LIBRARY
Te Tumu Herenga
THE UNIVERSITY OF AUCKLAND

RESEARCHSPACE@AUCKLAND

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand). This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage. http://researchspace.auckland.ac.nz/feedback

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form

Fiction and the Emotions

Grant Tavinor

A thesis submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy
in Philosophy, The University of Auckland, 2003

Abstract

Our participation with works of fiction involves the emotions. Just what this involvement amounts to is a question of great interest. I argue that the literature on the involvement of the emotions in fictive practice has usually made several mistakes. Firstly it has assumed that the emotions are better understood than they actually are. I argue that the theory of emotions common to most of the theories of fiction and the emotion is far from controversial. Revising our conception on what emotions are may allow us to revise our conception of the nature of our emotional involvement in fiction. Second, I argue that two forms of problem are often conflated when philosophers discuss fictional and the emotions. There are semantic and ontological problems, and substantive psychological problems. Often these problems are confused for each other. Distinguishing between the two forms of problem will help to clear up the significant disagreements that exist in the literature. My thesis is broadly naturalised in intent, in the hope that this focus will be productive in explaining the role of the emotions in fictive practice.

Acknowledgements

I would foremost like to thank my supervisor for thesis Associate Professor Stephen Davies. He was always generous in his criticisms, be they of a philosophical, grammatical, or ornithological nature.

Thanks go also to the Department of Philosophy at Auckland University for providing an atmosphere so conducive to philosophical study.

I would also like to the Foundation for Research, Science, and Technology for their provision of a Bright Futures Scholarship for this research.

Finally I would like to thank the family and friends who were supportive of my study, even when I was particularly cranky.

Contents

1. THE PARADOX OF FICTION

1.1 Introduction
1.2 Solutions to the Paradox

2. THE EMOTIONS

2.1 A Theory of the Emotions and the Theory of Fiction
2.2 The History and Theory of the Emotions
2.3 Eliminating “Emotion”
2.4 Approaches to the Emotions
  i.) Affective Emotions
  ii.) Borderline Cases: Sexual Arousal, Startle, Laughter
  iii.) Higher Cognitive and Social Emotions

3. A THEORY OF FICTION

3.1 The Theory of Fiction
3.2 Cognition and Pretence
3.3 Walton and Make-Believe
3.4 Fictive Practice
  i.) Isomorphic Participation
  ii.) Representational Participation
3.5 Fictional and Factual Idioms
3.6 Real and Fictional Actions

4. FICTION AND THE EMOTIONS

4.1 Fictional Emotions
4.2 Thought Theory and Genuine Emotions
4.3 The Usage of Emotions Terms
4.4 The Paradox Resolved
4.5 A Cognitive Theory of Fiction and the Emotions

5. FICTION AND THE AFFECTIVE EMOTIONS

5.1 Affect
5.2 Affect and the Imagination
5.3 Naïve Affective Reactions
5.4 Affect as a Reflexive Prop
5.5 Affect, Fiction, and Behaviour

6. FICTION AND THE COGNITIVE AND SOCIAL EMOTIONS

6.1 Cognitive Emotions and Fiction
6.2 Understanding Fictional Folk
6.3 Who Loves Janet Weiss?

7. FICTION NATURALISED

7.1 Fictions, Emotion, and Behaviour
7.2 The Ubiquity of Fiction

8. CONCLUSIONS: EMOTIONAL PARTICIPATION IN FICTIONS

8.1 Representational Participation
8.2 Isomorphic Participation: Theoretical and Reflexive Attitudes to Fictions
8.3 The Paradox Resolved
8.4 Emotional Participation
8.5 Representational Participation and Emotion
8.6 Isomorphic Participation and Emotion: Theoretical, Affective, and Reflexive Emotional Attitudes
8.7 Conclusion: The Heterogeneity of Emotional Participation in Fiction

1. The Paradox of Fiction

1.1 Introduction

> BLOOM
>
> (*Communes with the night*.) Face reminds me of his poor mother. In the shady wood. The deep white breast. Ferguson, I think I caught. A girl. Some girl. Best thing could happen to him… (*He murmurs*.)… swear that I will always hail, ever conceal, never reveal, any part or parts, art or arts… (*He murmurs*.)… in the rough sands of the sea… a cabletow's length from the shore… where the tide ebbs… and flows…
>
> (*Silent, thoughtful, alert, he stands on guard, his fingers at his lips in the attitude of a secret master. Against the dark wall a figure appears slowly, a fairy boy of eleven, a changeling, kidnapped, dressed in an Eton suit with glass shoes and a little bronze helmet, holding a book in his hand. He reads from right to left inaudibly, smiling, kissing the page.*)
>
> BLOOM
>
> (*Wonderstruck, calls inaudibly.*) Rudy!
>
> RUDY
>
> (*Gazes unseeing into Bloom's eyes and goes on reading, kissing, smiling. He has a delicate mauve face. On his suit he has diamond and ruby buttons. In his free left hand he holds a slim ivory cane with a violet bowknot. A white lambkin peeps out of his waistcoat pocket*.)
>
> (James Joyce, *Ulysses*)

It makes me feel sad, that passage. It is sentimental I know, but the chapter to which it forms the conclusion had been *very unsentimental* up until that point. The chapter is filled with distressing visions of the night; prostitutes, the ghosts and corpses of family members, and drunken, violent soldiers confront Bloom and Stephen as the 16th of June 1904 passes into the following day. When they find themselves outside the brothel under cool streetlight, Bloom has a vision, seeing his deceased infant son, strangely costumed, and reading Hebrew scripture. It is symbolic of course. The passage draws together a great deal of the detail from the previous five hundred or so pages of *Ulysses*, and makes explicit the thematic relationships that stand between Bloom and Dedalus and the various symbolic personas they correspond to. The vision combines regrets Bloom has about his life, what his own son might have been, and explains his attachment to, and care of, the drunken Stephen Dedalus. Dedalus, motherless, and a wanderer, had that day conversed on Shakespeare's *Hamlet* in the National Library. Bloom thinks Stephen, despite the bad teeth, has the potential to become an all-round cultured man. Something he would both want for his son and himself, but for the premature death of his son, and his own poetic banality – he misconstrues a song by Yeats, *Who goes with Fergus*? for Stephen talking about *some girl*. He has been denied on both counts. Bloom settles for enjoying these things vicariously through Stephen's eyes. The son and the father are linked so strongly throughout *Ulysses* they are almost one. Later Bloom invites Stephen to stay at his home, but on Stephen's polite refusal, it is settled the younger man will instruct Bloom's wayward wife. That Bloom's desire to have his vision fulfilled can only fail makes me sad. It *is* sad.

The paragraph you have just finished reading, though perfectly sensible, on close inspection is *strange*, perhaps even *paradoxical*. Bloom and Stephen do not exist; they never did, and never will, as they are the literary creations of the Modernist James Joyce. Yet I write about them as though they do: Dedalus, motherless, and a wanderer, had that day conversed on Shakespeare's *Hamlet* in the National Library; Bloom thinks Stephen, despite the bad teeth, has the potential to become an all-round cultured man. What can I mean? How can I mean it? Obviously there must be more to these sentences than is *literally* implied by their *surface* or *customary* meanings.

It gets worse. Above I write "That Bloom's desire to have his vision fulfilled can only fail makes me sad." *Makes me sad*; how can it? I do not believe either Bloom or Stephen exist, or have ever existed. Emotions are commonly conceived to be comprised of, among other things, beliefs about the existence of their intentional objects. Surely being sad about Bloom and Stephen's situation would *at the least demand that I believed in their existence*? What could I have meant when I wrote the passage then? Also, what could I have been *feeling* that I was inspired to tell you I felt sad? Should we conclude that not just I, but all those who write or speak in such a manner – and that, I think, would be most of us – are inconsistent, or worse, irrational?

And of course the problem is not just with highbrow literature such as *Ulysses*. Indeed it is even more in evidence with the fiction of popular culture. I felt incredibly queasy watching the first thirty minutes of *Saving Private Ryan*, though I was aware the characters I was watching, even though based on actual people, were only fictional. I know people who avoided *Saving Private Ryan* knowing its likely content. Horror movies, pulp romance novels, soap operas, and even videogames, all have the potential to arouse emotions in their appreciators, even though they are quite aware they are participating with fictions.

Such questions are usually considered under the rubric of *The Paradox of Fiction*, or of the *Fictional Emotions*. The subject is rapidly establishing itself a history and an extensive literature. Philosophers such as Colin Radford, Kendall Walton, Susan L. Feagin, Gregory Currie, Noël Carroll, and Peter Lamarque have taken a central part in the debate. There are numerous bit players. The participants are *far* from full agreement as to the nature and status of our seeming emotional interaction with fictions and fictional people, however.

In my thesis I intend to pick and choose from the current literature, highlighting various issues, in an attempt to locate the centre of the discussion. I also introduce characters not usually heard on this stage. In general I apply a *naturalised* focus to these questions. In particular I discuss the work of philosopher Paul Griffiths, emotion theorist Paul Ekman, neurologists Joseph Le Doux and Antonio Damasio, and others. My focus will be on what could be described as a difference between what might be said *logically, semantically, ontologically* about fictive interaction, and what might be said about the *substantive psychological* features of fictive practice. Among my most important conclusions will be that there is an *apparent* logical or ontological divide between fictional characters and their worlds, and us, and that this does seem to throw up the paradox of fiction. I doubt, however, that there is anything particularly emotional about this paradox; it applies to all of our apparent relationships with fictional characters and worlds *qua people and situations*. It is a general class of problem that arises when we attempt to frame ourselves in respect to fictional entities and situations. There is, though, a *specifically emotional* problem – moreover, a set of problems – concerning the emotional appreciation of fictions. Their solutions involve a substantive discussion that inevitably will dip into the growing scientific literature on the emotions.

Obscuring the two classes of questions is the fact that the substantive psychological issue is often confused for the philosophical or logical one, and also, that answers to the logical, semantic, and ontological questions, often end by attempting to answer the psychological and emotional questions. A major symptom of these confusions is the enduring conflict between those who say the emotions that arise from fictive interactions are *fictional emotions*, and those who say they are *real emotions*. This distinction may strike the newcomer to these debates as a bizarre one to make, and indeed, many active participants in the debate still think it bizarre to posit such a thing as a fictional emotion. Yet I will argue that if logical and substantive issues are sufficiently clarified, the postulate is a sensible and necessary one. Moreover, my eventual conclusion will be that the emotions we have for fictions *are both real and fictional*. The conclusion you make depends on what question you think you are answering: a philosophical one, or a psychological one.

Indeed, the two groups of theorists that Robert Yanal classifies as “make-believe” or “pretence theorists,” and “thought theorists” (1999), often have subtly different focuses on the questions of fictive appreciation. Characterised logically*,* and taking into account semantic issues, I believe make-believe theorist Kendall Walton (1978; 1990) is correct when he says that emotions aroused by and relating to fictional characters are *fictional emotions*. Substantively, and here I am agreeing with opinions expressed by, among others, thought theorist Peter Lamarque (1983; 1996), the emotions are *real emotions*. I hope my reasons for concluding this will become clear by the time you get to the end of this thesis.

I will also note here that the substantive characterisation of the emotions will make a positive contribution to my approach to the debate. It is likely the prevalent characterisation of the emotions only adds to the philosophical confusion surrounding these issues. Indeed, I think part of the disagreement inherent in the literature is directly attributable to a difference in how make-believe theorists and thought theorists characterise the emotions. I will attempt to clear up this conflicting usage by giving an extensive treatment of the emotions that makes explicit many things previously only hinted at or assumed.

There is one crucial distinction I need to make at this very early stage. Given that I argue that it is essential we distinguish language used in fictive idioms from that used in factual idioms [3.5], I will attempt to preserve a terminological distinction between the usually synonymous terms “fictive,” and “fictional.” I use the former when referring to states and processes that are involved in *fictive practice*. I use the latter term to identify content that pertains to a *fictional world*, and as such, content that when reported or described belongs behind a *fictive operator*. In this sense, James Joyce’s *Ulysses* is a fictive book. Yet the book Leopold Bloom buys for his wife Molly, *The Sweets of Sin*, is a fictional book. It is only fictional that there is a book called *The Sweets of Sin*. It may be impossible to preserve this distinction throughout this work because I think it is often the case these two senses are equivocated in the philosophical literature. In setting up the issues, and discussing the works of other philosophers, I will often have to fudge the terms. If we look into how the terms are fudged in those discussions, I think it will be clear how confusion is prone to entering this debate.

This thesis will first involve, in this introductory section, a brief account of the paradox of fiction and emotion. I discuss the work of Colin Radford (1975), Kendall Walton (1978; 1990), and Peter Lamarque (1983; 1996), on the problem of our emotional interaction with fictional characters. I take these three writers to be reasonably representative of the state of the debate. They will, anyway, provide an introduction to the debate that is suitable for my purposes. I also make some provisional observations on the importance of fiction and fictive practice as topics of study, because I think that of the topics commonly dealt with in the philosophy of art, fiction and the emotions is one that has particularly wide-reaching implications.

Chapter Two is comprised of a discussion of the emotions and of emotion theory. It is important if I am to explain exactly what the emotional interaction with fiction is comprised of, and what it is not comprised of, that I get a clear idea of what emotions actually are. Griffiths, Ekman, Le Doux, and others from the science and philosophy of the emotions will play an extensive role in this section. Subsequently, my ideas may not be philosophically mainstream; though, with an increasing interest in the philosophy of biology, and naturalised philosophy in general, they are becoming more so. I provide arguments for why I think they are the correct conclusions to hold. One of the things that will come out of this chapter is my commitment to the idea that there is a *kind* or *class* boundary (*or several*) to be drawn within the vernacular class “emotion.” Emotions are a more disparate group of phenomena than traditionally conceived. Most importantly for the topic under consideration, this will mean it is unlikely our emotional interactions with fiction will take a single form. A substantive characterisation of our emotional involvement in fictions might involve a variety of explanations and theories.

Chapter Three deals in more depth with the notion of fictive practice, as I attempt to square away the traditional philosophical problems of fiction, opening the way for my substantive characterisation of the involvement of the emotions in fiction. I will ask what state – or set of states – fictive practice is comprised of, and what it is not compromised of. Of particular interest will be the recent cognitive theory of pretence set out by Shaun Nichols and Stephen Stich (2000). Some of the philosophers who have proposed solutions to the paradox of fiction, Walton, Lamarque, Feagin, and Currie, and others, will also be discussed as I attempt to frame a balanced view of fictive practice. I introduce the distinction between *philosophical* and *substantive psychological* vantage points on the issue, and suggest how the problems each gives rise to can be solved. The philosophical problems are solved by paying particular attention to the pragmatic and idiomatic location of the language we use in our dealings with fictions. Our psychological involvement with fictions is explained by detailing the cognitive structure of pretence.

In Chapter Four, I set out these philosophical and substantive issues with the emotions in mind. I hope to provide a solution to the paradox of the fictional emotions, and a provisional cognitive theory of the role of emotions in fictive practice. I will suggest the problems of our emotional involvement with fictions are species of more general problems.

More detailed solutions to the substantive class of questions will be provided in Chapters Five and Six. Chapter Five deals with the role of “affective emotions” in fiction. Chapter Six deals with the role of the “higher cognitive” or “social” emotions in fiction. These are the two classes of emotion I arrive at in Chapter Two, and in these two chapters my commitment to naturalism in philosophy will be further pressed into service.

In Chapter Seven I make some observations concerning the wider implications of the role emotions play in fictive practice, especially concerning behaviour, and on the role of the emotions in propagating the *cultural ubiquity of fictive practice*. Here evolutionary considerations are made, especially concerning the recent work of *evolutionary psychologists*. I try to give an even-handed, if provisional, appraisal of attempts to characterise the arts within evolutionary terms.

