musical works during live performance. I believe I have shown this, and we are therefore safe to infer that artistic merit gained through such interactivity is the sole preserve of the gamer, and not also of the musical performer.

Players of videogames are not performers then, or at least not typically, but are they more like audiences or more like artists? What are the knowledge, acts and intentions necessary and sufficient for a person to fulfil the art-role of a player? An artist requires significant knowledge of the work she is about to, or in the process of, creating. An audience requires far less, indeed no specific knowledge of the work is necessary in order for an audience to appreciate, say, a novel, or a play as works of art, save that they are meant to be appreciated in that way. Players of videogames are not clearly one or the other. A player need not know the future details of the plot of the game she is playing, but on the other hand it is true that a player may well know the kinds, if not the specific, choices that her character will make;

whether they will be 'good' or 'evil' for instance. A player will know that her avatar is about to turn left, open a door or throw a punch in a way that a designer of the game will not necessarily know.

Furthermore, the knowledge barrier of entry for players of videogames is considerably higher than that which we would usually require of an audience of any other kind of artwork. Even the most cursory playing of a videogame requires the mastering of certain commands, usually involving a degree of motor skill. To play a videogame a player must know *how* the game works, and even when this amount of knowledge is relatively little, it is still considerably more than is required for any other medium. A viewer of Picasso's *Guernica* might well be, in some sense, a better audience if they know something about the history of art, or the history of Spain, but it is still entirely possible to appreciate *Guernica* as art knowing nothing of either of these things. While the player of *Bioshock* need not know anything about Randian objectivism to engage with it as art, though that will certainly make the experience a richer one, a player must at the very least know the appropriate way to manipulate a controller in order to play the game at all.

While the *amount* of knowledge necessary to play a videogame is greater than that needed to engage as an audience with many other kinds of work, it is not clear that the required knowledge is of a sufficiently similar *kind* to that we would normally require of an artist to settle the question of whether players are audiences or artists right now. Players may know, and indeed decide, features of certain characters or the plot, but they do not and cannot know other important aesthetic features of the work that an artist might be expected to; the style of art, the specific dialogue, the twists in the plot etc. On the other hand, since the vast majority of videogames are collaborative works, there is no expectation that any of the individual artists have knowledge of all these things as they are creating the work either. Indeed, if *ECV* is to be plausible at all, it must necessarily be understood as claiming that players are *collaborative* artists, so

provided they knowingly contribute some of the artistically relevant features to their experience, there need be no requirement that they know more than that.

In short, the very nature of interactive media means that players of videogames are going to bring more knowledge to the work than other kinds of audience, but not so much as to unquestionably make them artists. There are two other features of different art-roles we have to investigate that may yet provide a solution. There is some intuitive support to the idea that playing videogames requires a creative intention of some description; when one plays a game one intends to bring into existence something that did not exist before, namely a playing, or playthrough, of that videogame. However, this by itself, nor considered in conjunction with the evidence concerning the knowledge requirement for the role of the player, does not suffice to show that players are artists. Performers have a similar creative intention; they also intend by their actions to bring into existence something that did not exist before: a performance of a work. Now, we already know that players are usually not performers, so it is unlikely that what players of videogames create is a performance. The mere fact that playings of videogames are not performances, however, does not guarantee that they are certainly artworks. Whether the creative intention of the player is of the requisite kind to count as an artistic intention depends entirely on what kind of thing it is that the player intends to create.

A similar concern forces us to withhold judgement on whether the kinds of actions that players perform have more in common with the traditional artistic role or that of the audience. It is perhaps obvious that players do *more* than traditional audiences, but does what they do involve the creation of a work of art? This is where we come to what I believe is at the centre of this debate, the ontological claim at the heart of the Extra Credits View.

### 7.3 The Fridge Magnet Poetry Problem

The Extra Credits View rests on an ontological claim about videogame works, that is, the things we appreciate when we appreciate videogames as art. We have seen this claim manifested in three formulations so far:

- 1) Videogame works do not exist without some contribution from a player
- 2) The art-role of the player requires a creative intention, the end of which is to bring a videogame work into existence
- 3) The art-role of the player requires a creative action which brings a videogame work into existence

The crux of all these claims is that *ECV* takes videogame works to have a particular ontological status. A videogame work, according to *ECV*, is not the thing we buy in our local game store, or download from Steam. It isn't the sum total of the programming code or of the Tavinor-esque affordances for fictional action. A videogame work, according to *ECV* is what happens when a player interacts with those things. Let us create some terminology to discuss this potentially confusing component of *ECV*. Because according to *ECV*, the videogame *work* doesn't exist until after some player interaction has taken place, that thing that we might normally think of as a videogame - a single work with many possible displays - is not a work, but rather a *device for creating new works*. Let us call a device of this sort an *Extra Credits Machine*, and the works created using it *Extra Credits Games*. The truth of *ECV*, then, turns on the question of whether those objects we buy from game stores, download from Steam, or what have you, are videogame works in their own right,

single works with multiple displays, or *Extra Credit Machines*, device for creating many different works with a single display each.

This is an instance of the problem that Lopes calls ‘the ontological challenge of Frigidaire poetry’.<sup>155</sup> The poetry kits commonly found adorning fridge doors consist of magnetic tiles with words or, in some cases, single letters printed on them that can be arranged into lines of verse. The thought is that if we were to consider the ontology of the fridge magnet poetry kit, we surely wouldn’t see each magnetic poetry kit as a ‘work’ with many different displays – one for each assembled poem – rather we would understand the kits as devices for making individual works with one display each. If our intuition here is right, then do we have any good reason for supposing that videogames are different? It seems like we must accept that either a fridge poetry kit is one work with many displays (like a symphony), or that videogames are tools for creating individual works (like an Etch-a-Sketch).

Lopes attempts to answer the point by bringing up an interesting parallel with musical works. George Gershwin’s ‘Summertime’ has been performed by many people, among them Ella Fitzgerald and John Coltrane. Their two performances are certainly not identical, but they are nonetheless instances of the same work – the intuition is that they are different displays of the same work. What makes the song Summertime (and any given videogame) single works with varying displays, according to Lopes, is that when we attend to the different displays of such works our appreciation of them is informed by the fact that they are individual displays among other possible ones derived from the same work. Part of the reason we might consider Coltrane’s interpretation of Summertime a worthy one is that we know of other ways that Summertime can be performed, and the choices made by Coltrane in how he performs the piece become relevant to our appreciation of it. The difference then, Lopes claims, between a

---

<sup>155</sup> (Lopes 2010)



device for creating many works and a single work with many displays is that when we appreciate the displays of the latter, we do so “as one of a range of possible performances.”<sup>156</sup>

The consequence of this claim for the problem at hand is that in the case of fridge poetry we do not appreciate the poems we create as among a range of other possible poems, but rather solely as works in their own right. However, I do not believe that Lopes is correct in that assertion. Our appreciation of fridge magnet poems is surely informed by the fact that the poets have had to make choices more limited than if they were simply working with the English language as a whole. Furthermore, we will also compare works from the same kit not in isolation from one another, but as works from the same kit.

Lopes also suggests, however, that when appreciating a work created by a device such as a poetry kit we do not do so in a way that “implies appreciating the [device] itself” and takes this as a further feature of the difference between devices and works. If this is true, then my above objection may be questioned on the grounds that we do not, in fact, appreciate the poetry kit in the same way as we would the work (as opposed to the performance of) *Summertime*. However, it is far from certain that an aesthetic appreciation of devices such as musical scales and instruments does not go hand in hand with appreciating the works created using them (one might appreciate the dissonance of the Dorian scale, or the ‘woodiness’ of the oboe’s tone as one is listening to the music created using them). Because of this it seems that Lopes is in danger of labelling scales and instruments works with multiple displays rather than devices for creating musical works. It appears therefore that neither of Lopes’ proposed criteria of difference adequately match our intuitions about what should count as a work and what should count as a device for making new works.

---

<sup>156</sup> Lopes (2010) p. 57

Is there a better way to answer the ontological challenge from fridge poetry? Consider the different aims of those who make and use objects like fridge magnet poetry or musical scales and those who make and use objects like videogames or songs. The aim of those who make devices to create works is to provide new ways of creating works in an existing medium. One can still make music without following the requirements of a particular tonal system, but those systems can regulate the making of music and in doing so, make the work produced a distinctive kind of music (e.g. modal music) just as following the rules of the road results in a distinctive kind of driving (legal driving). Similarly, with fridge poetry, the point of it is not to simply allow its users to be poets (we don't need fridge magnets for that), but rather to restrict the syntactic options available to its users, from the whole of the English language to just those letters, words and phrases printed on the magnets, so that creative use of the available magnets will produce a new and distinctive form of poetry. The role of the device for creating new works can therefore be described as imposing regulative rules on a pre-existing activity. John Searle has famously distinguished between rules of this sort – rules that regulate pre-existing activities, and constitutive rules.<sup>157</sup>

Constitutive rules don't regulate pre-existing activities; rather they create the possibility of new activities taking place. To again take chess as an example, the rules of chess do not regulate the pre-existing past time of pushing pieces of wood around a chequered board. The rules of chess constitute the game, so if one is following the rules of chess then one just is playing chess. Just as devices for making new works seem to play a regulative role in creative activity, the things we want to describe as single works with many displays appear to play a constitutive role in such activity. When Gershwin composed *Summertime* he created a series of prescriptions concerning the

---

<sup>157</sup> (Searle 1995)

melody, rhythm etc. the following of which count as a performance of that work – the rules for what counts as a performance of Summertime create the possibility of playing Summertime in exactly the same way as the rules of chess create the possibility of playing chess. Now, when a design team create a videogame like Bioshock they are certainly not regulating a pre-existing activity as in the case of fridge magnet poetry. What they are doing is creating a situation in which it is possible to play Bioshock at all, by supplying the fictional environment in which the game takes place (the Randroid dystopia of Rapture) and prescribing the rules and actions that constitute its playing.<sup>158</sup> Just as the rules of chess create the game by limiting the legal moves of the pieces, so the possible experiences delineated by a videogame's designers are limitations of the possible actions in the fictional world of the game. In Bioshock, to play the game is just to work your way through Rapture, and to work your way through Rapture is to play the game. In Flight Simulator, to play the game is just to pilot the virtual aircraft with which you are presented in the game, and to pilot virtual aircraft presented to you in Flight Simulator is to play the game. In both cases, the rules that govern these fictional activities (concerning physical interaction, choices of aircraft etc.) are the same rules that make it possible to engage in them. If one were to hack the software of Flight Simulator to enable the player to walk from place to place, one would no longer be playing Flight Simulator, just as if one were randomly pushing chess pieces around a board one would no longer be playing chess. Similarly, to play Summertime is to follow Gershwin's prescriptions, and to follow Gershwin's prescriptions is to play Summertime. However, when it comes to devices for making multiple works, the prescriptions are regulative, so while to create a