In the final chapter I attempt to draw the strands of the argument together and conclude what emotional fictive participation is, and what it is not. This will necessarily be located within an understanding of what fictive participation *per se* is. The more specifically philosophical questions, and the solutions as provided within Chapters Three and Four, will also play a role here. Emotional fictive participation will turn out to be a diverse set of states and processes.

1.2 Solutions to the Paradox

Colin Radford (1975) queries how it is that we can feel pity for Anna Karenina. He finds it astonishing that the fact we do seem to pity the fictional Anna Karenina is accepted so readily in everyday thought, and that people will so readily assert they *really pity Anna Karenina*. He draws analogies by which he intends to show the peculiarity of such an apparent pity. Radford asks us to suppose we are hearing an account of some tragic set of events. Suppose for example, Bob, our usually credible friend, tells us of his experiences in ‘Nam; how he lost his buddies and his legs to a mortar shell in Da Nang; how he spent six months in a Charlie torture camp, before crawling on his leg stumps fifty kilometres through leech infested jungles to the safety of an American outpost; of his continuing nightmares, drug dependence, and psychological torture; how he wakes in the night screaming and sweating; and how he has never been able to eat Vietnamese take-out since. Being a compassionate person, we think ourselves likely to feel concerned for Bob and his tragic circumstances. We may even shed a silent tear for his horrendous ordeal.

But on noticing Bob’s particularly healthy looking legs, and finding Bob merely borrowed most of his tale from the plot of last night’s episode of *Death Squad Commandos*, our compassionate feelings would likely wane. We might even be angry for being duped by Bob. Radford then turns our attention to fictions such as *Anna Karenina*. In such cases we are *aware from the outset* that the tragic circumstances are made-up, yet emote and express our concerns as we might were the story true. How is it we are able to feel emotional and to express these emotions for fictions such as *Anna Karenina*, and *Death Squad Commandos* for that matter, when we are under no illusion the situations depicted are real? The discrepancy between the cases needs explaining.

Radford canvasses several possible explanations of this behaviour. Going over these proposed solutions here will give some taste of the nature of the paradox, and further, how *not* to solve it. Firstly he asks whether it is simply the case that while observing fictions we *momentarily forget* they are fictions. This cannot be right, because as well as reducing the educated theatregoer’s response to that of a child – it might occasionally be that children *do* forget the fictional status of what they are observing – the explanation can make no sense of why we do not *act* in accordance with our feelings toward fictions. If we simply forget Mercutio is a fiction it should puzzle us why we never feel the impulse to prevent his death, or to notify the police that the killer wore tights and spoke in an incomprehensible English dialect. We do not forget fictions are fictions; *doing so would prevent us from interacting with them as we do*.

Second, Radford considers that we *suspend our disbelief*; the idea originally expressed by Samuel Taylor Coleridge in the philosophical and autobiographical work Biographia Literaria (Engell and Bate, 1983). Radford thinks this little more than a catchphrase or slogan, with little explanatory import. It suggests that while appreciating fictions, we do not direct our attention to their fictional status, but this leaves the paradox unexplained. What it seems is needed is a *positive* source of belief in the substance of the fiction, not merely a temporary ignorance of the fictionality on the part of the appreciator. Also, we need a state that involves the recognition of the fictional status of the fictional content, rather than a temporary ignorance of its status. These two attitudes are quite different, and only the first will explain the experience of fiction, because the latter equates fictive participation with a psychological error.

Third, Radford considers that it is just a *brute fact* about people that they can be concerned for fictional folk. This again is a non-starter. Besides being an arbitrary and uninformative explanation – just what sort of *brute fact* is it? – it leaves unexplained the particular difficulty identified by Radford's original analogy. Why, when the case of being duped by a friend seems so close to that of reading a fiction, do we continue to emote concerning the latter when there we know the fictional status of the material *beforehand*?

The fourth solution Radford offers contends that the fictional case is not so strange. Perhaps there are other cases in which we are rationally and emotionally involved with situations that we are aware have not happened or will never happen. Radford thinks situations like these might arise, but they will not be close enough to the fictional case to warrant our regarding this emotional response as also rational. In a case in which a mother hears of the death of a friend's children, and becomes emotionally concerned about her own children, it is the *possibility* her own children might be hurt in a like manner that makes her emotional, something that is certainly real. The case of Anna Karenina is different: there does not seem to be a manner in which we might be concerned that such a tragic set of circumstances could befall us in the future. It would certainly be unwarranted and irrational to feel occurrent fear for such a small possibility. Moreover, and this will also prove a problem with the following explanations, it seems it is *Anna* we pity not some other, real person. The explanation does not accord with the phenomenal character of our experience, or the way we speak about our fictive experiences.[1]

Fifth, Radford considers the possibility that the fear we feel for Anna Karenina might be focused on others in her situation. Saying "I pity Anna Karenina" is a paraphrase of "I care for people in Anna Karenina's situation." This demands we expect there to be some other person in such a particular situation, and when Lionel Cosgrove straps on the lawnmower and starts shredding zombies in the climactic scenes of Peter Jackson's *Brain Dead*[2], I am just not sure who has ever (or could ever) be in such a situation, that I would fear for their plight. Likewise with Gregor Samsa's strange predicament in Kafka's *Metamorphosis*, Frodo's in *The Lord of the Rings*, or Gonzo's in *Muppets from Space*. Again, it seems plainly to flout the fact that it is *Anna* we pity. At least, this is the way we commonly speak and feel. It also begins to look as if something here really is paradoxical: what is Anna's situation that we might pity others who happen to be in it? Anna does not exist, and is in *no* situation! Such an explanation plainly does not get us very far.

---

[1] Though in itself it is no reason to reject a theory if it does not accord with the *phenomenal character* of our experience. Often this is rather a minor part of theoretical desiderata. If we were to hold to it, atomic theory, and the electromagnetic theory of light, would both be considered more difficult than they presently are. If things really were as they seemed, science would have a much easier time of it. Part of such theories should be an explanation of why things are not as they seem. Indeed, I will eventually (after significant clarification of what is at stake in these matters) offer a theory that in some ways does not accord with the character of our experiences of fiction.

[2] Released in the United States as *Dead Alive* (1992).

The sixth solution opens up the possibility that being moved by real events is a different *kind* of thing to being moved by fictions. Emotions directed at fictional characters and situations may be a *sui generis* form of emotion. While belief in the reality of the object of the emotions is necessary for the former, *genuine* type of being moved, in *fictionally being moved*, no such beliefs are necessary. Radford thinks this gives up on exactly what needs explaining, that is, why the two instances of pity are alike in most respects, yet different in respect to the presence of beliefs. He also finds it counterintuitive that the emotions we have concerning fictions are of a different order to common or garden emotions. Both feel similar, and though we might contend the fictional emotions lack the intensity of real emotions, anecdotal evidence I have encountered suggests that often people think the most deeply felt of their emotional experiences are the *fictive* ones. Also, our sadness over the death of Mercutio and the sadness we might have over a real young man's death seem *superficially* focused on the same things: that the death was such a waste, that it was the end to a life that held much promise. Real emotions and fictional emotions share much of their *dynamic structure*.[3] Like many of our cognitive attitudes and interactions with fictions, such as judgments and inferences, the fictional emotions seem *functionally* or *structurally isomorphic* to their genuine cousins. The fictional emotion is tempered just by the fact that we know Mercutio's life in reality never held any promise. Mercutio is, after all, a fictional character, and fictional characters, *really*, are not the type of things that could be alive.

Radford's enquiry seems at a dead end, though the way he will turn has, if we have read the article carefully, become quite obvious: in being moved by the plight of Anna Karenina, or saddened by the death of Mercutio, we are being *irrational*. Emotions directed toward fictional characters "involve us in inconsistency and so incoherence" (1975:79). To back this claim up Radford gives two other examples of ways in which we are similarly irrational. Tennis players often gesture to the ball, urging it over the net. I have seen (and participated in) something similar in cricket: coaxing of the ball to roll a little quicker to the boundary line is a common cricket practice, even though most cricketers are aware the ball is not listening. This does not seem to be a case of irrationality, but of a simple *game* one might play, in which the utterances made do not imply a commitment to customary illocutionary force.[4] Radford's second example of common human irrationality, the fear of death, is even less convincing. In line with a rather venerable philosophical position, he holds that, since death is like an *endless and dreamless sleep*, we have nothing to fear. This is unpersuasive; because it is not the dreamless sleep most of us fear, but the fact we will not wake up, ever.

---

[3] This will be an important idea when I discuss the cognitive structure of fictive practice [3.2, 3.4, & 4.5].

[4] Though the tennis behaviour, especially when non-verbal, could also be something else that is certainly odd, and yet not I think irrational. I note that appreciators of computer fictions, like tennis players, get up to some exceedingly odd behaviour that can seem irrational [5.5]. Gamers make many actions in playing games (especially in my experience, driving games) that contradict their beliefs about how they can change the fictional world of the game. The actions of tennis players may be similar. If they are, I think such behaviours are simply *non-rational*.

Though he has shown convincingly there is something strange about our emotions concerning fictions, I doubt Radford has provided a satisfactory solution to the quandary. Currie thinks that, in his conclusion that the fictive emotions involve us in incoherent behaviour, Radford "is like one who says that being in two places at once is an incoherent idea, so that anyone who is in two places at once is engaging in incoherent behavior. The problem is posed by an inconsistency between propositions, and we can't solve that problem by talking about incoherent behavior" (1990:187).

I am not sure Radford characterises the problem in an adequate manner. The characterisation and proposed solution seem to be little more than a "Wow! How odd!" They also seem to sell the experience of fiction short. There are many theorists, with what we might call a *humanist* disposition, who would not be particularly happy to see the emotional experience of fictions as a psychological aberration or curiosity. For many of these theorists, the role the emotions play in fictive experience form, or have the capacity to form, a central part of our *moral and emotional education*. It simply will not do that the emotions are irrational or inconsistent.

Walton also comes to this conclusion. His article "Fearing Fictions" (1978) provides the seed of an explanation he fills out in Mimesis as Make-Believe (1990). I will deal with it in more detail later [3.3 & 4.1] but for the moment here is a brief outline of his *pretence* or *make-believe* theory of fictive appreciation. People play games with fictions in a manner similar to the way children play games of make-believe. Walton asks us to imagine how children might play a game of make-believe in which stumps are to be seen as standing proxy for bears (1990:21ff). In doing so, they treat the stumps with the expectations they might hold toward real bears; and by convention they run from the stumps-as-bears, warn friends of their presence, or even bravely attempt to wrestle with them. Fictive appreciation likewise involves a convention-bound participation with such *props*. But here the participation is comprised of psychological attitudes. Appreciators of fiction make interpretations and judgments concerning fictions, and sometimes have emotions concerning them. Fictive audiences *make-believe* in the contents of the fiction the author has provided them with, thereby participating with the contents of the "work-world" of the fiction. Walton also thinks audiences make-believe that they are observing this fictional world, placing themselves in a *relationship* to the content, and generating fictional facts about themselves. Walton calls this the "game-world" of the fictive appreciation (1990:58). Consequently, the emotions resulting from fictive appreciation are "fictional" emotions. It is fictional, not true, that appreciators have emotions for fictional characters and situations (1990:241ff).

As it stands this is ambiguous and could, and has, given rise to confusion, as I will argue later [4.1-4.4]. Walton does not mean the emotions we have toward fictional characters are merely play-acted. Children need not consciously adopt expressions of excitement when encountering a stump; the reaction is natural. Emotions concerning fictional characters are not seen as *contrived* in Walton's theory. Later I will argue this common mischaracterization of Walton's position[5], combined with Walton's sometimes less than clear use of emotion terminology, has been a bar to progress on this issue.

Lamarque (1996), among others, thinks Walton's account fails. Central to his criticism is the contention that the emotions we have toward fictional characters are not fictional emotions, as Walton would have us believe, but *very real emotions*. Lamarque judges that Walton was forced to conclude the emotions were fictional emotions because their apparent *objects* were fictional people or situations. Walton, as a cognitivist concerning the emotions, is committed to the view the emotions are constituted primarily by propositional attitudes such as beliefs. Though I will have a great deal more to say of this and other conceptions of the emotions later [2.1-2.4], here it will suffice to note that an instance of fear for oneself would be comprised of (among other things) a belief on the part of the subject that they were in danger. As Lamarque observes, Walton concludes that because such an attitude cannot be granted to an observer of a fictional blob of slime that is insidiously slithering toward them through a movie screen, the resultant emotions movie viewer Charles might have in such a case are to be seen as fictional. Fictional characters and situations do not exist, and we do not believe them to exist; whatever it is we feel toward fictions, it is not a real emotion.

Nevertheless, Lamarque thinks the fear or pity we often feel when appreciating fictions is genuine fear or pity. He contends the central question we should be asking when approaching this issue is "What are we responding *to* when we fear Othello and pity Desdemona?" (1996:114). Contra to Walton's contention that these responses and psychological attitudes exist within games of make-believe, Lamarque thinks they exist "in the real world" (1996:115). Lamarque's theory of fiction, as most extensively outlined in the early chapters of his Fictional Points of View (1996), and in his work with Stein Haugom Olsen (1994), contends that, from an "external" point of view, fictional characters are referred to as sets of descriptions, rather than as *people*. It is "internally" that reference to characters is reference to people, and then only other internal characters make such a reference.[6] What we respond to when we react psychologically to fictions are the senses of the sets of descriptions comprising externally characterised fictional characters. Lamarque also holds the view that the propositional or predicative contents that constitute the substance of thoughts need not be believed in order for them to be a proper cause of the emotions. "Unasserted thoughts," as both Carroll and Feagin argue, are the proper *causal objects* of our emotions concerning fictions

---

[5] See especially, Carroll, 1990, 73-74; Lamarque, 1996:127; Yanal, 1999:52.

[6] The internal/external distinction arises elsewhere in the literature (Crittenden, 1991; Dilworth, 2002) and will be discussed later [3.5]. I argue that the distinction between *fictive and factual idioms* can better explain the issues the internal/external distinction is aimed at explaining.

(Carroll, 1990, 1998; Feagin, 1996:78-81).[7] While still being committed to a cognitive conception of the emotions, emotions concerning fictions are allowed to take place without the presence of beliefs. The cognitive states that plug the gap are unasserted thoughts, or suchlike.

Lamarque thinks this has four interesting consequences (1996:118). Firstly, he thinks the more imaginative the involvement we have with such a cognitive state, that is, the more that it is held before the mind and the more detail with which it is held, the more psychologically affecting it is likely to be. Second, because unasserted thoughts can arouse emotions, we can be genuinely frightened without being in danger. He allows that there may be good, perhaps evolutionary, reasons for this, a theme I will discuss in later sections. Third, it need not even be likely I could be in a situation for me to be legitimately made afraid *by* the thought of it. Fourth, and most importantly for my concern here, is the fact that emotions elicited by unasserted thoughts or similar cognitive states are genuine and not fictional emotions. This is the conclusion Lamarque offers against Walton’s contention that appreciators of fictions emote fictional emotions.

There is something here that needs resolving. Lamarque seems correct in his observation that, since the fear is directed at or caused by the contents of a thought, it is genuine and not fictional fear. But still, as he admits a little too casually (1996:125) “it can not be literally true that Charles fears the slime, for the reasons Walton gives”; the horror movie slime, obviously, does not exist. Lamarque does not seem to me to have satisfactorily resolved the paradox of fictional emotions. We have reasons both to call the fear genuine – for the fact that it is directed at the contents of thoughts – and fictional – because it simply cannot be literally true that Charles is afraid *of* the slime. In this thesis I will attempt to resolve this difficulty.