---

<sup>158</sup> Rules in videogames often take the form of algorithms programmed into the game itself, rather than declarative rules of the sort found in board game rulebooks. This makes the rules of videogames more difficult to break, but they are no less present for that. The rule that my avatar cannot run through walls, for example, partly constitutes Bioshock just as much as the rule that Knights cannot move diagonally partly constitutes chess.

fridge poem is to create a poem, to create a poem is not necessarily to create a fridge poem. In sum, when the rules imposed by a thing regulates an antecedent form of creative action that thing is a device for making new works, but when a thing prescribes rules the following of which are constitutive of a non-antecedent activity, it is a single work with many displays.<sup>159</sup>

With this understanding of the ontology of many-displayed artworks we can comfortably say that videogames belong in that category, and that because videogames have a constitutive function (they create the possibility of new experiences) rather than regulative (they do not regulate pre-existing activities), they do not count as devices for making new (single displayed) works. Note that, if one prefers an appreciation-focused view of the ontology of art such as the one offered by Lopes, and would rather apply his criterion to the problem of whether videogames are *Extra Credits Machines*, he would come up with the same answer. When I play Bioshock a part of my appreciation of the game is precisely the fact that I could have played it a different way; the rich variety of options is precisely what makes some games more fun to play than others, and Lopes is absolutely correct that this kind of appreciation would not be possible if videogames were *Extra Credit Machines* and their works *Extra Credit Games*.

#### 7.4 Videogames as Artworks

*ECV*, on reflection, must be false. The view that videogames are possessed of distinctive artistic value by virtue of the creative role that

---

<sup>159</sup> When deciding whether an object is a single work with many displays or a device for making new works. There may be other objects or concepts that can be distinguished the same way (e.g. Searle's own 'status functions'). Non-repeatable works, works with single displays (such as works of sculpture) do not fit into either of these categories by virtue of their non-repeatability.

their audiences play in their experience of the work, it turns out, depends on an implausible view of the ontology of videogame works. Videogames are not *Extra Credits Machines* because they encode constitutive rules, allowing a new activity to take place. They do not regulate pre-existing activities and are therefore works with many displays and not devices for creating new works. Furthermore, as Lopes points out, a significant part of our appreciation of videogames turns on the fact that our individual playing are only one of a range of possible ones. The story we help to tell by guiding our characters through a game, although it may be the only *source* of our aesthetic engagement with the work, is not *the work itself*.

In closing, I offer a word of encouragement. I do believe that videogame works are artistically valuable, and furthermore I believe that this derives from distinctive artistic capabilities rooted in their inherent interactivity.<sup>160</sup> It is on this premise that the most persuasive cases have been made to support the value of videogames as an artistic medium. More than anything, therefore, I think it is incumbent on those of us who want to see videogames taken seriously as an art form to place their interactivity at the centre of any theories of their artistic value that we may offer. While *ECV* is, obviously, inspired by that interactivity, it has an unwelcome consequence in that *Extra Credits Games*, the objects that are the artworks according to *ECV*, are not interactive. According to *ECV*, a player interacts not with the *Extra Credits Game*, but with the *Extra Credits Machine*; an *Extra Credits Game* has a single display, and has no room for interactivity in its appreciation. This last consideration demonstrates how a poorly considered ontological claim can have far reaching unintended consequences for how we can argue for the artistic merits of different kinds of artworks. It also provides a good reason, I think, for those who might have been attracted to the idea of players-as-artists to feel

---

<sup>160</sup> I have argued for this view in detail in Baker (2012)

that there is certainly a better way to argue for the value of the videogame medium; one that maintains the interactive status of its artworks.

### 7.5 Videogames and Empathic Engagement

Following from my analysis of the role of the player in videogames and its consequences for videogame ontology, we are in a position to describe how my views on empathic engagement with stories might shed light on our experiences with interactive artworks.

The conclusion that players of videogames are not artists is good news for an empathic account of videogaming. As I have suggested earlier, one of the roles of artists in the empathic engagement of stories is to construct the narrative perspective with which the audience engages. This might require some understanding of the ways that empathy with particular perspectives can be finessed so as to achieve the desired result, but it does not require empathy *per se* in principle. If it turned out that videogame players were artists, then, it would look like they were not empathising as a part of their engagement at all, but rather constructing an opportunity for empathy for someone else, or themselves, to engage in later

I want to finish here by just sketching some thoughts on how a theory of empathic player engagement might be developed from all of the preceding arguments and observations. The first, and I think most interesting, thing to say is that I don't think that empathy for the player's character will turn out to be nearly as significant an impact on a player's experience as might first be assumed. For one thing, in many cases the player's character simply acts as an extension of the player into the world of the game; many characters are deliberately empty vessels, not targets for empathy so much as blank canvases that we can think of as we please. There is no room for empathy in such a

relationship, since there is nothing to understand about the other person.

Another reason empathy is likely to be surprisingly rare in video games is because, as I have argued in detail elsewhere, a large part of the distinctive artistic capacities of videogames lie in representing things *about* the player's character *to* the player, by virtue of manipulating the way the player interacts with their avatar.<sup>161</sup> Such a representation might consist in my character suddenly being unable to move as quickly as before; this might represent to me that my character is injured or unhappy, but that is unlikely to encourage me to feel in empathy for that character, I would be more likely to be frustrated. Such interesting representational capacities, however, do have the potential to leverage the narrative perspective to new and engrossing levels of audience engagement. It is even a promising thought that videogames, by leveraging representations of this kind, can make use of a perspective of a different kind, call it a 'game perspective', with which players can empathise. Game perspectives might generate ways of seeing that are distinct from those with which we have become familiar, characterising scenes not in terms of their prominence in a story, but in terms of their relation to the player's efforts to play the game well.

---

<sup>161</sup> *Ibid.*

## 8. CONCLUSION

In this thesis I have offered and defended a conception of empathy, and also offered a theory of narrative engagement based on that theory of empathy. The idea that empathy is connected to narrative is nothing new, but I hope that I have found a way to show that empathy is plausibly central to storytelling practices in a way that has been dismissed by many up until now.

The thesis argued for two headline proposals; the Complex Theory of Empathy and the Empathy Model of Narrative Engagement. While I did not make it explicit during the main body of the thesis, I should point out that I do not think that either one of these proposals relies on the success of the other. If someone were to find one persuasive and the other not, I see no problem in that. Having said that, the form that I suggested Empathic Engagement with Narratives takes is certainly inspired by the Complex Theory of Empathy, and in particular the centrality of being-with imagining to the reliability of the process. Indeed, of the many insights which contributed to the formation of the preceding ideas and arguments, the one I think is the most significant is the thought that Perspectives, or something very like them, play an undeniably powerful role in our lives; in the stories we tell, and in the stories we tell *about* our own lives. If there is any idea that counts as a running thread through all the disparate chapters of this thesis (save perhaps the last), it is this. Given that fact, it would be remiss of me not to mention the great debt I owe to the work of Richard Wollheim and Peter Goldie, who made similar thoughts so central to their own philosophy.<sup>162</sup>

---

<sup>162</sup> Wollheim (1985), Goldie (2012)



If the importance of perspectives to our psychological lives is the single most important theme in the preceding work, there are several other key claims that I would like to emphasise along with it.

The first is that I sincerely believe that much that was interesting, useful and remarkable about the concept of empathy has been lost due to the modern fixation with seeing affective matching as its sole or principal purpose. If there is one point I wish taken from my discussion of empathy, let it be that.

Having said that, there is much to recommend the Complex Theory when set against its competitors. The focus on empathy as a means to understand, to a significant degree, the inner life of another person is, I think, the right way to go. And once you have decided to go that way, although the psychological effort I described as a part of the Complex Theory of Empathy is significant, it is by no means insurmountable. Crucially, I think it is better to describe a faculty like empathy in terms of what it is capable of, and not just what it typically results in. Since we can come to understand other people by the means described in the Complex Theory, it seems trite to me to play down its capabilities just because we often won't avail ourselves of its full potential. The Complex Theory also has the distinct advantage of treating successful empathy as a justified ascription, rather than an affective match, and so we might find that empathy succeeds more often on the Complex View than it might on some other less complex, but equally cognitive views of empathy. The success condition on the Complex View seems to me eminently sensible in any case; providing we can justifiably come to some understanding of another person through our empathy, what reason is there to say we failed?

A further key point that I want to emphasise in closing is the central place of perspectives in stories especially. Even if one does not buy into the empathy model of empathic engagement, I hope I have been persuasive in arguing that we are far better served in thinking about our emotional (and other) responses to stories as continuous parts of a

whole story-length response to the narrative, and not merely as discrete events prompted by individual sentences. We should treat a story as a whole thing, especially for the purposes of analysing our mode of engagement with it, which we should also treat as a whole thing. Incidentally, though, if we are persuaded that perspectives play a key role in determining how our responses to stories develop and change, then an empathy model of narrative engagement, by virtue of its affinity with perspective shifting, will be a natural next step.

In closing, let me leave you with an excerpt from a story I was reading towards the end of writing this. It serves firstly as a nice series of thoughts with which to stop, secondly as a decent summary of a future question that someone interested in the thoughts expressed herein might investigate, and finally, as an excellent part of a great story, whose feelings I enjoyed sharing immensely.

*“I think the act of reading imbues the reader with a sensitivity toward the outside world that people who don't read can sometimes lack. I know it seems like a contradiction in terms; after all reading is such a solitary, internalizing act that it appears to represent a disengagement from day-to-day life. But reading, and particularly the reading of fiction, encourages us to view the world in new and challenging ways...It allows us to inhabit the consciousness of another which is a precursor to empathy, and empathy is, for me, one of the marks of a decent human being.”*