There is another paradox of the fictional emotions. It also concerns the content of fictions and our emotional reactions to it. The *paradox of tragedy* has a long history, with one of the most influential treatments dating back to David Hume’s work. This paradox questions how fiction appreciators can enjoy or take pleasure from fictive works when the works in question seem to give rise to emotional states that are customarily thought of as unpleasant. How can we enjoy disagreeable content? The literature on this topic is extensive (Feagin, 1983; Williams, 1998). It is also becoming clear the paradox has some cousins in the form of the *paradox of horror* (Carroll, 1990, 1995; Gaut, 1993, 1993b) the *paradox of tearjerkers*, or even the *paradox of disturbing reality television shows*. The latter paradox is certainly puzzling; why do people enjoy watching amateur video footage of *actual* car and airplane crashes, animal attacks, and crimes? I will not have a great deal to say specifically on the paradox of tragedy or the paradox of horror in this thesis, but I think the theory of fiction and the emotions I develop may have implications for them.

---

[7] The term “unasserted thought,” now widely used, is attributable to Roger Scruton (1974). He introduces it in noting “Clearly there are modes of thought that involve not the assertion of “*p*”, but the more elusive ability simply to hold the proposition that *p* before one’s mind, to entertain *p* as a possibility, or as a supposition. Indeed much of our more complex thought processes – imagination, for one – are of this kind, and we know exactly what it is to say “*p*” unasserted” (1974:88).

***

Besides the mystery of the paradox of fiction and the emotions, there is much else that is of interest in the nature of fictions and fictive practice and the emotions. And it should be of interest to a wide variety of researchers. Scientists, both of the mind and brain, and of the social sciences, should have an interest in fictive practice; fiction should not solely be the concern of the partition of the humanities that concerns itself with philosophy, literature, drama, and film. The reasons for this are manifold. Foremost, fiction and fictive practice are *ubiquitous cultural traits*, observable in present and historical societies, even those unrelated in terms of a transmitted cultural heritage; fictive practice seems to be a *pancultural* trait. Fictive practice, in some form, like language, may even be a form of the "metaculture" that evolutionary psychologists John Tooby and Leda Cosmides think "may be explained as the expression of our universal psychological and physiological architectures in interaction with the recurrent structure of the social or non-social world" (1992:117-118).

This ubiquity makes the research into the nature of fictive practice especially pressing in at least two ways. First, *why* is fictive practice so culturally ubiquitous? What are the universal psychological and physiological architectures the expression of which might explain the prevalence of fiction? The philosopher and the scientist should have an interest in the features of a human practice as common as fiction, for what it will tell us about the mind. In this regard it is similar to other ubiquitous cultural traits such as language. With the scientist, I might ask whether the features we *discover* within our philosophical research into fictive practice, that is, those features we need to posit to make the apparently paradoxical practice philosophically sensible – the affective engagement, the cognitive engagement – are the types of things that might be candidates for features of our universal psychological architecture. At this stage in proceedings, philosophy, with its designs on bridging the divide between the humanities and the biological sciences, sits in an ideal position to make fictive practice accessible to scientific study. Much the same has happened as other disciplines have ascended from philosophical conjecture into hard science. In this connection fiction is important because it opens a door toward further knowledge about our species, both of our transmitted culture and of the universal psychological and physiological traits that ground that culture.

The second reason why the ubiquity of fiction presses the need for its study upon us is quite different. Fiction is an industry, perhaps one of the biggest industries in the modern world. Hollywood itself, the most famous geographical centre of the fiction industry, is likely only the tip of the iceberg, even within the film industry. India's production of films is prodigious, and it seems that, like nuclear weapons, or a space program, the mark of a developed and significant country in the modern world is its ability to support a film industry. The tiny South Pacific country of New Zealand, though unlikely ever to develop a nuclear arsenal or space exploration program, has such filmic pretensions, mostly in the persona of filmmaker Peter Jackson who is spending around 500 million New Zealand dollars of *someone else's money* on a filmic presentation of the fictive works of J. R. R. Tolkien. The production of fiction for television is likewise an enormous industry. Soap operas now dominate the cultural landscape of many societies. Like it or not, the cultural domain of the Western world (and the increasingly westernised non-Western world) is ruled not by the art gallery and symphony hall, but by soaps such as *Coronation Street*, *The Days of our Lives*, *Shortland Street*, *Home and Away*, and innumerable Mexican, Brazilian, Korean, and Spanish soap operas. The production of written fiction also continues to grow in our electronic age. As well as raising the ire of fundamentalist Christians, *Harry Potter* books sell millions of copies to young readers. The fiction of King, Grisham, and Clancy also sells rather well. The electronic age has even spawned a new fictive monster: videogames. Videogames, such as those run on the immensely popular Playstation and Playstation 2, Microsoft X-Box, and Nintendo games consoles, and on personal computers, almost universally involve fictional elements. The games industry is enormous, now rivalling the film and music industries in financial terms. The sophistication of these fictions is also growing, as the technology and fictive media develop. Some videogames now involve complex, film-like interactive fictional adventures. Initial steps have been taken to theorise about the nature of videogames (Herz, 1997; Poole, 2000), though the nature of videogames as fictions has hardly been touched upon.

A further feature establishing the importance of fictive practice is more traditional. Fiction is apparently a potent agent of personal and social change. There have been many claims made for the ability of fictions to educate us, and to improve our moral and emotional lives. There is also a tradition of theory concerning the alleged deleterious effects of fiction on our education, emotional lives, and our societies, starting perhaps with Plato's notorious comments in the *Republic*. A generous sample of the positive claims can be found in Martha C. Nussbaum's books (1986, 1990). Groups who call themselves the moral majority or suchlike often make the negative claims. The *Harry Potter* series of books has been singled out by various Christian groups as involving children with the *dangerous influence of witchcraft*. Though I will not enter directly into this venerable discussion in this thesis, I think that all such claims will be verified only by an accurate assessment of the psychology of fictive practice – or, as in the case of *Harry Potter* books, an accurate assessment of the *dangers of witchcraft!*

***

There is one further thing I would like to note at this stage. For some, the following work might seem to blur the uneasy line between philosophy, and psychology and the biological sciences. And I am aware many are deeply uncomfortable with the blurring of this line. I will make many references to scientific studies, and I will make many attempts to fit the phenomena of fiction and the emotions within the framework of those studies. There will be a lot of *psychologising* of philosophical issues; perhaps even worse, there will be attempts to *physiologise* some of those issues. I think this is inevitable in a project like this. This work is foremost an attempt to *demote a philosophical paradox into a scientific problem*. Such a study will inevitably be a mix of philosophy and science.

Interdisciplinary work opens itself to a variety of charges, foremost that the treatment of the material from the discipline foreign to the researcher, in my case the various mind and biological sciences, is insufficient. I do not believe this to be a problem deriving from the very nature of the combination of philosophy and science in research; I think there are excellent works of this dual nature. Philosophy has a great deal to learn from science, and the opposite is also true. The number of popular (and serious) scientific or philosophical works that commit rudimentary philosophical or scientific blunders that affect the substance of those works, attests to this fact. Any faults that arise in such an interdisciplinary project are due to the researcher's practice in the varying disciplines. I am quite at ease with the fact that my scientific practice in this work (not to mention my philosophical practice!) is likely to be flawed or imprecise. Anyway, *being wrong is not the worst of flaws*; I should think it worse to be charged with indeterminacy, or with proposing an untestable hypothesis. The heuristic structures that have *evolved* or *spread* within the academic world – depending on whether we are prone to Darwinist or epidemiological metaphor – will account for these flaws. But I see no reason to expect a study that trades in a philosophical paradox for a scientific problem will be flawed in its nature; the rise of the natural sciences has had just such a structure.

Such a trend toward empiricalism – especially with a biological flavour –seems to be a problem with some because the natural sciences are striking ever closer to the properties that combine to form our *humanity*. The fatal blows naturalism delivered upon God are fresh in the memory. Several recent philosophical works ponder and worry about the significance of these blows. Daniel Dennett thinks Darwinism a "universal acid" that cannot help but cause us to revise our positions in all branches of the study of our species (1995). I will have no direct concern with these issues.

2. The Emotions

2.1 A Theory of the Emotions and the Theory of Fiction

In this work I spend a considerable time discussing the emotions, perhaps more so than is common in other studies on this topic. Indeed, some of the philosophical works on fiction I consider later make only the most perfunctory of mentions of the emotions, usually signalling their allegiance to an established model. Why then, do I need a theory of the emotions, and why am I paying more attention to this part of the topic than is usual? It is partly because of what I have found in the course of my study into the emotions. For reasons that will become clear throughout this chapter, I think it likely that the usage of the concept “emotion” in a scientific sense, that is, where we are concerned with making *predictive and explanative generalisations*, as I should be in this case, needs considerable revision. The *vernacular* usage of emotion terms, and scientific conceptions that respect those vernacular terms’ generalisations, does not fit with the discoveries of research programs within much of the current science of the emotions. It may turn out the emotions do not form a *natural kind*, and are instead a disparate group of phenomena united only by *superficial* or *arbitrarily defined* features. I will attempt to establish this conclusion in a following section where I draw from recent work within the philosophy of biology.

This is important for the topic under consideration because the couching of the paradox of the fictional emotions has characteristically made reference to a theory of the emotions that respects the generalisations inherent in the vernacular or folk-psychological concept of the emotions, namely, *cognitivism* or *propositional attitude theory*. If this conception of the emotions turns out to be flawed, then not only will my answer to the paradox likely be affected, so too will be the form in which I frame the questions I ask of the phenomena surrounding emotions and fictions. Given a radical shift in the theory of the emotions, the problem may turn out to be quite different from how it is commonly conceived. Philosophers have often relied on a cognitivist or propositional attitude theory of the emotions, and this has played an extensive role in their characterisation of the problem of fictional emotions, and its solutions. The focus in these studies is typically on the missing beliefs and desires central to the propositional attitude account of the emotions; fictional emotions are paradoxical because they lack the beliefs deemed central by simple cognitivist theories of the emotions. Yet, in respecting this common-sense based characterisation of emotions in which they are seen as cognitive representations of classes of situations, philosophers may have assumed the range of occurrences of emotional episodes in fictions is better defined, established, and understood than it is.[8]

---

[8] Indeed, this seems to be reflected in the fact that theories of the fictional emotions often seem predominately *a priori*, seeing little need to establish the existence of the phenomena they wish to characterise. Walton discusses childhood games of make-believe (1990), but his argument cannot

The most cursory of observations suggests a considerable complexity and heterogeneity in our emotional dealings with fictions, not all of which seems immediately captured by a simple propositional attitude theory of the emotions. There are likely arrays of fine discriminations that can be made concerning the emotions we have for fictions that have not always been acknowledged. If we look closer at the *typical* case of an emotion directed at or caused by a fiction, I contend we will find it less than typical. Indeed, the idea that there is a typical emotional response to fiction may be undermined. We often have *feelings* in conjunction with fictions even without ever reporting these feelings; and the opposite is often the case where we report emotional episodes despite a lack of feeling. Some episodes involve us talking about the emotions fictional characters are having or are going to have given the plot developments, and further, about their emotional motivations. Others involve us talking about our emotional *relationships* to the characters. Some of the episodes are ambiguous, in that it is questionable whether they involve emotions as traditionally conceived, though they seem to give rise to many of the same issues: is feeling sick, offended, or startled by a fiction being emotional about it? How about laughter, sexual arousal, anxiety, nausea, frustration, or agitation? All of these seem to raise issues about how we relate to fictional people and situations. Obviously in these cases, what I count as an emotion will play a role in whether I consider it as part of my study; the nature of the emotions become of prime importance. Later I will note evidence for some emotions that might be elicited by fictions that do not fit the propositional attitude stereotype at all, in that they are *non-cognitive* [2.4, 5.3]. Given this variety, it is not particularly clear cut what a theory of fiction and the emotions needs to explain; both because I am not presently clear about exactly what constitutes an emotion, and just when it is we are having an emotional episode involving a fiction.

I should ask clearly at the outset of this study – what do I need to explain about fiction and the emotions? And how do I discover what type of material needs explaining? Given I might revise my conception of the emotions, the traditional form of these questions – such as, *how do we have emotions for fiction where the prerequisites for those emotions, beliefs and desires, do not exist?* – may not be the most accurate or productive way to frame the issue. There are two ways to provide answers to these provisional questions. The first is that I pay a closer attention to the nature of the

---

be mistaken for a set of empirical observations. It is a hypothetical story designed to ease us into his theory of make-believe. Most of the evidence for our emotional involvement in fictions is anecdotal, and the form that seems to make the paradox obvious – appreciators describing themselves as feeling emotions for characters – is certainly a paradigm of anecdote. But we might ask ourselves whether we can always trust anecdote. As I will argue later in this chapter we cannot always believe people when they say they are emotional [2.2]. Of course, one of the options in solving the paradox is to simply deny people have emotions for fictions. A notable exception to this empirically uninterested attitude can be found in Nichols and Stich's recent "cognitive theory of pretense" (2000).

emotions, focusing on the emotion concepts I use to characterise mental states, perhaps revising their *intension* and *extension* along the way. That is the purpose of this chapter. The second comes in a complex form. Having a suitable model of the emotions, perhaps even having rejected the vernacular categorisation altogether, I should re-examine the traditional question, asking how the new model explains the lack of beliefs and desires, thereby explaining the question in its own terms. Yet I also re-examine the episodes in which fiction and the emotions interact, expecting to find my revised conception of the emotions will suggest new lines of inquiry because of the revised intension and extension of the emotion concept(s) as I now understand them. This latter tack might mean I include in my study material previously thought off topic. I think answering this range of questions constitutes a difficult task, and is dealt with in the following chapters.

At this point I would like to propose a preliminary list of candidates for explanation. These are the phenomena, elicited by fictions, or directed toward fictional content, that I think a successful and rounded theory of the fictional emotions should attempt to cover. This collection of phenomena has been shaped by the conclusions of my study, so at this stage it might strike the reader as an odd collection that does not fit with the traditional conception of the paradox of fiction. My arguments for seeing this revised extension of phenomena as the important range of mental, physiological, and behavioural facts are scattered throughout the following text. But perhaps they all share in the view that the commonsense characterisation of both the emotions and of the problem of fiction emotions is not the most productive way to approach this area of study. Among the episodes I think I should be interested in are:

- Feelings (including feelings of fear, elation, anger, agitation, surprise, and sadness).
- Physiological episodes such as racing heartbeats, sweating, sobbing, shivering, and (especially when elicited by visual fictions) startle, disgust, dizziness, nausea.
- Indeterminate feelings such as *stomach butterflies*, *warm fuzzies*, and *hair prickling*.
- Various forms of affective and cognitive arousal, including sexual arousal, concentration, and attention orientation.
- Vocal, facial, bodily, and other emotional *expressions*.
- Laughter.
- Emotional ascriptions and self-reports (e.g. “That [fictional episode] was sad,” “I feel sad”).
- Ascriptions of emotional *attitudes* toward fictional characters and situations, and reports of *relationships* with fictional characters (e.g. “I feel sad for Janet” “I like/love/hate Janet”).
- Prediction and explanation of fictional characters’ emotions, psychological states, actions, and motives (e.g. “Janet feels confused and afraid”).

It is likely there are areas of overlap in this preliminary list, but my observations in what follows are made while keeping these phenomena in mind.

2.2 The History and Theory of the Emotions

The emotions have always been a subject of philosophical interest. This is not surprising considering the central role they play in human life. Philosophers such as Plato, in various works, Spinoza in his *Ethics*, Nietzsche, and Sartre (1939), have been interested in the value of the emotions, and how one should respond to them in life. Much of the current analytic philosophy of the emotions, rather than concerning itself with questions of their value and their place in the moral or aesthetic human life, seeks specifically to pin down their nature. This is foremost an attempt to situate the emotions within the natural world, be it the biological or social world.[9] Much of the study of the emotions now falls within the bounds of the philosophy of mind. Beware; the theory of the emotions is murky territory!