— John Connolly, *The Book of Lost Things*

## REFERENCES

- Adelmann, Pamela K. and Zajonc, R. B** (1989), "Facial Efference and the Experience of Emotion", *Annual Review of Psychology*, 40, pp.249-80
- Ainslie, G. and Montresso, J.R** (2002) "Hyperbolic discounting lets empathy be a motivated process" (Peer commentary in Preston, S and de Waal, F., "Empathy: Its ultimate and proximate bases", *Behavioural and Brain Sciences*, 25, pp.1-72)
- Alward, P.** (2006) "Leave me out of it: *de re*, but not *de se*, imaginative engagement with fiction," *The Journal of Aesthetics and Art Criticism*, Vol. 64, No. 4, pp. 451-459
- Anscombe, E.** (1958) *Intention*, Oxford: Blackwell
- Arpaly, N** (2011), "Open Mindedness as a Moral Virtue", *American Philosophical Quarterly*, 48:1
- Ascott, R** (2007) *Telematic Embrace: Visionary Theories of Art, Technology and Consciousness*, Berkeley, CA: University of California Press
- Baker, A** (2012) "Videogames as Representational Art", *Postgraduate Journal of Aesthetics*, 9:2
- Barron, W.** (2015) "3 Ways Empathy Can Help You Build a Killer Strategy", *Adweek*, 7<sup>th</sup> October, (<http://www.adweek.com/news/advertising-branding/3-ways-empathy-can-help-you-build-killer-strategy-167415>)
- Batson, D.** (2011) *Altruism in Humans*, Oxford: Oxford University Press
- Battaly, H.** (2011) "Is Empathy a Virtue?" in Coplan, A. and Goldie, P. (eds.) *Empathy*, Oxford: Oxford University Press
- Bavelas, J. et al.** (1987), "Motor mimicry as primitive empathy". In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development*, Cambridge: Cambridge University Press.

**Bonomi, A & Zucchi, S**, “A Pragmatic Framework for Truth in Fiction”, *Dialectica*, 57:2

**Booth, W** (1961), *The Rhetoric of Fiction*, Chicago: Chicago University Press

\_\_\_\_\_ (2005), “Resurrection of the Implied Author. Why Bother?” Phelan, J & Rabinowitz, P. (eds.). *A Companion to Narrative Theory*. Oxford: Blackwell, pp. 75-88.

**Brackett, M. et al** (2004) “Emotional Intelligence and its relation to everyday behavior”, *Personality and Individual Differences*, 36, pp.1387-1402

**Brackett, M. and Mayer, J.** (2003), “Convergent, Discriminant, and Incremental Validity of Competing Measures of Emotional Intelligence”, *Personality and Social Psychology Bulletin*, 29, pp1-11

**Brosterman, M.** (2015) “Why Millennials’ Cyber Over-Sharing is Making the World a Better Place for My Children”, *Huffington Post*, 30<sup>th</sup> September, ([http://www.huffingtonpost.com/megan-brosterman/why-millennials-cyber-ove\\_b\\_8201880.html](http://www.huffingtonpost.com/megan-brosterman/why-millennials-cyber-ove_b_8201880.html))

**Budd, M.** (1995) *Values of Art: Pictures, Poetry, Music*, London: Allen Lane

**Burg, B** (2011) “The Beauty of Empathy”, *Bob Burg Blog*, August 31<sup>st</sup>, (<http://www.burg.com/2011/08/the-beauty-of-empathy/>)

**Bury, Liz** (2013) “Reading literary fiction improves empathy, study finds”, *The Guardian*, 8<sup>th</sup> October, <http://www.theguardian.com/books/booksblog/2013/oct/08/literary-fiction-improves-empathy-study>

**Byrne, A** (1993), “Truth in Fiction: The Story Continued”, *Australasian Journal of Philosophy*, 71:1, pp.24 – 35

**Camp, E.** (2004) “The Generality Constraint and Categorical Restrictions”, *Philosophical Quarterly*, 54:215, pp.209-231

\_\_\_\_\_ (2006) "Metaphor and that certain 'je ne sais quoi'",  
*Philosophical Studies*, 129, pp.1-25

**Carroll, Noel** (1990) *The Philosophy of Horror, or Paradoxes of the Heart*, London: Routledge

**Chatman, S** (1990) *Coming to Terms: The Rhetoric of Narrative in Fiction and Film*, Ithaca: Cornell University Press

**Coplan, A** (2009) "Empathic Engagement with Narrative Fictions",  
*Journal of Aesthetics and Art Criticism*, 62:2, pp141-152

\_\_\_\_\_ (2011) "Will the Real Empathy Please Stand Up? A Case for a Narrow Conceptualization", *Southern Journal of Philosophy* 49(s1), 2011(a) pp.40-65

**Currie, G.** (1990) *The Nature of Fiction*, Cambridge: Cambridge University Press

\_\_\_\_\_ (1997) "The Paradox of Caring", in Mettje Hjort and Sue Lavers (ed.) *Emotion and the Arts*, Oxford University Press, pp.63-77.

\_\_\_\_\_ (2010) *Narratives and Narrators: A Philosophy of Stories*, Oxford: Oxford University Press

**Currie, G and Ravenscroft, I** (2002) *Recreative Minds*, Oxford: Clarendon

**Davies, D.** (2009) "Dodd on the 'audibility' of musical works", *British Journal of Aesthetics*, 49:2, pp. 99-108

**Davies, D.** (2011) *Philosophy of the Performing Arts*, Oxford: Blackwell

**Davies, S** (2011) "Infectious Music: music-listener emotional contagion" in Coplan, A and Goldie, P. (eds.) *Empathy: Philosophical and Psychological Perspectives*, Oxford: Oxford University Press

**Davis** (2002) "Too early for a neuropsychology of empathy" (Peer commentary in Preston, S and de Waal, F., "Empathy: Its ultimate and proximate bases", *Behavioural and Brain Sciences*, 25, pp.1-72)