Traditionally philosophers have cut the mind up into three different *chunks*. Firstly, *sensations* are seen as those mental entities that have a *qualitative feel*. We have red sensations, warm sensations, cold sensations, painful sensations, each putatively having a distinct and primitive *phenomenal character* by which it is defined. We normally think mental access to these senses is *primary*: what is a red sensation? Well it is sort of *reddish* really. Philosophy's modern concern with this category of mental phenomena is the *problem of qualia*, which seems to some the most formidable barrier to formulating a naturalistic theory of consciousness (Jackson, 1982), though others see the worry as inconsequential (Dennett, 1988, 1991).

Second among the traditional mental categories are *thoughts*. Western philosophy, especially since Descartes, has often been focused exclusively on thought when it has concerned itself with the mind. That tradition has it that a person is essentially a *thinking thing*. Typically thoughts have been characterised as *propositional* in nature, that is, as having *content* rather than having a phenomenal character as feelings and sensations do. Like declarative sentences in language, thoughts are foremost *about* things, though not necessarily things in the world; thoughts have the curious capacity to be about *non-existent* things, a fact that is of central interest in the philosophy of fiction. This aboutness is often referred to as *intentionality*, or, when it is language that is the concern, *intensionality*, *semanticity*, or *meaning*.

---

[9] The division between the biological and social world, which reflects the popular concern with *nature vs. nurt*ure, has been much discussed recently (Barkow, Tooby, and Cosmides, 1992; Dawkins, 1986; Griffiths, 1997; Sterelny and Griffiths, 1999; Tooby and Cosmides, 1992). It is now widely believed the division is artificial, in that it misrepresents complexities in the *development* or *ontogeny* of an organism. The division will be important to my argument here because the current partition in emotion theories seems to be between *nativists* and *social constructionists*. In as much as the arguments for these positions trade on the nature/nurture distinction – and they do to a considerable extent (Griffiths, 1997) – and in as much as it is an artificial distinction, their characterisations of the emotions will be undermined.

Some philosophers have recently claimed the just noted similarity of thought to language exists because there is literally a “language of thought” (Fodor, 1975, 1987), though one need not hold thought is language-like to hold it has content; one might swap the *language* metaphor in favour of a *mapping* metaphor. Recently in the philosophy of mind an important debate has arisen in which the capacity of thoughts to have propositional content, and also to be *psychologically significant* causal states, has been questioned. Several philosophers hold that the states a future mature psychology deals in will not rely on the generalisations inherent in the current content-based psychology (Churchland, 1981; Dennett, 1987; Stich, 1983). Numerous other philosophers (sometimes numerically identical with those in the first category!) disagree strongly (Fodor, 1987; Sterelny, 1991; Stich, 1996). This is a particularly vexing issue, on which the last word is far from being spoken, and it is impossible for me to say much about it here. But because my thesis will deal with issues of content, intentionality, and mental categorisation – especially concerning *threats* to *folk-categories* – I will occasionally make reference to this debate.

Third among the traditional mental categories are the emotions. It is perhaps the case that of the three categories, philosophy has had the least concern with the emotions; though there have been exceptions in Plato, Spinoza, Nietzsche, and Sartre. If we give credence to how people commonly speak of emotions, then the emotions seem to stand in an ambiguous space between sensations and thoughts. Sometimes emotions are characterised as sensations, as when a person says, “I was so angry it felt like my blood would boil.” “Feeling” is often used as a synonym of “emotion.” Robert Solomon notes the intuition: “a seemingly self-evident truth is that, what ever else it may be, an emotion is first of all a feeling” (Solomon, 2000). Here it is the *phenomenal character* of the experience that is seen as constitutive of the emotion. It is clear enough to most of us that the emotions, particularly, in my experience, fear, anger, and sadness, *do* involve significant and often distinctive feelings.

However, at other times emotions seem more closely allied to thoughts: I might feel guilty *that* I ate the cheesecake I discovered in the refrigerator when I came home from the pub late one night. If my flatmates found out about it I would certainly be embarrassed that I ate all that cheesecake. These examples have a conjunctive structure called a *propositional attitude* in that they involve attitudes directed toward the contents of thoughts, which are often characterised as propositions. In the examples I was both guilty and embarrassed *that* “I ate the cheesecake,” or *that* “I was discovered to have done so.” Characterised in this way, emotions share an obvious similarity with thoughts, in that they are intentional states and have content. Other mental states thought to have such a propositional attitude structure are doubts, desires, and intentions.

Recently, emotion theorists have typically been from one of two camps that correspond to the feeling/thought division noted directly above. Feeling theorists (of which there are a number of species) think sensations are fundamental to the nature of the emotions, and to the differences between the various emotions. Propositional attitude theorists, who are particularly influential at this time, think that thoughts, especially the conjunctive clauses noted above, are more central to the nature of the emotions. They also think it is *classes* of propositional attitudes that allow us to *differentiate* between the emotions. Thus, when approaching the emotions, philosophers, psychologists, and other theorists have typically attempted to *subsume the emotions within a category hoped to be better understood, or more established within the philosophy of mind and psychology; be it thought or sensation.*

Feeling theory is often linked to the James-Lange feedback theory (James, 1884), but the idea that the emotions are essentially internal qualitative states goes further back than that. Yet this conception of the emotions began to find its mature form within the *hydraulic* theory of mind that was prevalent in the Nineteenth Century, and which was shared by Darwin (1872) and William James (1884). The mind on this view is comprised of a trigger system that releases *nervous energy* on the perception of the environmental stimulus, and channels this energy into various actions that respond, adaptively, to the environment. James thinks “that the nervous system of every living thing is but a bundle of predispositions to react in particular ways upon the contact of particular features of the environment … The neural machinery is but a hyphen between determinate arrangements of matter outside the body and determinate impulses to inhibition or discharge within its organs” (James, 1884:190). Behaviour follows perception in a causal and mechanistic manner, with the organism fitting its environment as a lock fits a key (1884:191). Our feelings or emotions on William James’ view are the conscious perception of these bodily responses. “Our natural way of thinking about these standard emotions is that the mental perception of some fact excites the mental affection called the emotion, and that this latter state of mind gives rise to the bodily expression. My thesis on the contrary is that *the bodily changes follow directly the* PERCEPTION *of the exciting fact, and that our feeling of the same changes as they occur* IS *the emotion*” (James, 1884:189-190, emphasis in original). Emotion is consciousness of the autonomic nervous system arousal that accompanies our reactions to perceived environmental stimuli. The feeling theory also aims to differentiate the various emotions according to the classes of situation we regularly face. Fear then, is the conscious perception of the various physiological systems that comprise the *fight/flight* reaction aimed at avoiding danger. Joy might be seen as the perception of the physiology of some sort of reward physiological system aimed at the orientation toward situations of benefit to the organism, or of satisfaction of the organism’s basic drives and desires. On this view, emotion is *just* the feeling or feedback of the physiological behaviours; “emotion” and “feeling” here are synonymous terms.

Darwin thinks feeling is the perception of an *overflow* of nervous energy (1872:74). His theory is more concerned with explaining the *expression* of the emotions in people via facial features and bodily postures derived from traits ancestral to our species and shared as homologues in other species, and which now are sometimes communicative acts. Darwin is less easily construed as a feeling theorist than is usually thought, in that he characterises the emotions as primarily involving physiological and behavioural symptoms. In this he is much closer to various modern theorists who I will discuss later. A modern proponent of a view that is in many ways very similar to the feeling theory is Antonio Damasio (1994, 1999). As I will argue however [2.4], like Darwin, Damasio's theory is rather more sophisticated than an identification of emotion with feeling, and avoids many of the problems feeling theory faces.

The predominant theory of the emotions in the last forty or so years has been cognitivism or propositional attitude theory (Averill, 1980; Greenspan, 1988; Kenny, 1963; Solomon, 1986). Its rise is related to the rise of *cognitive science* in the second half of the Twentieth Century. The story of this rise is complex and is related elsewhere (Griffiths 1997; Le Doux 1998), and will be gone into in more detail later, but several things are important to my argument at this stage. Propositional attitude theory is often couched very much in opposition to the feeling theories of James and the nativist theories of Darwin. There are noted problems with feeling theories. The three most important problems are, firstly, there are a range of *cerebral* emotions that seem to have no discriminable qualitative component. Second, it is a notable problem that empirical findings have shown even animals without the possibility of feedback from bodily regions still exhibit emotional symptoms, and presumably feelings (the experiments typically involve animals, especially dogs, with severed spinal cords).[10] Third, the autonomic nervous system arousal involved in different emotions does not seem to be *all that different*, at least not different enough to allow us, on the basis of the feeling, to differentiate the variety of emotions we attribute to ourselves. I might expect a successful theory of the emotions not only to identify the basic nature of the category "emotion," but also to differentiate between the various emotions. Feeling theory fails on this last count, it is claimed, because the feelings involved in the various emotions are not sufficiently different between the emotions.

It is interesting that in his work James can be seen to be already struggling with a problem that is inherent in traditional emotion theory: the inability to fit all the phenomena covered by the vernacular usage of the concept "emotion" within a single satisfying framework. Having postulated his theory that emotions are essentially the feeling of bodily responses, James faces the problems that some emotions fail to have significant bodily correlates, instead having a primarily "cerebral" origin (1884:201). Propositional attitude theory faces the opposite problem with emotions that lack a suitable cognitive component, one case of which is the very fact that fictions can inspire emotions. Furthermore, James' machinations in avoiding the problem (including the resort to talk of *standard*, and presumably *non-standard* emotions) are strikingly similar to some of the arguments commonly used in defence of propositional attitude theory.

Emotion theorists have subsequently looked elsewhere for the basis of the emotions. An experiment carried out by Stanley Schacter and Jerome E. Singer, which I will deal with in detail later in this chapter, suggests where to look. Schacter and Singer observe that experimental subjects are unable to differentiate between the physiological effects of adrenaline and the physiological effects of normal emotion (Schacter and Singer, 1962). This is the conclusion they draw from the fact that the experimental subjects experiencing the effects of administered adrenaline consistently report emotions when confronted with emotionally relevant stimuli, while control subjects, not aroused by adrenaline, fail to report emotional episodes when confronted by the same emotional stimuli. The feeling of the episode cannot be used to define the nature of emotion because the feelings of emotions and of administered adrenaline arousal are apparently indistinguishable. *A fortiori*, feelings cannot be used to distinguish between different emotions.

---

[10] Damasio thinks this complaint a red-herring, as the severing of the spinal cord would eliminate neither the signals sent from the viscera to the brain through the blood stream nor the nerve signals from the head and face which enter at the level of the brain stem, and such signals are significant (Damasio, 1999:289-291).

Rather than feelings being the important aspect of emotional reactions, propositional attitude theorists claim it is *thought* or *cognitive appraisal* of the type attributable to Schacter and Singer's experimental subjects, that is most important to identifying emotional states. Emotion is stereotyped as a cognitive appraisal of a situation, *perhaps* involving physiological arousal and phenomenal content or feeling. The cognitive states invoked are typically beliefs and desires. Fear then, is not a feeling, but is the apprehension or belief that *you are in danger* (and perhaps also the corresponding desire that *you not be in danger)*; surprise might be characterised as the apprehension of salient novelty in your environment; sadness as the appraisal of loss, or immanent loss, and so forth. The two features we expect of a successful framework for explaining the emotions are met by a cognitive stereotype. Fundamentally, emotions are cognitive appraisals of features of the world; and *differentiating* between the various emotions are the classes of features, such as danger, novelty, and loss, to which the cognitive appraisals typically affix.

For the cognitivist the emotions are primarily states involving *informational transformation*, and are thereby apt to be included in a cognitive theory of the mind. Emotional states are thus also apt to be construed in *functional* terms, or at a *task description level*, where the nature of the implementation of the functional role is essentially irrelevant to the identification of the mental state. Furthermore, inherent in the functional roles propositional attitude theory assigns to emotions – such as appraisals of danger, novelty, and loss – are a respect for the categories and generalisations inherent in *folk-psychology*. Propositional attitude theory maps fairly neatly onto the generalisations inherent in the vernacular usage of emotion terms.

However, propositional attitude theory shares with James' feeling theory the difficulty in subsuming all of the states referred to by the vernacular concept "emotion" under one successful framework. Feeling theory is unable to account for the cerebral emotions and other cases emotion where bodily feedback is non-existent, or impossible due to experimental manipulation, pathology, or injury. Propositional attitude theory has problems in cases where there is a lack of a suitable cognitive component, specifically a lack of the presence of beliefs and desires as traditionally conceived. The evidence for such non-cognitive emotion states is widespread and compelling, and though propositional attitude theory might wriggle, it is likely it will not escape the fact that some emotions lack a suitable cognitive component. I will now spend considerable time and effort discussing these worries because they bear directly on the topic of this thesis, because of their importance in the future of psychology and philosophy, and because they are very interesting.

In The Emotional Brain (1998), Joseph Le Doux plainly says some things many people, especially propositional attitude theorists, would find contentious: emotions are “not merely a collection of thoughts about situations” (1998:71); they are instead best to be identified with things (and even then, a diverse set of things) *in the brain!* Emotions are not stereotyped cognitive appraisals, but are the activations of particular neurophysiological mechanisms and structures, or in more general terms, emotions are “biological functions of the nervous system” (Le Doux, 1998:12). In Le Doux’s work, the functionalism that led to the cognitive revolution in psychology goes out the window to be replaced by talk of emotions as physiological episodes in the bodies and brains of actual humans and their close *phyletic* relations. This is a replacement of the task description level familiar from cognitive science, with a characterisation at the level of *implementation*. While this move is strictly consistent with the continuation of the use of explanations at the level of task description, what Le Doux thinks a concern with the level of implementation *will show* is that the predominant analysis of emotion at the abstract level, that is, cognitivism or propositional attitude theory, is fundamentally flawed. He also thinks discussing the emotions at the level of implementation will allow us to discover which of the many possible implementations that might satisfy the task descriptions provided by cognitive science, evolution has provided us with.

Before he begins to characterise the anatomical structures involved in the emotions, Le Doux offers a concerted criticism of cognitivism’s methods, presumptions, and conclusions. Much of his criticism stems from empirical findings; Le Doux is himself an experimental neuroscientist. He makes it clear that two senses of cognitive action are to be made distinct, and while it is a *truism* to hold emotions are cognitive states in one of the senses, to hold that emotions are cognitive states with the other sense in mind, is far from uncontroversial (1998:68). First is the *narrow* sense of the manipulation of mental representations or content such as beliefs and other propositional attitudes, perhaps even in the form of the *quasi-logical* manipulation of belief-like sentential states that I noted is suggested by various philosophers of the mind (Fodor, 1975; 1987). The second *wide* sense of cognitive action is broad and includes the brain’s processing of any form of information, however *reaction-like* the transition between informational states is, or however loosely the term “information” is used. It is the former Le Doux is denying as constitutive of emotion, something he thinks is shown by the arguments and experimental results he presents and that I will shortly look at. Indeed, to deny that the latter is involved in emotion is, as Paul Griffiths notes, to attribute the work of the emotions to some sort of “magic” (1997:26).

But in fact the former narrow sense of cognition *is* held by various theorists to be the substance of the emotions. Propositional attitude theorists see emotions as consisting partly (perhaps predominately, and certainly necessarily) of propositional content such as beliefs concerning situations (Averill, 1980; Greenspan, 1988; Kenny, 1963; Solomon, 1986). To be afraid is partly but necessarily to believe I am in danger (and perhaps to desire I not be in danger). Le

Doux is denying this is the case; it is not necessary that I believe I am in danger in order to experience fear, and the fear certainly does not *consist* of the cognitive appraisal. “Something else is needed to turn cognitive appraisals into emotions, to turn experiences into emotional experiences. That something, of course, [in instances of fear] is the activation of the system built by evolution to deal with dangers” (Le Doux 1998:284).[11]

In fact I might be fearful without even knowing what I was fearful of (Le Doux 1998:298). Le Doux thinks various studies on the effect of “subliminal perception” (1998:56-63) or “mere exposure” (1998:53-55) on preferences and on emotional conditioning show that affective emotional processing is unconscious. In the various studies Le Doux describes, subjects without the possibility of conscious access or recall to various cues presented to them nevertheless exhibit the ability to become emotionally affected by the exposure to the cues. It is not merely that the stimuli are unconsciously processed, but that they are non-cognitively (narrow sense) processed.

In phobias, I might be afraid in ways that *contradict* the beliefs I have, or those I would attest to. Phobias are a notable problem for cognitive theories to account for because they cannot be easily subsumed under the cognitivist framework. They might be reconciled with cognitivism, perhaps expeditiously, by being labelled *irrational* or *atypical* emotions. This move is essentially an attempt to theoretically capture an alleged emotional type that does not correspond to the functional role cognitivists see as constitutive of their emotional typology. Many people are afraid of things they obviously do not believe are dangerous. I am convinced flying is an extremely safe way to travel. This does not negate my hanging around in the airport bar before getting on most planes (after which I can honestly say I *love* flying). Many people are terrified by thunder and lightning, and it seems a stretch to conclude that just because the fear is not accompanied by a suitable cognitive component, such as a belief that one is likely to be struck by lightning, that the fear is atypical; in children, the fear can be quite typical. Phobias are not *irrational. As* Le Doux might claim, they merely reflect the *non-rational* nature of much affect. Imagine for a moment I am a sufferer of *pukekophobia*: an irrational fear of pukekos.[12] The phobia will not be based on any irrational thoughts I may have about pukekos, but the fact that my affective systems have been somehow

---

[11] Le Doux’s claim that cognitions without *something extra* would be *cold* (1998), is an echo of James’ claim over a hundred years before that grief without physiological feedback would be a “feelingless cognition that certain circumstances are deplorable, and nothing more” (James, 1884:194).

[12] Pukekos are black and blue-feathered birds of the *rallidae* family, now often seen living on roadsides in rural New Zealand. Despite certain opinions to the contrary, pukekos can fly, though they look especially un-airworthy when they do so. I know of no particular rational reason to be afraid of pukekos.

conditioned[13] or *primed* into eliciting fear reactions in the presence of pukekos. My phobia consists of my having a reaction of fear on inappropriate occasions.[14]

Care should be taken not to turn what might be the exception into the rule, lest some think I am placing too much emphasis on the anomalous results of these studies. That one might turn the exception into the rule seems a particular threat to the brain sciences, because in the main their method, especially when concerned with our human brains, must by necessity deal with the brains of the pathological or injured, because actual experimental manipulation is ruled out for obvious ethical reasons. In this case, I think the weight of these studies is increased because of their situation within a wider theoretical view of affect that embraces naturalism and a *bottom-up* methodology. The prevailing view in neurophysiology and neuropsychology, such as that practiced by Le Doux, Damasio, and many others, is that affect is unconscious and even sometimes non-cognitive. The modern neurophysiology of the emotions has a strong underpinning in evolutionary considerations, drawing perhaps from Paul MacLean's influential view that affect is based on phyletically ancient neural structures (MacLean, 1980, 1990). The idea that affect is unconsciously implemented fits comfortably within this view, given what many scientists think about phyletically ancient cognition.

Le Doux argues that cognitive appraisals or self-attributions of emotion states often arise *after the fact*, that is, once the emotional significance has already been established by the non-cognitive and unconscious mental mechanisms that are the object of his experimental observations (Le Doux, 1998). Cognitive states do play a role in *eliciting* or *inducing* affective emotional responses. But such cognitive content identified by cognitivists as the core of the emotions often arises as the physiological perturbation of an emotional episode enters our *cognitive workspace* – that is, our extended consciousness (Damasio, 1999:195-233), where it is that we are aware *we* are experiencing something – and is assimilated with the conscious evidence of our present situation. We often note the affective disturbance and ascribe an emotional state to ourselves. Cognitive scientists call this the *attribution* in an emotional episode. I think it is involved *confabulation* in the Schacter and Singer experiments. The notion of an *emotional attribution* or *ascription* will be important in what follows.

The Schacter and Singer experiments are important in making clear the distinction between what Le Doux sees as constitutive of emotion and what cognitivists have typically seen as central. Schacter and Singer conclude that the important step in the substance of emotion is the cognitive appraisal of the sort available to self-report; it is cognitive appraisals that distinguish the emotion states from each other. This fits well with the criticisms of James' feedback theory identified earlier; that the emotions are not sufficiently differentiated by their qualitative component. What is important is not a feeling, which, anyway, is ambiguous between the various emotions, but a cognitive appraisal of an environmental feature such as the danger or novelty of the situation. Such a feature provides a suitable candidate for a *functional stereotype* of the various emotions; the appraisal of danger provides the cognitive stereotype of fear; the appraisal of novelty provides the cognitive stereotype of surprise. This conclusion was of great influence on the subsequent propositional attitude school of emotion theorists. It identifies as the core of the emotions the narrow sense of cognitive action or information processing I discussed earlier: a propositional attitude. This theory is very much like the cognitivist characterisation of beliefs, in that it provides a *satisfaction condition* for the identity of beliefs. Construed in this way, emotions are in part intentional states.

---

[13] I am aware of the, perhaps ugly, *behaviourist* connotations this term has. Its use here signals no commitment to behaviourist theory.

[14] Philosopher of the arts Susan Feagin makes what I think is a bit of a blunder on this matter. She argues that phobias (1996:212) are constituted by thoughts you cannot stop yourself having. But this gets things around the wrong way, and reflects a bias toward cognitivist theories of the emotions that exists in her work.

I think Le Doux helps us to see that Schacter and Singer, and those subsequent to them, draw the wrong conclusion from the experimental results. Rather than finding physiology irrelevant, and cognitive identification as the *key* to the emotional episode, all the experiments seem to show is that people are prone to attribute or ascribe emotional states to themselves to explain physiological disturbances such as the injected adrenaline. Further, that in doing this they will take cues from the surrounding environment. We need not expect this last point to be exhibited only in situations of experimental manipulation; occasionally the morning after a late night a person might think himself or herself a little depressed, when in fact they are just suffering from a mild hangover. Coffee has the opposite effect. But to think hangovers or coffee binges somehow alter the content of a person's propositional attitudes is to get things mixed up. To conclude that the cognitive appraisal is the basis of emotion seems unwarranted, as Schacter and Singer's subjects *did not have emotions*. Rather, drawing from the evidence of their environment, they *confabulated*, and posited the existence of emotion states to explain the internal affective disruptions that presented themselves as feelings. Furthermore, Le Doux concludes these results show we are often on a very unstable cognitive footing with our emotional episodes (1998:298-299).

The experiments *do* show feelings are ambiguous, but this does not discount the role physiology plays in providing the basis of emotion, and in discriminating between the emotions; it only counts against a very rudimentary feeling theory, which it is doubtful even Darwin held out hope for. Even given we are on an unstable footing when identifying our emotions, this says nothing against the claim that emotions are *discrete physiological states*. Establishing that we often confuse a set of twins, Bob and Carl, does nothing to undermine the claim that Bob and Carl are different people after all. Schacter and Singer's experiment does not establish that there is not a qualitative component to some emotions that can be used to identify these emotions. The experiment establishes only that, where there is an ambiguous feeling, a confabulated attribution might occur. It does seem likely joy and fear can in *practice* be differentiated by their subjective feelings, though this is not what in principle individuates them into *kinds*. Pain is not so close to pleasure that we cannot in practice distinguish the two states. Furthermore, sincerely reporting an emotion state is not sufficient for having an emotion state. This *confabulation* in such cases of cognitive constraint is confirmed by various sources, such as the familiar “split-brain” studies (Gazzaniga and Le Doux, 1978), and studies of confabulation in self-reports (Nisbett and Wilson, 1977).

A major factor in this discrepancy is that the type of *information* used by subjects to appraise their emotional states is often not of the form available to the emotion systems themselves, and vice versa; the information that induces affective states may not always be of a kind available to higher cognition and to consciousness (Le Doux, 1998:53-64). What experiments have turned up are some surprising facts about the availability and transfer of information within the mind/brain. The *conscious part*[15] or highly cognitive part of the mind, at least that part that is *available* in self-report and introspective observation, does not always *know* what the rest of the mind is up to. Indeed, the *rest of the mind* might *know more* than the conscious or narrowly cognitive subject has access to (Nisbett and Wilson, 1977). Stephen Stich thinks such features of mind, especially evident from various empirical research programs, place a strain on the folk-psychological generalisations at the basis of modern cognitive science in that they open a *schism* within the mind that is unprecedented in, and irreconcilable with, folk-psychology (1983:230-237). Likewise, part of the problem with the cognitivist model of the emotions, I think, is that the preferred generalisations and categories of cognitivism do not always seem to fit with the findings of empirical science.

This noted schism in mental structure might be due to the fact that several of the sub-systems in the brain can be seen as “modular,” in that they do not freely trade information with other systems. Jerry Fodor introduced the notion of modularity to the philosophical world in the short work Modularity of Mind (1983). His candidates for modularity are typically perceptual systems. Among the marks of modularity are informational encapsulation, autonomy of function, mandatory function, and introspective opacity.[16] The affective emotions Le Doux and others discuss are candidates for being at least partially modular systems, especially in that they have autonomy of function, are mandatory, and exhibit varying degrees of informational encapsulation, as I will argue. If they do have such features, the self-report Schacter and Singer identify as the basis of emotion is based on information not necessarily in the form that plays a role in the affective processing. Seeing affect systems as modular states, or somewhat modular states, would explain the discrepancy between affective states, and cognitive appraisals of emotional situations. If the physiological states at the basis of the emotions were modular it would be expected that our reports of their activation might be fallible.

---

[15] Though this is not to claim there is such a discrete entity as the *conscious part* of the mind, or what Dennett criticises as the “Cartesian theater” (1991).

[16] Note though that “modularity” does seem to be used in a variety of other ways at various places in the philosophical and scientific literature. Evolutionary psychologist and linguist Steven Pinker seems to use the term in a broader sense than Fodor (Pinker, 1994). I think what shows Pinker is not always using a precise sense of the term “module” is, immediately before listing his proposed set of modules, his calling them “families of instincts” (1994:420). This is a better characterisation of the psychological materials he lists. There are specific modules *within* these broad areas of interest – affect program fear modules explain some of what he refers to as the “Danger” module (1994:420) – but the broad areas of interest Pinker identifies seem to be functional characterisations of the proper content of a study of human psychology, rather than actual mechanisms. Pinker’s modules often seem to be his suggestions for a correct *taxonomy* of human

Important for understanding unconscious effects such as phobia is the fact that narrow cognition, or more generally, cognition that is in principle available to conscious introspection, is not the only way in which memory, such as that of pukekos, is accomplished in the brain. It has been shown that there is an “emotional memory” independent of familiar conscious, “declarative,” or “explicit” memory (Damasio, 1999:43-47; Le Doux 1998:145-146, 181-182). Unpleasant appraisals of pukekos can quite easily occur even if I have no consciously available memories of pukeko experiences. Le Doux also argues for an emotional memory that might be modular to conscious cognition in such a manner. Illustrating this separation between conscious and emotional memory is the fact that the emotional memory often remembers conditioned, or *affectively primed*, stimuli, long after explicit or declarative memory has forgotten about them, or even where explicit memory had no access to the priming incident in the first place (Le Doux, 1998:238). A pukeko may have bitten me in early childhood as the result of an unfortunate visit to a New Zealand native bird sanctuary, and even though I have now forgotten the incident, pukekos might still tend to elicit affect fear in me. Every time I see one on the side of the road I tend to cringe, and then try to run the damn thing over. The fact that emotional memories can persist long after explicit memory of the conditioning or priming event has faded means a subject’s explanations of subsequent emotional arousal may be wrong in the way I noted in the Schacter-Singer experiments (Le Doux 1998:203-204). I may attribute the physiological perturbations I have around pukekos to a general dislike of the birds of the *Rallidae* family. This suggests an explanation of phobias, where the emotional memory is especially salient and resistant to *extinction*.

The interesting idea of an emotional memory quite separate from familiar conscious memory is again illustrated in an example Peter Lamarque has drawn from the psychopathology literature (1996:155), and used in his work in the philosophy of art. *Capgra’s syndrome* involves patients confabulating stories about their close personal associates as a result of a failure of their emotional memory; this seems to reverse what is the case for phobias. Though sufferers are able to consciously identify their family and acquaintances, the emotional engagement ordinarily aroused

---

psychology. Stich also notes yet another use of the term “modularity,” in the sense that a process is modular when it is typically performed on different occasions by a single system in the brain (1983:238).

by familiar people does not arise. Sufferers of Capgra's syndrome no longer feel any emotional ties to their families and friends, even though they are quite conscious of whom the people are. Consequently they confabulate stories about the people being automatons, robots, impostors, or ghosts, to explain away their lack of emotional attachment to people to whom they should have strong emotional ties. Damasio illustrates an almost mirror image of this deficiency. In some patients with persisting amnesia, because of the memory deficit the explicit recognition of people and situations is extinguished, but the emotional attachment to familiar people is retained (Damasio, 1999:43-47). Such patients are likely not to remember the people they meet on a daily basis, and yet will display preferences and emotional bias toward those with whom they are *familiar*, and will avoid those who have treated them poorly in the past, even though they do not remember the occasions on which they were treated poorly.

It is for reasons such as this that Le Doux thinks Schacter and Singer, along with the propositional attitude theorists, really have no reason to conclude against physiology as being constitutive of affect or emotion. This is because cognitive appraisals might really often be distinct from the processes behind affect elicitation and constitutive of affect. Indeed this is just what the Schacter-Singer results seem to show. The fact that people may have an often very uncertain access to their emotional states might be called the *problem of attribution*; our mental attributions may not always be accurate, and the particular mental states one attributes to oneself may not always be present.[17] Indeed, eliminativists, who I will discuss shortly, believe whole *classes* of mental attribution will fail. Further, it means that, as an instrument of research, self-report cannot be relied upon. The notion of attribution will have important consequences for our conclusions concerning the fictional emotions, as some of the cases of fictional emotions seem to involve attributions where the putatively typical adjuncts of attributions, people and situations, are missing [6.3].

***

Emotions then, do not necessarily involve beliefs or other complex cognitive states. They may even be "non-cognitive" (Zajonce, 1980). It is obviously at this stage I will have to motion toward the debate concerning the character of beliefs and whether and in what form they will play a part in a mature scientific psychology. Though I am reluctant to enter this debate here, I will have to make some brief observations. We are often squeamish about attributing higher cognitive beliefs and desires, that is, the postulates of folk-psychology, to other animals. And if we are willing to attribute these states to other animals, it is usually those that are relatively close to us phyletically, and which share our way of life to some extent. But Le Doux's work suggests the animals to which we would have trouble attributing beliefs and desires share much of the brain architecture that in us comprises affect programs such as fear. Pukekos likely have affect fear (Le Doux 1998:172-174), but do they believe things? Le Doux thinks even lizards are likely to have affect fear (1998:172). This again seems to assert the difference between affect and higher cognition.

---

[17] This runs against Walton's contention in Mimesis as Make-Believe (1990:247) that we all have "privileged access" to the nature of our emotional states. I think what the Schacter-Singer experiments really show is that this is just wrong, but I am not surprised to read Walton writing it, as it fits perfectly with his propositional attitude leanings.