- Davis, M.H.** (1980) "A Multidimensional Approach to Individual Differences in Empathy", *JSAS Catalog of Selected Documents in Psychology*, 10(4): 85
- Davis, C** (1990) "What is Empathy, and Can Empathy be Taught?", *Physical Therapy*, 70, pp.707-11
- Debes, R** (2010) "Which Empathy? Limitations in the mirrored "understanding" of emotion", *Synthese*, 175:2, pp.219-239
- Dewey** (1934), *Art as Experience*, New York: Perigree
- Dimberg, U** (1982) "Facial Reactions to Facial Expressions", *Psychophysiology*, 19:6, pp.643-7
- Dodd, J.** (2010) "Confessions of an unrepentant timbral sonicist", *British Journal of Aesthetics*, 50:1, pp. 33-52
- Eisenberg, N., and J. Strayer**, eds. (1987) *Empathy and Its Development*. Cambridge: Cambridge University Press
- Esposito, N** (2005) "A short and simple definition of what a videogame is", *Proceedings of DIGRA 2005*
- Ficarra, B** (2011) "How Empathy Can Help Empower Patients", 8<sup>th</sup> October 2015, The Huffington Post, ([http://www.huffingtonpost.com/barbara-ficarra/why-empathy-empowers-pati\\_b\\_783138.html](http://www.huffingtonpost.com/barbara-ficarra/why-empathy-empowers-pati_b_783138.html))
- Friend, S** (2012) "Fiction as a genre", *Proceedings of the Aristotelian Society*, 112:2, pp. 179-209
- Gaut, B** (2006) "Art and Cognition" in Kieran, M (ed.) *Contemporary Debates in Aesthetics and the Philosophy of Art*, Blackwell, pp. 115-126
- Gendler, T.** (2000) "The Puzzle of Imaginative Resistance", *Journal of Philosophy*, 97:2, pp.55-81
- Gibson, J.** (2015) "Empathy" in Carroll, N. and Gibson, J. (eds.) *Routledge Companion to Philosophy of Literature*, Oxford: Routledge

**Goldie, P** (1999) "How We Think of Others' Emotions" *Mind and Language*, 14:4, pp. 394-423

\_\_\_\_\_ (2000), *The Emotions: A Philosophical Exploration*, Oxford: Oxford University Press

\_\_\_\_\_ (2002) "Emotions, Feelings and Intentionality", *Phenomenology and the Cognitive Sciences*, 1:3, pp. 235-254

\_\_\_\_\_ (2003) "Narrative, Emotion and Perspective" in Kieran, M and Lopes, D (eds.) *Imagination, Philosophy and the Arts*, Oxford: Routledge

\_\_\_\_\_ (2006) "Wollheim on Emotion and Imagination", *Philosophical Studies*, 127:1, pp.1-17

\_\_\_\_\_ (2012) *The Mess Inside: Narrative, Emotion and The Mind*, Oxford: Oxford University Press

**Goldman, A** (2000) "The Mentalizing Folk" in Sperber, D. (ed.) *Metarepresentation*, Vancouver Series in Cognitive Science, New York: Oxford University Press

\_\_\_\_\_ (2006) *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*, Oxford: Oxford University Press

\_\_\_\_\_ (2011) "Two Routes to Empathy: Insights from Cognitive Neuroscience" in Coplan, A and Goldie, P. (eds.) *Empathy: Philosophical and Psychological Perspectives*, Oxford: Oxford University Press

**Gordon, R.** (2002) "Empathy, Simulation and PAM" (Peer commentary in Preston, S and de Waal, F., "Empathy: Its ultimate and proximate bases", *Behavioural and Brain Sciences*, 25, pp.1-72)

**Hatfield et al.** (1993) "Emotional Contagion" *Proceedings of the American Psychological Society*, 2:3, pp.96-100

**Harold, J** (2000) "Empathy with Fictions", *British Journal of Aesthetics*, 40:3, pp.340-55

\_\_\_\_\_ (2007) "The Ethics of Non-Realist Fiction: Morality's Catch-22", *Philosophia*, 35:2 pp.145-59

**Hatfield, Cacioppo and Rapson**, (1992), "Primitive Emotional Contagion" in Clark, M. (ed.) *Review of Personality and Social Psychology: Emotion and Social Behavior*, Thousand Oaks, CA.: Sage

**Hatfield, E., Rapson, R, and Le, Y.** (2009) "Emotional Contagion and Empathy" in Decety, J. and Ickes, W. (eds.) *The Social Neuroscience of Empathy*, Cambridge, MA: MIT Press

**Hess, U. and Blairy, S.** (2001) "Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy", *International Journal of Psychophysiology*, 40, pp.129-41

**Hills, A** (2009) "Moral Testimony and Moral Epistemology", *Ethics*, 120:1, pp. 94-127

**Hoffman, M.** (1982) "Development of Prosocial Motivation: Empathy and Guilt" in Eisenberg, N. (1982). *The Development of prosocial behavior*. New York: Academic Press

\_\_\_\_\_ (2000) *Empathy and Moral Development: Implications for Caring and Justice*, Cambridge: Cambridge University Press

**Hopkins, R** (2007) "What is Wrong with Moral Testimony?", *Philosophy and Phenomenological Research*, 74:3, pp. 611-634

**Hoque, M. et al.** (2012) "Exploring Temporal Patterns in Classifying Frustrated and Delighted Smiles", *IEEE Transactions on Affective Computing*, 3:3, pp. 323-34

**Hume, D** (1751) *An Enquiry Concerning the Principles of Morals*, in Aiken, H.D. (ed.), *Hume's Moral and Political Philosophy*, New York: Hafner Press, 1948.