This is a *very* controversial question, and there is much to be taken into consideration. Many theorists would be glad to accept pukekos have beliefs in the exact same sense we have beliefs. Instrumentalists such as Dennett are an example (Dennett, 1987). I do not think this controversy affects my claims; the worries introduced are often more extreme and devastating for cognitivism than are the criticisms of cognitivism I offer here. It all obviously depends on what we think a belief is. We might be realists, and think beliefs are fundamentally representational or even sentential states manipulable by quasi-logical (*content-based*) rules. This is what cognitivists (*propositional* attitude theorists), interested as they are by complex *information processing*, are committed to in their characterisation of emotions. In this case, the claim that pukekos and lizards have beliefs of a kind that will satisfy their having emotions is at the least controversial; and so the observation they do have affect is likely to be difficult for cognitivists.

There are alternative characterisations of beliefs that are satisfied by pukekos and lizards, and so pukekos and lizards might be counted as having emotions by a cognitivist, however. Beliefs might be states of an organism readily described by "intentional systems theory," the instrumentalist calculus Dennett thinks belief-desire psychology amounts to (Dennett, 1987). Yet this is of no use to the cognitivist. Despite a willing ambiguity in Dennett's theory, I believe this is *not* a realist characterisation that would be of use to propositional attitude theorists. Indeed, Stich, who once had a similar view to Dennett's, thinks cognitivism's inability to extend its content-based psychological generalisations to peripheral cases – the young, the pathological, and to animals – is a fundamental fault with cognitive science (Stich, 1983).

Alternatively, cognitivists might decide to be *liberally functionalist* about belief states: beliefs are *whatever* explains the behavioural regularities that are the basis of the folk-psychological generalisation behind the usage of the concept "belief." Stich now seems to think such a move is available to the belief realist as a means of escaping the conclusions of eliminativism (1996). For a functionalist of this type, pukekos probably *do have* beliefs and therefore emotions: there is something about their nervous systems that fixes their functional role in their environment comparable to the functional role we have to our environment, and which we call belief. The particular functional role we characterise as belief would involve a story about behavioural commitments in counterfactual situations, for example. But if belief is characterised in this broad and liberal functionalist sense, then *even modular content may count as belief*. This cannot be what cognitivists mean by the beliefs involved in emotional cognition. Human beliefs invoked in the cognitivist story about propositional attitudes are something *very particular*. They are those states identified by Schacter and Singer as the subject of self-report, and which are typically identified with the content-based higher cognitive functions in the brain. If the cognitivist adopts a *liberal functionalism* in response to my criticism, then it is not necessary that beliefs of a type available to self-report are constitutive of emotions. The cognitivist has collapsed into the position I advocate here; sometimes, what we call “emotions” are not higher cognitive appraisals of situations. And this is why these important worries concerning the nature of belief are not a threat to the argument I am presenting here: they pose more fatal problems for the cognitivist than the criticisms I am offering. Cognitivists can hardly avoid my criticism of their characterisation of the emotions by introducing instrumentalism or a broad form of functionalism. Cognitivists are foremost realists about mental entities such as beliefs, and furthermore, have a very definite view about what sort of entities beliefs are.

These issues broach the previously signalled debate concerning whether the affective emotions could truly be non-cognitive. Robert Zajonce’s classic article on this topic has led to a significant debate concerning whether the mental states that subserve affective states could really be sufficiently reaction-like to deserve the designation “non-cognitive” (Le Doux, 1993b; Parrot and Schulkin, 1993a, 1993b; Scherer, 1993a; Zajonce, 1980). I believe what this debate shows is that the types of worries expressed in this chapter are a threat not only to naïve forms of propositional attitude theory but also to more sophisticated forms of cognitivism not so committed to seeing beliefs as indispensable to the emotions.

***

Le Doux goes on to identify the emotions with neurophysiological systems, and in particular, in the case of his specialist subject, fear, with various parts of what was labelled the limbic system.[18] He explains, via the introduction of various experiments he and others have recently carried out, how fear has been identified with the activation of a set of neural structures, in which the *amygdala* is particularly important. I will go into further detail, introducing other theorists and their ideas, when I discuss the affective emotions later in this chapter [2.4]. Such a characterisation allows his theory to circumvent the classic complaint cognitivists have concerning feeling or nativist theories. Focusing on the James-Lange theory, cognitivists complain that the feelings that accompany emotional episodes are not sufficiently different in character to differentiate all the varied emotions. But their complaint is misdirected. Le Doux’s revised conclusion on the Schacter-Singer experiments show exactly what the cognitivists urge; the phenomenology of an episode cannot be used to characterise the emotional event, but then, following the understanding that introspective discrimination is not a sure fire path to establishing a mental ontology, it is not meant to. What individuates the emotions is not a subjective feeling; *it is a substantive neurophysiological or physiological difference*. Fear is different from joy because the former is run on a different neural structure to the latter, is aroused by different eliciting states, and is wired up to different behaviours.

---

[18] It is no longer labelled the limbic *system* because it is not a system; the various parts of the brain once called the limbic system are no longer seen as either neurophysiologically *discrete*, or *specific* to emotion, as once thought (Le Doux 1998:98-102).

Paul Ekman concurs and elaborates on these points: typically an emotion is “complex, entailing a number of different response systems […] the changes occurring within each (or most) response systems are distinctive for one as compared with another emotion” (Ekman, 1980:81-82). For Ekman, directing the progress of an emotion is an “affect program” (1980:82).

In fact, as far as the substance of the emotions goes, feelings are a red herring. Le Doux provides a convincing story of how feelings arise in the emotional consciousness (1998:267-303). The term “emotion,” if I want to get a scientifically interesting story, should not be seen as synonymous with “feeling.” Emotions are complexes of physiological mechanisms. Feelings arise when the effects of these physiological complexes bubble up into the cognitive workspace of consciousness. In Le Doux’s work, and in that of Damasio (1994, 1999) it can be seen that William James’ story about bodily feedback is currently experiencing a revival, though it is not taken as an explanation of the emotions, merely of the feelings that often attend emotional episodes. Damasio also thinks emotions and feelings are quite distinct, and that many animals without the possibility of having feelings (because they lack the requisite consciousness) still have emotions (Damasio, 1999:37). I discuss feeling later [2.4].

So, following Le Doux and other brain scientists, I end up at something of an impasse. Emotions such as fear, sadness, anger, surprise, disgust, and joy, are *things in the brain*; specifically, the activation of neural structures built by evolution to help the organism respond adaptively to its environment. Yet this does not fit well with many of the findings of modern emotion theory, particularly that of philosophers and social scientists. In particular, the *emotion in the brain* theory does not seem to do justice to the apparent *cultural* and *social* aspects of some emotions. Many emotions, it is held, have a distinctly cultural nature, and vary amongst cultural epochs and divisions. Catholics have their guilt; the Japanese, the deep respect for authority called *amae*; the Christian ascetics had an obsessive religious devotion that led them to perform extraordinary acts of literal and metaphorical self-flagellation; and audiences of artworks, apparently, have (or once had) the culturally particular emotion of *catharsis*. Moreover, most people can attest to having emotions without any obvious physiological arousal, and emotions that persist for years; it is improbable to claim such an emotion might be based on the continued activation of a discrete physiological network such as those described by Ekman and Le Doux, as these are typically *short-term* responses (Ekman, 1980:80-81). Paul Griffiths’ work in the philosophy of biology aims in part to reconcile the findings of the affect theorists with the view of emotion as culturally emergent.

2.3 Eliminating “Emotion”

Our problem is that the emotion theory one finds in much of the scientific literature – particularly that holding emotions are neural and physiological episodes – is unlikely to prove satisfying to a large proportion of those who want to discuss the emotions, even though it is based on a sound and productive scientific program. Some will think that obvious counterexamples to the theory, proving the proposed stereotype fails to capture some set of apparent emotional phenomena, show the theory is naively and biologically *reductionist*. Such theories, it will be thought, do not do justice to the important role of higher cognition and culture in the nature of emotional responses. It is equally unlikely that arguments in the higher cognitive and cultural directions will eventually provide a stable stereotype able to capture all the putative emotional episodes, however. They are bound to seem to some, especially biologists, unrealistic in their claims for the cultural source of the emotions and their degree of cultural variance. It may seem a particularly pessimistic induction on my part, but I do not believe the solution to this bind exists in providing more and better arguments in either of the *biological* or *cultural* directions. The solution will come in undermining various assumptions that exist within emotion theory. Foremost, it will come in undermining the assumption that there is one scientifically adequate *category* corresponding to the vernacular *concept* "emotion."

Critical to this argument is the claim that emotion concepts are terms within a folk-theory of psychology. "Anger," "sadness," "guilt," and their like, are terms in a folk-psychology aimed at explaining and predicting the behaviour of ourselves and those around us. They stand alongside folk-explanations of phenomena such as water, fire, floods, and stars, as relatively naïve attempts to discern the nature of portions of the universe. Further, our particular folk-theory has been around for a *long time* as far as explicit theories go – at least as long as Thales' now discredited theory that the elemental substance is water. We should not be surprised therefore, noting the difficulty science had in explaining these other features of the natural world – it took several brilliant scientific minds and years of ingenious empirical study to work out what stars were – if it turns out that our present explanations of our behaviours, vocabularies, feelings, and physiology (those things notable in our emotions) are off the mark (Churchland, 1981).[19]

Folk-psychology is on a par with the *folk-physics* among the platitudes of which is *what goes up must come down* and the like, and that allows us to make predictions of the behaviour of the middle-sized objects around us. Like folk-physics, folk-psychology *might turn out to be wrong*. Indeed, physics as practiced by everyday folk unaware of even the most rudimentary of scientific physical theory, often turns out to be disastrously wrong (Stich, 1996:11-13). I am not sure this has been subjected to study, but I would be prepared to bet (based on some very informal anecdote gathering) that a sizable proportion of the population would be willing to assert that if the *world stopped spinning then everything would float off into space*. This ability of theory to be wrong theory is perhaps one of the central reasons why philosophers invoke the notion of a folk-theory. Another realisation I might make is that, in dissecting our concepts and their implications, we will not necessarily find things out about us. Rocket scientists would get things disastrously wrong if they based their calculations on folk-physics.

---

[19] Stich thinks this and many other arguments offered in support of eliminativism are particularly weak (Stich, 1983). There is neither time nor reason for me to launch a defence of eliminativism here.

The application of the term "theory" to those of our everyday descriptions and explanations that invoke emotion terms will seem counterintuitive to most of those unfamiliar with the philosophy of mind. Indeed, many philosophers think the appellation inappropriate. There do seem significant similarities between emotion terms, and the terms of a developed science, however. People use emotion concepts to pick out *categories of items within the natural world* – such as *angry people*, *sad situations* – and to make *predictive generalisations* or *inductions* on the basis of those categories. We might reason from the facts that sad people tend to cry a lot, that personal loss tends to make people sad, and that our acquaintance has just suffered some terrible loss, to the prediction that our acquaintance might begin to cry a lot. The term "sad" is the basis for a set of generalisations about which classes of situations tend to lead to which classes of behaviours. Folk-psychology bases its law-like generalisations in the postulation of the existence of an *inner causal mechanism*. The behavioural and qualitative regularity is due to a mental state that causes the behaviour when the appropriate situation is presented. These inner states are beliefs, desires, intentions, emotions, and their like.

Often, like developed scientific theories, the states that are postulated are *unobservables*. Though one might think our emotions are observable in that we *feel* them, this equally could be used to argue *divine revelations* are observable phenomena. What is observed in these cases is just the feeling, and as I noted in the case of the Schacter-Singer experiments, there is a further step to be taken in establishing the *observation of an emotion* from the *observation of a feeling*. Furthermore, this inner causal mechanism and the emotion generalisations that are based on it are used as *explanations* of people's actions; our friend is crying a lot *because* they have suffered some profound loss, and not because they are continually slicing up onions. In this case the former theory is to be preferred to the latter because it best conforms to our *observations*; New Zealand has just lost *another* rugby union test to France, and onions are nowhere to be seen.

There are differences between folk-theories and developed scientific theories, however. Scientific theories are explicitly developed in the course of research, whereas folk-theories are tacit theories in that they are neither explicitly developed, nor, evidently, necessarily explicitly or consciously held in the mind. Using a folk-theory does not involve manipulating explicit rules in a manner that resembles calculating from a scientific or mathematical theory. Indeed, if the belief-eliminativists are correct, folk-theory will not be held in the mind as a set of contentful states such as beliefs![20] Folk-theory need not be thought of as a mental list of propositions or as existing in a

---

[20] Stich thinks this counts against the eliminativist argument that the folk-theory of beliefs is mistaken. As one of the horns of an eliminativist dilemma, he argues that, because of a form of

"declarative linguistic format" in the mind (Stich, 1996:10), but can be minimally defined as a set of *knowledge structures* that make possible the prediction and explanation of people (or indeed, animals, where they share behavioural and cognitive regularities comparable to those of our species[21]). This minimal view of folk-psychology as a knowledge structure is to be expected if there turns out to be no language of thought. Also, if I stick with this liberal usage, other species are also likely to have their own folk-psychological abilities, perhaps homologues of ours, and so language is not a prerequisite for folk-psychology. This is an obvious difference from any scientific theory. The use of "theory" as a description of these knowledge structures might turn some philosophers against the view, but the role of these putative structures in *prediction* and *explanation* is likely to convince most that the term is appropriate.

It is even more likely "theory" does capture something of the state of affairs if I note much folk-psychologising *does* exist as verbal and written declarative expressions, minimally formalised in its commitment to a consistency of theoretical terms such as "beliefs," "desires," and "emotions." Folk-psychology does not exist solely as practiced in the mind, but is reported and written about by everyday folk in their dealings with others, and in their cultural practices. Folk-psychology exists as a set of *platitudes*. It is in these expressions of folk-psychology that its generalisations are most obvious, and one does not need to look far to find them; listen to a few casual conversations, especially of the gossipy variety that occurs over coffee, to find evidence of folk-theorising. The products of cultures, in the form of reported utterances, historical documents, newspapers, television talk shows, and art works, are also a rich source for what people have believed about human psychology. Significantly, one of the most obvious uses of folk-psychology in cultural practice is in the narratives, sometimes fictional, we tell each other for entertainment.

It is at this point that a common methodological practice sees an opening. Some theorists have seen an opportunity for the study of the emotions via an explication and formalisation of the generalisations found in folk-utterances, namely, a *conceptual analysis* of emotion term usage (see

---

mental holism inherent in it, if something like the connectionism that eliminativists often advocate is correct, then the mental states that underlie folk-theory will not be a type of state that can be characterised as having discriminable content (1996:130). Folk-theory then, *does not make any singular claims about the world*, let alone claims that might be false. Folk-theory cannot be incorrect! "But wait," I hear you say, "this argument assumes that connectionism is an accurate characterisation of the mind doesn't it? Therefore, Stich seems to be using a form of eliminativism to argue against eliminativism!" At this point one feels as if one is watching a *Star Trek* episode where Captain Kirk must go back in time to stop himself threatening the space-time continuum by going back in time to stop himself threatening the space-time continuum… But of course Stich intends us to avoid this ingenious *reductio* by picking the other horn of the dilemma, in which it is not clear folk-psychology is under threat.

[21] Or, as Dennett thinks, even coke vending machines (1987:290-295).

especially Harré, 1986). Rom Harré advocates “Unravelling the basis of usage [of terms such as “anger”] will lead us deep into the heart of emotion theory…” (1986:4-5). But this is ambiguous: if Harré means such a study would lead us deep into the understanding of *folk-theory* of emotion, he is certainly correct; we would find out a great deal about what people believe about emotions, and how they use the emotion language to explain, predict, and characterise the people and situations around them. I doubt if this is what Harré hoped to establish; surely he wanted to talk about the emotions themselves, not merely what people believe about them? The more exciting claim then, is that unravelling the basis of usage (of terms such as “anger”) will lead us deep into the heart of a theory *of what emotions are, despite what people believe about them (which might be false).* This is certainly far from uncontroversial. On this view we might expect unravelling the usage of terms such as “witch” or “ghost” would lead us deep into a theory of what witches and ghosts are. It would not, because there are no such things as witches and ghosts. Such a study might coincidentally tell us something about how people behaved around witches and what they believed about them, or what putative witches looked like, but it would tell us absolutely nothing about witches. Witch theories were just that, theories, and they were also wrong. So too might folk emotion theories be wrong, and, if they are, the dissection of emotion talk is going to be far from a good theoretical method.

Griffiths subjects Robert Plutchik’s derivation of an “emotion solid” to similar criticisms (Griffiths, 1997:74-75). In an attempt to characterise the nature of the emotions, Plutchik conducts a “semantic field analysis” of various emotion terms, asking subjects to compare emotions such as grief, terror, loathing, or surprise in terms of similarity and dissimilarity. He then maps the results onto an emotion solid to display the relationships between the various emotions. Furthermore, Plutchik makes evolutionary and ecological hypotheses concerning the relationships found. Griffiths makes the telling point that Plutchik’s model of the emotion “is essentially a model of what English speakers think about emotions. Plutchik’s method is based on the assumption that ordinary English speakers already have a correct theory of emotion!” (1997:75). Plutchik places too much faith in folk-theory, or does not see it for the theory that it is. Because conceptual analysis tells us only what people believe about psychology, it may be unlikely to turn up substantial truths about the mind unless we can assume people *already have an accurate theory of psychology*.

Yet there is good reason to think conceptual analysis is partially immune to such criticisms. One could argue physics and psychology are not comparable fields of study. There are crucial differences between the role of concepts in physics, and their role in psychology. Foremost is the fact that in psychology, concepts play a *causal* role. Unlike in physics, what people believe about psychology will change the nature of their own psychology and human psychology in general, *because beliefs are psychological episodes*. The state of believing that people who are sad tend to cry a lot might conceivably play a causal role in producing crying behaviour in sad people.

However, though a successful characterisation of psychology and the emotions must take this causal role into account – in a similar connection Griffiths notes that a theory of concepts must take into account the fact that concepts are politically and socially dynamic states (Griffiths, 1997:198-200; see also, Hacking, 1995) – I do not believe conceptual analysts can take much comfort in this. Though concepts like “love” – if love is a *socially constructed* emotion – will play a causal role in that those judged to be *in love* will be interpreted differently, and will adopt expressions and behaviours typical of those *in love*, it is far from clear that the causal mechanisms structuring these behaviours are implied by the generalisations inherent in the usage of love concepts themselves. “Love” concepts might just have it wrong about how the notion of *being in love* works; indeed this seems likely if love is held to be, as, drawing on my observations of daytime television, I expect it often is, some type of *spiritual connection*, that involves *soul mates*, and *angels*. The folk-view, of course, does not have it that love is socially constructed. We should not expect that implied within our concept usage is a proper theory of how concepts, situated as they are within a population of language users subject to the causal influences of intuitions, traditions, and other cultural aspects, do their causing. There are reasons why the generalisations inherent in our concept usage, where these concepts are themselves causal entities, might be as misleading about the underlying causal mechanisms, as I noted language about physiological episodes often is. I will have a little more to say on this topic later, when I discuss the social-construction of the emotions [2.4].

Rather than deserving our immediate and untested deference, the concepts of folk-psychology must take their place alongside other theoretical posits, subject to the methods of science, and so are *falsifiable by future discoveries*. Folk-theories are possibly false theories. This means not only that we might change our mind about the properties various emotions have, but we might also *eliminate* the category from our scientific vocabulary, or even likewise eliminate the entities themselves. In the first case, we might find a folk conception of a particular emotion is wrong in its details, and that our conception of the particular emotion should be revised. This kind of revision is a common practice in science. The plum-pudding model of the atom had it that the atom was *plum-puddingish*, yet the work of Rutherford showed otherwise, and atom theory was revised.

The second situation is a case of *category elimination*, where the *extension* and *intension* of a concept might be revised in light of new discoveries. The extension is revised when phenomena are either left out of the category, or newly included. The intension or meaning of the corresponding concept is revised to fit the new extension. An example is the revision that often happens in biology when new evolutionary relationships are discovered. One can imagine the tribulations that occurred when the monotremes were discovered. (Actually one does not have to imagine, as Umberto Eco has written an entertaining summary of the tribulations [Eco, 2000:241-248].) With the discovery monotremes were mammals, the extension and intension of the category term “mammal” was forced into revision. The extension of the category was changed to include the platypus and echidna, and further, the intension of the concept was altered so that *nipple bearing* was no longer among the necessary conditions of being a mammal, and being oviparous was no longer a sufficient condition for counting something outside the category. Therefore, in a sense there was no natural class that corresponded to the pre-revision sense of the category.

The final case is one of what I might call *entity eliminativism*. Here it is a case of a theoretical entity being dispensed with altogether because it is thought nothing in the newly revised theory comes sufficiently close to the previous entity to deserve its name. The traditional examples are *phlogiston* and *caloric* in chemistry and physics. I might also say that various modern bodies of theory, from evolutionary theory to cosmology, have seen the elimination of God as a theoretical entity in serious science. Though there is much dispute concerning this latter form of theory revision, in particular in connection with beliefs, I will not have much to say about it here. It is not clear anything quite as exciting as entity elimination is going to happen in the case of the emotions. I will be far more interested in the second case: category elimination.

Since terms like anger, fear, love, and guilt share many similarities with the theoretical terms of an established science, this makes them prone to many of the same standards we apply to scientific terms; especially when they are to be incorporated into a developed scientific theory. Should the emotion terms come up short of a methodological standard we set for scientific terms and concepts, then the emotion terms should not play a role in future descriptions of the domain they putatively applied to. To retain their place in a serious science, categorial concepts need to pick out categories that are rich with predictive and explanative generalisations; that is, scientifically useful concepts need to be “projectable” (Griffiths, 1997:174). Where relevant, generalisations based on the concept should hold in future instances of the category. If a new substance is discovered to be of a *kind* with a previous substance, it should be expected to share many of the properties of the type. Given a new discovery about the substance, we should expect this discovery to be reflected in the rest of the category.

Griffiths contends that the basis of projectable categories are “causal homeostatic mechanisms” that explain *why* the concept is rich with generalisations (1997:188ff). Legitimate categories are based on underlying natural features or “essences” that bind the category together, though “essence” has shed many of its traditional implications:

> The concept of causal homeostasis entails a very broad conception of the essence of a category. An essence can be any theoretical structure that accounts for the projectability of a category. The microstructural essences that have received so much attention in discussions of natural kinds are merely one kind of essence. Even the stock examples of natural kinds do not all have microstructural essences. The fundamental kinds of chemistry, such as the elements of the periodic table and categories like alkali and acid, have their properties because of their internal microstructure. But biological taxa, the other classic example of natural kinds, turn out to be united by external forces. Biological taxa at all levels of the taxonomic hierarchy form projectable categories because their members are descended from a common ancestor…
>
> (Griffiths, 1997:188-189)

Important here is that the “causal homeostatic mechanisms” have varied origins, and as in the case of species, this might imply that a category is not united simply by *material similarity*, but by another theoretical posit such as *history*. This means *material dissimilarity* need not be a bar to counting an object within a category. Birds and dinosaurs are materially very different, but at a certain, relatively close level, belong to the same taxonomic grouping.[22] Conversely, it means material similarity is no reason in itself to count objects as residing in the same category. Bats and birds have similar features, yet bats are no more easily categorised with birds than humans are.

In formulating respectable categories, scientific theories should attempt to characterise the homeostatic mechanisms at their basis. In fact, the causal homeostatic mechanisms *are* the identities of the categories science deals with. A species *just is* a *geographically isolated group of interbreeding organisms linked by descent from a common ancestral population*, because this is why the properties of the category are projectable to all normal individuals within it. Likewise, an acid or alkali just is a substance with a particular microstructural essence. Given this, a theory of the emotions should attempt to discover a causal homeostatic mechanism at the basis of emotional phenomena, and this theoretical postulate, just is, as Griffiths says, “what emotions really are.” A theory of the emotions is foremost a theory about the *mechanism* that binds the emotions into a projectable category.

However, in the case of the vernacularly defined emotions, it is not clear that a mechanism exists that can bind the vernacular category together. The wide disagreements contained in the folk-theory of the emotions may in part show that the folk-category of the emotions has no rich and coherent set of projectable generalisations. I noted earlier the inability of traditional theories of emotion to find a framework under which all apparent emotional phenomena might be subsumed. Feeling theories cannot contend with the cerebral or unfelt emotions, propositional attitude theories with the non-cognitive emotions, and emotion in the brain theories with the putative social or cultural emotions. The mechanisms proposed by each – feedback from bodily arousal, cognitive appraisals of environmental situations, and kinds of physiological and neurophysiological events – all seem incapable of binding together the diverse set of phenomena termed the “emotions” in folk-theory. A simple (though perhaps expeditious) way to resolve this difficulty would be to conclude that there is no single framework or theory under which all these phenomena can be subsumed. Given Griffiths’ view of the nature of scientific categories, a correlate of this would be to conclude there is no single causal homeostatic mechanism binding the folk-category together. This is exactly the conclusion Griffiths urges.

Griffiths thinks the folk-category “emotion” is an *arbitrary* category, based on *superficial* material or functional similarities, and will play no role in a mature science of the mind. The vernacular, or folk-psychological concept “emotion,” and research programs that respect its categories, muddle together various different phenomena. Though these phenomena may share *superficial* similarities that might have led to the construction of the folk-category, the progress of science is revealing significant disunity within the putative kind. The examples of emotional pathology familiar from the literature attest to this. The folk-category “emotion” is arbitrary in the way the Aristotelian category of “superlunary” is arbitrary (Griffiths, 1997:1). The concept picks out no projectable category on which to base predictive generalisations.

---

[22] Of course, *every* living thing belongs to the same taxonomic grouping at some level of generality, barring the improbability that life began twice or more on the Earth.

The superficial similarities that make the vernacular category seem so natural are likely to be regularities in the *classes* of environmental situations the emotions respond to. The folk-theory has it that people respond to danger with fear, to loss with sadness, and to novelty with surprise. Cognitivism elaborates on this model by proposing the mental mechanisms at the basis of this regularity are propositional attitudes, the paradigm of intentional states in cognitive science. The recognition of the environmental states of danger, loss, and novelty happens via a cognitive appraisal. Cognitivists have not seen the possibility of mental states affixing to particular classes of environmental situations, in the *absence* of paradigm intentional states such as propositional attitudes, however. The unseen alternatives are simpler, often unconscious, even non-cognitive and reaction-like affective states. Though very different to cognitive appraisals, these fit the vernacular concept “emotion” because in a sense even simple affective states, despite often being mere reactions to simple perceptual stimuli regularly associated with particular situations, are *about* those stimuli because they arose as adaptations to those situations. I think this is enough to show that the vernacular emotion concept, should it be based on such functional features, does not cut deep enough to be interesting to the scientist.[23]

Given the relative success of the various research programs in progress, we might expect a revised conception of the phenomena covered by the vernacular concept “emotion” will postulate a *variety of kinds*, each bound by its own causal homeostatic mechanism. The *emotion in the brain* and *social* emotion research programs are turning up much that is of interest, yet that is unlikely fall under a single super-theory. The separation of emotion research into a variety of programs concentrating on different phenomena would be a productive way to continue with emotion theory. In as much as this happens, however, emotion theory, because it is based on the assumption there is a projectable category corresponding to the vernacular concept “emotion,” will have been eliminated.

---

[23] Of course this takes me into the difficult territory of the origin of mind and intentionality, an area in which there is presently much interest. Several such works are consistent with the view just expressed in that they see intentionality as not solely a matter of highly cognitive representational states. Richard Dawkins, admittedly dicing with metaphor, notes that organisms are “descriptions” or even “pictures” of their environments (1998:239-242) and Dennett, a philosopher particularly interested in evolution, goes as far as to say that complex cognition is but a special case of the intentionality exhibited in adaptation/environment relationship (1995:200-207). Kim Sterelny develops a similar view of cognitive evolution in a recent collection of articles (2001).

There are several promising approaches to the conceptually revised study of the material that once comprised the emotions, though it must be admitted that is still very murky territory. I am much more convinced by Griffiths’ negative project than I am by the observations he makes about what will replace the eliminated emotion theory. Griffiths thinks three promising approaches are beginning to cohere in the emotion literature: “affect program” theory, the various theories of “higher cognitive emotion,” and theories of “disclaimed action emotion” (Griffiths, 1997).

Yet it is still far from clear what the causal homeostatic mechanisms behind these categories are, and indeed, I can conceive these categories themselves might be subject to a more radical form of Griffiths’ negative project. I question whether the affect programs are all of a kind. Given that they likely evolved at different times, are implemented in different parts in the brain perhaps not forming anything like a *system*, and that there are borderline cases such as startle and laughter, it remains possible that “affect program” should be eliminated as a category.

Yet it is undeniable there is promising movement in emotion theory, particularly within the neurosciences. In the next section I will discuss those movements that seem to have a bearing on the question at hand: fiction and the emotions. My collection of the emotions under various headings, “affective” emotion, and “social” or “cognitive” emotion, is not intended as a strong claim about scientific kinds, but is made primarily for expository reasons. I make no claims about the probable outcome of a mature science, as I am in no position to do so, but I use groupings that will best cater to the kinds of problems I think exist when fiction and the emotions interact. At the beginning of this chapter I claimed that a careful study of the emotions might affect our characterisation of the problem of the fictional emotions; I hope it is now beginning to become clear why.

***

Before I move on to the next section, I want to cover myself against a charge that might be levelled against the category eliminativism I advocate for the vernacular category “emotion.” The charge originates from philosophical functionalism. Griffiths thinks that in our search for the emotions, we should be attempting to find causal homeostatic mechanisms that stand behind mental categories, and these mechanisms will provide the basis of discriminable mental kinds. Further, it is thought it is unlikely that a single homeostatic mechanism would be found that unites the diverse material termed the “emotions” in folk-theory. It might be suggested, however, that a *functional* characterisation might be formed that binds the apparent emotional states together? Mental types might be defined by a functional characterisation such as that which defines the categories “wing” and “eye.” Both wings and eyes have evolved several times in evolutionary history, and so the various extant examples of wings and eyes differ in both historical origin and method of implementation, two places where we might have been tempted to ground the categories. “Wing” and “eye” are grounded on functional observations. The vernacular concept might be based on a superficial observation of functional regularities in mental and behavioural activity, so perhaps the functional schema can be worked into a respectable homeostatic mechanism to ground the vernacular concept?

I think it is unlikely a functional characterisation is what we are really after in this case, for a variety of reasons. Functionalism’s ability to satisfy our intuitions about our world, along with the manner in which it allows us to establish broad scientific categories, explains functional characterisation’s current popularity, especially in the philosophy of the mind.[24] The functional characterisation of mental states allows a level of generality in forming *task-descriptions* of the mental states we would expect to fit with the empirical data scientists are collecting about the mind. Without this *top-down* level of generality, it is unlikely the mind sciences would progress very far very fast. In biology, functional characterisation *is* an important part of empirical study, yet only when pursued in conjunction with several provisos I will now argue for. Functionalism is best seen as a heuristic mechanism, as it ignores difficulties with the notion of *function* in that the notion is *artificial*; it often seems *arbitrary* to values and assumptions that are already inherent in our approach. Furthermore it is insufficient for establishing biological kinds (which many putative mental kinds are), as these are established only when informed by a *historical* characterisation that sees the origin of an organism or traits as having final say. Le Doux also voices his opposition to the functional characterisation of mental states, particularly because it abstracts away from the true theatre of the emotions, the human body and brain (1998:39-41). I am also sympathetic with this view.