**Iacobani, M.** (2011) "Within Each Other: Neural Mechanisms for Empathy in the Primate Brain" In Coplan, A., & In Goldie, P.



(2011). *Empathy: Philosophical and psychological perspectives*. Oxford: Oxford University Press.

**Ickes, W.** (1997), *Empathic Accuracy*, New York: Guilford

**Iseminger, G** (2004) *The Aesthetic Function of Art*, Ithaca: Cornell University Press

**James, R. et al.** (2002) "Empathy: A unitary circuit or a set of dissociable neuro-cognitive systems?" (Peer commentary in Preston, S and de Waal, F., "Empathy: Its ultimate and proximate bases", *Behavioural and Brain Sciences*, 25, pp.1-72)

**Kania, A.** (2005) "Against the Ubiquity of Fictional Narrators", *Journal of Aesthetics and Art Criticism*, 63, pp.47-54

**Kant, I.**, (1911/1790), *Critique of Judgment*, James Creed Meredith (tr.), Oxford: Clarendon Press.

**Kieran, M** (2003) "In Search of a Narrative: Narrative Understanding, Imagination, Sympathy and Empathy" in Kieran, M and Lopes, D. (eds.), *Imagination, Philosophy and the Arts*, London: Routledge

**Kohut Heinz** (1984) *How Does Analysis Cure?*, Chicago: University of Chicago Press

**Kristof, Nicholas** (2015) "How Do We Increase Empathy?" *The New York Times*, 29<sup>th</sup> January,  
([http://www.nytimes.com/2015/01/29/opinion/nicholas-kristof-how-do-we-increase-empathy.html?\\_r=0](http://www.nytimes.com/2015/01/29/opinion/nicholas-kristof-how-do-we-increase-empathy.html?_r=0))

**Lamarque, P and Olsen, S** (1994) *Truth, Fiction and Literature*, Oxford: Clarendon

**Lamm et al.** (2011), "Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain", *NeuroImage*, 54, pp2492-2502

**Liao, S** (2013), "Moral Persuasion and the Diversity of Fictions", *Pacific Philosophical Quarterly*, 94:3, pp. 269-289

**Levenson, R, Ekman, P, and Friesen, W.** (1990) "Voluntary Facial Action Generates Emotion-Specific Autonomic Nervous System Activity", *Psychophysiology*, 27:4, pp. 363-83

**Levinson, J.** (1996(a)) "Film Music and Narrative Agency" in *Contemplating Art: Essays in Aesthetics*, Oxford: Oxford University Press, pp.143-83

\_\_\_\_\_ (1996(b)) *The Pleasures of Aesthetics*, Ithaca: Cornell University Press

**Levy, P** (1997) "The Art of Cyberspace" in Timothy Druckrey ed. *Electronic Culture: Technology and Visual Representation*, New York: Aperture

**Livingston, P** (2005) *Art and Intention: A Philosophical Study*, Oxford: Oxford University Press

**Lopes, D. M.** (2001) "The ontology of interactive art", *Journal of Aesthetic Education*, 35:4, pp. 65-81

**Lopes, D. M** (2010) *A Philosophy of Computer Art*, Oxford: Routledge

**Lotze** (1856) (cited in Depew, D. (2005) "Empathy, Psychology, and Aesthetics: Reflections on a Repair Concept." *Poroi*, 4:1, pp. 99-107)

**Matravers, D** (2003) "Fictional Assent and the (so-called) 'Puzzle of Imaginative Resistance'" in Kieran, M. and Lopes, D. (eds.) *Philosophy, Imagination and the Arts*, Oxford: Routledge.

\_\_\_\_\_ (2014) *Fiction and Narrative*, Oxford: Oxford University Press

**Mayer et al.** (2004), Emotional Intelligence: Theory, Findings, and Implications, *Psychological Inquiry*, 15:3 pp. 197-215

**McBride, Karyl** (2011) "How Empathic Parenting Is the Antithesis of Narcissism", *Psychology Today*, 8<sup>th</sup> October 2015,

(<https://www.psychologytoday.com/blog/the-legacy-distorted-love/201108/how-empathic-parenting-is-the-antithesis-narcissism>)

**Miller, C.** (2011) "Defining Empathy: Thoughts on Coplan's Approach" in *The Southern Journal of Philosophy*, 49(s1)

**Moran, R** (1994) "The Expression of Feeling in Imagination", *The Philosophical Review*, 103:1, pp. 75-106

**Mulgan, G** (2015) "Trotsky, Blair and the New Politics", 12<sup>th</sup> December 2015, *The New Statesman*,  
(<http://www.newstatesman.com/politics/uk/2015/10/trotsky-blair-and-new-politics>)

**Ninan, D** (2010) "De se attitudes: Ascription and Communication", *Philosophy Compass*, 5:7, pp. 551-67

**Parr, L** (2002) "Understanding Other's Emotions: From affective resonance to empathic action" (Peer commentary in Preston, S and de Waal, F., "Empathy: Its ultimate and proximate bases", *Behavioural and Brain Sciences*, 25, pp.1-72)

**Peacocke, C** (1985) "Imagination, Experience and Possibility" in Foster, J. and Robinson, H. (eds.), *Essays on Berkeley: A Tercentennial Celebration*, Oxford: Oxford University Press

**Portnow, J et al.** (2011) "Extra Credits: The Role of the Player" *Extra Credits Youtube Channel*, Retrieved May 2013 (<https://www.youtube.com/watch?v=1XlfeXpiSuQ>)

**Portnow, J and Grip, T** (2011) "Extra Credits Addendum: Discussing the role of the player" *The Escapist*, 24<sup>th</sup> June,  
(<http://www.escapistmagazine.com/articles/view/editorials/misc/8976-Extra-Credits-Addendum-Discussing-the-Role-of-the-Player>)

**Preston, S., and F. de Waal.** (2002). "Empathy: Its Ultimate and Proximate Bases." *Behavioral and Brain Sciences* 25: 1-72

**Preston, S** (2007) "A perception-action model for empathy" In Farrow, T & Woodruff, P. (Eds.), *Empathy in Mental Illness*, Cambridge, MA: Cambridge University Press

- Prinz, J** (2010) "The Moral Emotions" in Goldie, P (ed.) *The Oxford Handbook of Philosophy of Emotion*, Oxford: Oxford University Press
- Quenqua, Douglas** (2013) "Seeing Narcissists Everywhere", *The New York Times*, 5<sup>th</sup> August  
(<http://www.nytimes.com/2013/08/06/science/seeing-narcissists-everywhere.html?pagewanted=all>)
- Reynolds, Molly** (2015), "The 7 Habits of Highly Empathic Women Entrepreneurs", *The Huffington Post*, 24<sup>th</sup> September,  
([http://www.huffingtonpost.com/molly-reynolds/the-7-habits-of-highly-empathetic-women-entrepreneurs\\_b\\_8183938.html](http://www.huffingtonpost.com/molly-reynolds/the-7-habits-of-highly-empathetic-women-entrepreneurs_b_8183938.html))
- Rimmon-Kenan, S** (1983), *Narrative Fiction: Contemporary Poetics*, London: Routledge
- Rochat, P** (2002) "Various kinds of empathy as revealed by the developing child, not the monkey brain" (Peer commentary in Preston, S and de Waal, F., "Empathy: Its ultimate and proximate bases", *Behavioural and Brain Sciences*, 25, pp.1-72)
- Rokeby, D** (1995) "Transforming Mirrors" in Simon Penny, ed. *Critical Issues in Interactive Media*, Albany: SUNY Press
- Schafer, R.** (1959) "Generative empathy in the treatment situation", *The Psychoanalytic Quarterly*, 28, pp. 342-373
- Schechtman, M** (2005), "Experience, Agency and Personal Identity", *Social Philosophy and Policy*, 22:2, pp. 1-24
- Searle, J** (1995) *The Construction of Social Reality*, London: Allen Lane
- Sharpe, R. A.** (2000) "The Empiricist Theory of Artistic Value", *The Journal of Aesthetics and Art Criticism*, 58:4 pp. 321-332
- Singer, T. et al.** (2006) "Empathic neural responses are modulated by the perceived fairness of others", *Nature*, 439, pp. 466-9
- Smith, A.** (1759), *The Theory of Moral Sentiments*, Haakonssen, K. (ed.), Cambridge: Cambridge University Press 2002.

**Smith, J.** (2015) "What is empathy for?", *Synthese*, Special Issue: The Future of Social Cognition (forthcoming).

**Smuts, A** (2009) "What is interactivity?", *Journal of Aesthetic Education*, 49:1, pp. 53-74

**Stein, E.** (1917) *On the problem of Empathy*. Washington, DC: ICS Publications

**Stichter, M.** (2011), "Virtues, Skills, and Right Action", *Ethical Theory and Moral Practice*, 14, pp. 73–86.

**Stock, K** (2005) "Resisting imaginative resistance" *Philosophical Quarterly*, 55:221. pp. 607-624

**Stramondo, J** (2010), "How an Ideology of Pity Is a Social Harm to People with Disabilities", *Social Philosophy Today*, 26, pp.121-134

**Tavinor, G** (2009) *The Art of Videogames*, Oxford: Wiley-Blackwell

**Titchener, E** (1909), *Lectures on the Experimental Psychology of the Thought Processes*, New York: Macmillan

**Tversky, A and Kahneman, D.**, "The Framing of Decisions and the Psychology of Choice", *Science*, 211:4481, pp.453-358

**Velleman, D** (1996), "Self to Self", *Philosophical Review*, 105:1, pp. 39-76

**Vignemont, F** (2009) "Drawing the boundary between low level and high level mindreading", *Philosophical Studies*, 144:3, pp. 457-466

**Vignemont, F. and Singer, T.** (2006) 'The Empathic Brain: How, When and Why?', *Trends in Cognitive Sciences*, 10

**Walton, K** (1990) *Mimesis as Make Believe*, Cambridge, Mass: Harvard University Press

\_\_\_\_\_ (2015) *In Other Shoes*, Oxford: Oxford University Press

**Walton, K and Tanner** (1994) "Morals in Fiction and Fictional Morality", *Proceedings of the Aristotelian Society*, 68, pp. 27-50

**Weatherson, B** (2004) "Morality, fiction and possibility",

*Philosopher's Imprint*, 4:3, pp. 1-27

**Wollheim, R** (1984) *The Thread of Life*, Cambridge: Cambridge

University Press

\_\_\_\_\_ (2000) *On The Emotions*, London: Yale University Press