There are difficulties when we come to the particulars of human psychology in using a functional characterisation to determine what is to count as a particular state or, important to my argument here, what is to count as an emotion. One of the chief virtues of functionalism of the mind is that it is liberal; it allows us to see as minds, or as mental events, things other than particular brain states that instantiate those properties in normal human subjects. A problem though is that in order to capture this liberal generality, functionalist theories may not be fine-grained enough to pick out what are significant *local* features of broadly functionally equivalent states. On a certain level of generality, reaction-like affective states induced by perceptual stimuli, and more complex or *deliberate* cognitive appraisals of environmental situations might *both* satisfy a single functional role or task-description, for example, in achieving danger avoidance behaviour, that is, in establishing vernacular fear. Yet below this level of functional generality is a great deal we might want to know about the states in question. Indeed, later I will be arguing that in the case of fiction and the emotions much does hang on how particular or definite we are in our conception of emotion.

---

[24] Though functionalism, in a benign sense, often seems to be the *theory of mind you have when you do not have a theory of mind*.

In cases where there is still more we want to know about our functional states below the shared level of functional identity already characterised, it might be argued we simply have not carried the detail of our functional analysis *far enough down*. But the problem is that the more fine-grained our characterisation of the state gets, the more definite, and less amenable to liberal functional characterisation the state becomes. At a more precise level of characterisation reaction-like affective states induced by perceptual stimuli, and more complex cognitive appraisals of environmental situations begin to take on significant differences, to the point that we begin to ask ourselves – are they still both of a *kind*? For example, could tokens of both processes amount to tokens of the psychological kind fear? Are they still the same functional state though *implemented* on differing architectures? This claim becomes difficult because the differing architectures themselves seem to add to the functional nuances that need to be described. Indeed, some have questioned the validity of the function/implementation duality that seems central to these questions (Lycan, 1987). In a sense the division between function and implementation is *artificial*, because it depends on our theoretical purposes or assumptions. For this reason it is unlikely to provide the source of biological kinds.

I think functional characterisations give broad conditions for type identity statements that will satisfy us to a particular level of generality and are important heuristic mechanisms. Yet the level of generality almost seems to be *arbitrarily set by the nature of our interests*; we might want to know how an organism locomotes, or how it flies, and then decide that, like certain insects, birds, and like aircraft, it has *wings*; we might wonder whether something is capable of *thinking*, and come up with a functional analysis of mind. Furthermore, when issues from outside science proper begin to have an influence on our categorical methods, functional characterisation may take on even more importance. Moral and political issues cannot help but having an influence on our use of concepts, especially when it comes to such a *touchy subject* as the mind (Stich, 1996:71-72). But when we begin to ask questions – such as “how does this thing work?,” “what it is like when it works?” – questions that will be crucial for our purposes, we will often need to descend to a level of more complexity that is less amenable to functional characterisation; to a level, that is, where the architecture *just is* the manner of implementation. This is nowhere clearer than in the connectionist theory of mind, where functional units come very close to looking like neurophysical units (P. S. Churchland and Sejnowski, 1990).

The brain is a messy thing on the level of implementation; it tends to do what works, and can often over-satisfy a task description. Evolution is not particularly parsimonious. It is quite possible, therefore, that the liberal mental functional roles identified by functionalist typologies, are achieved in the brain – the thing that constitutes the substance of the mind *around here* – by mental entities best described as *different in kind*, even though *roughly* functionally equivalent. It might become clear, once we arrive at a suitably complex characterisation of the emotions, that the *stuff we call fear*, might be both, to continue the present example, reaction-like affective states induced by perceptual stimuli *and* more complex, deliberate, or “ponderous” cognitive appraisals of environmental situations (Ekman, 1980:84).

But at this point, if we want to know whether these two things are *really* the same kind of thing, in the sense of a *projectable psychological kind*, their functional equivalence has nothing to say on the matter; foremost, we are in search of a causal homeostatic mechanism. Griffiths thinks in the case of the affective emotions, that the causal homeostatic mechanism amounts to their sharing a common origin or historical source. In nature there are many roughly functionally equivalent traits that are different biological kinds; the eye (*functionally*, a biological apparatus for seeing, that exploits electromagnetic radiation of a particular range of wavelength), has been reinvented many times in nature. *Convergent evolution* (Sterelny and Griffiths, 199:245-246) has allowed there to be various evolutionary trajectories that have produced eyes, and so there are various biological kinds we usually group together and call eyes. Though functionally similar, the eyes of insects and those of humans work in a significantly different manner. Basing our predictions about insect eyes on what we know about human eyes is unlikely to be successful, as there is no reason beyond the broadest of physical constraints why the properties of one should be projectable onto the other.

The problem with functional categories in establishing biological kinds is that they explicitly leave out the *causal* aspect of Griffiths’ theoretical desiderata. A way to remedy this complaint is to introduce a mechanism via which functional similarity is seen to rest on a causal relation. A possible mechanism is *ecology*. The claim in using this notion to substantiate a category would be that traits are of a kind when they are adaptations generated as responses to ecological regularities. Wings would then be a biological kind because they all exploit a feature of the environment, the atmosphere, to an adaptive end, that is, locomotion. Griffiths criticises this move, mainly because there is no developed science describing the manner in which organisms might interact with ecological regularities to produce adaptations (Griffiths, 1997:240). There is also the familiar scepticism about whether a thoroughgoing theory of adaptation could get off the ground to contend with (Gould and Lewontin, 1979). Griffiths thinks we are left with a commitment to “adaptive-historical” methods of analysis of trait kinds, as these provide an obvious mechanism on which to base trait identity: comparative “cladistic” relationships (Griffiths, 1997:112-114).

Such evolutionary observations can only inform our view of the mind (though Griffiths and others often reserve their doubts for *evolutionary psychology*). The fact that the brain has evolved over time has its consequences. For one, we should expect similar features to those presently comprising the human mind, or *homologies* of our mental traits, might be discerned in our close genetic relatives. Cladistic theory will play an important role in informing our view of the structure of the mind, and in how to study the structures. Darwin, Le Doux, and Ekman would agree wholeheartedly with this point. Their views of emotion are of a mental trait (or more likely *set* of mental traits) that has evolved over time. Emotion is likely to be present in the animals with which we share a genetic past. Le Doux suggests evidence for this will be turned up by a comparative empirical study of the various physical systems that instantiate emotional activity in the brains of our genetic relatives. This assumption is illustrated by the very fact that Le Doux's study is carried out on those genetic relatives.[25]

2.4 Approaches to the Emotions

In this section I intend to give a brief outline of the various approaches to the emotions that seem particularly viable at this point in time, and descriptions of some of the features we might expect the mental states described by these approaches to have, especially those pertinent to the issues of fiction and the emotions. Two predominant forms of emotion will emerge from this discussion, *affect*, and the *higher cognitive* or *social* emotions. My ordering of the emotions as affective emotions and the cognitive or social emotions reflects *expository purposes*, not a commitment to actual *kinds*. Indeed I am sceptical these have been established at this point. This thesis is more of a survey, and less of a theory. And it also includes phenomena most would not call emotions, though I think I should pay attention to these in the light of the purposes of this study. I will now set out their likely basic features so I can use them in my characterisation of the relationship(s) between fiction and the emotions.

i.) Affective Emotions

Griffiths singles out a first approach to the emotions as "affect program theory" (1997:77-99). In this study I will henceforth refer to a class of *affective emotions* the features of which are drawn out of various research programs. What sort of thing is an affective emotion and how is it elicited? Our most promising sources are the work of Ekman on affect programs, Damasio's work on the links between emotion and higher cognition, and Le Doux's specific ideas on fear in the brain. Affective states are features of the brain's physiology, or as Le Doux says, "biological functions of the nervous system" (Le Doux, 1998:12). They are *short-term* neurophysiological patterns of response, elicited by situations that have been established through learning in both the species and the individual as worthy of a response. As such, affect programs are "complex, co-ordinated, and automated" (Griffiths 1997:77) in being comprised of facial and bodily responses, vocal changes, endocrinal responses, attention orientation, musculo-skeletal responses, and autonomic nervous system arousal. It is typically proposed that there are six types of affect responses; they correspond to much of what is referred to by the vernacular emotion terms surprise, joy, disgust, sadness, anger, and fear. I will stick to the terms "affect program" or "affective emotion" when talking of these specific states. In this thesis I will also be a little more liberal in my usage of the term "affect," including other states, such as startle, anxiety, nausea, and various forms of arousal, including sexual arousal, under the “affective” banner. I think it is important to take note of these things in connection with fictive participation, and that these states share many similarities with the strict affective states.

---

[25] Most of the experiments were carried out on animals such as rabbits, cats, dogs, rats, and monkeys. These animals are reasonably close to us on the phyletic tree. Ethics committees are no doubt the curse of a neurophysiologist's experimental career: one can almost sense Le Doux's desire to try his stuff out on a *really close* phyletic relative. Perhaps the French?

Ekman proposes to give a “neurocultural” theory of emotion that will reconcile the work happening in the neurosciences on the emotions with the established view that emotions have an identifiable cultural component (1980). Emotions are “brief, often quick, complex, organised and difficult to control” (Ekman, 1980:80), features that are hallmarks of physiological reactions, and yet the situations in which they occur are culturally variant. Anger and sadness in one part of the world might be elicited by the loss of a pig to a predator, and in another by the loss of a rugby union test match to the French. The reconciliation is achieved by the postulation of a neurophysiological mechanism capable of being triggered by environmental features the significance of which might be *learned*. Central are the notions of an “automatic appraisal mechanism,” an “affect program,” and a set of “elicitors.” The automatic appraisal mechanism “must be constructed so that it quickly attends to some stimuli, determining not only that they pertain to emotion but to which emotion, and then activating the appropriate part of the affect program. The automatic appraisal may not only set off the affect program and the responses it directs but also may initiate the processes that evoke the memories, images, expectations, coping behaviours, and display rules relevant to the emotion” (Ekman, 1980:84).

The automatic appraisal mechanism must be quick because appraisal often happens without awareness. Because the appraisal mechanism is automatic, the affect programs are hard to control, and can be hard to fake. This is a source of what has been traditionally noted as the *passivity* of the emotions; emotions are passive in that they *happen to us* rather than being brought around by our agency.[26] Some of the recent work in evolutionary psychology explains this passivity, at least in some cases, as an adaptation. Steven Pinker argues that the passivity of anger and of our sense of fairness, allows us to enforce deals that only *pay-off* when they are considered one of a series of reiterated deals (Pinker, 1997:407-414).

The affect programs triggered by the automatic appraisal mechanism are complex in that they involve facial, vocal, and bodily expressive responses, autonomic nervous system arousal, musculo-skeletal changes, central nervous system arousal, and other behavioural and cognitive effects such as attention orientation and stimulation, which are in the “largest part [genetically] given, not acquired” (1980:82). Ekman’s studies have paid particular attention to facial expressions of the emotions (Ekman, 1972; Ekman and Friesen, 1975), something that also played a prominent role in Ekman's intellectual forebear, Darwin (Darwin, 1872). The dynamic of a particular emotion is in general concordant to the type of situation that elicits that emotion, in that it involves responses apt to aid in coping adaptively with those situations. Ekman believes affect is perhaps literally a set of *programs* in the brain that respond to environmental features and provide adaptive responses. Furthermore, they comprise an affective *system*. "For there to be such complexity and organization in various response systems, there must be some central direction. The term *affect program* refers to a mechanism that stores the patterns for these complex organized responses, and which when set off directs their occurrence" (Ekman, 1980:82).

---

[26] Sartre, however, thinks this false. He holds that emotions involve our attempting “magically” to alter our circumstances and to become a passive object in order to escape responsibility for our actions (Sartre, 1939). Like just about everything in Sartre’s view of life, emotions involve us in “bad faith.”

Affect programs are typically non-cognitive in Le Doux's "narrow" sense [2.2], playing out in a reaction-like manner. Though affective emotions may be elicited by sophisticated perceptual or cognitive information, once triggered the affect program runs a pre-determined course of physiological reactions.

Culture is allowed to play its desired role in the theory because of the distinction between elicitors and the affect program proper. The affect program is *given* yet the elicitors, the portions of the environment to which the automatic appraisal mechanism responds, may be learned and culturally variant. Ekman uses "the term elicitor to refer only to those stimuli that are identified by the automatic appraiser as specific for one or another emotion" (1980:85). The automatic appraisal mechanism that responds to the elicitors is not to be seen as a mere "stimulus-response reflex arc" because the features of the environment the mechanism responds to – the elicitors – will often be contextually complex, and "there is no emotion for which there is a universal elicitor, uniform in its specific details, which always calls forth the same uninterruptible set of emotional responses" (1980:85). Though the link between the appraisal of an emotionally relevant elicitor and the production of an affect program is biological fixed or given, the link between the elicitor and the appraisal mechanism may vary among cultural situations, not to mention personal situations.

Le Doux sketches a similar story, though he concentrates on fear, and his theory is a rather concrete description of features of the brain. Fear is elicited in an automatic fashion by the activation of brain networks in the amygdala. Like Ekman's theory, the elicitors of fear are classes of features in the environment, specifically those signalling danger in the environment. A fear reaction leads to the arousal of various neurophysiological and physiological systems, and then naturally to the behaviour the physiological arousal prepares the subject for. Also like Ekman, Le Doux does not see the elicitors as biologically fixed; they may sometimes be context sensitive, even highly cognitive states, since the amygdala is also influenced by signals from cortical regions where complex representations arise (1998:163-165). Yet the amygdala is biased to produce fear in a quick and dirty manner, with the cortex often merely modulating the response.

Le Doux was led to inquire into the amygdala as the neural basis of fear both because it had previously been established to have an influence on the various physiological reactions typical of affective fear, such as autonomic nervous system arousal, "reflex modulation," and freezing reactions (1998:158), and because it has direct links to the sensory systems that would be needed to appraise dangerous environmental stimuli, particularly to the auditory thalamus and other early sensory regions (1998:159). The activation of the portions of the amygdala important to fear involves what Le Doux calls the “low and the high road” in that projections both from the phyletically ancient and cognitively rudimentary thalamic areas, and from more modern and cognitively complex cortical regions, can influence fear reactions (1998:161). Though he admits the thalamic connections, which are strongly “biased towards provoking responses,” and therefore might be the source of false responses, might be vestiges from our evolutionary past, he contends they have in fact been retained because *quick and dirty* responses are ultimately adaptive:

> Imagine walking in the woods. A crackling sound occurs. It goes straight to the amygdala through the thalamic pathway. The sound also goes from the thalamus to the cortex, which recognizes the sound to be a dry twig that snapped under the weight of your boot, or that of a rattlesnake shaking its tail. But by the time the cortex has figured this out, the amygdala is already starting to defend against the snake. The information received from the thalamus is unfiltered and biased toward evoking responses. The cortex’s job is to prevent the inappropriate response rather than to produce the appropriate one. Alternatively, suppose there is a slender curved shape on the path. The curvature and the slenderness reach the amygdala from the thalamus, whereas only the cortex distinguishes a coiled up snake from a curved stick. If it is a snake, the amygdala is ahead of the game. From the point of view of survival, it is better to respond to potentially dangerous events as if they were in fact the real thing than fail to respond. The cost of treating a stick as a snake is less, in the long run, than the cost of treating a snake as a stick.
>
> (Le Doux, 1998:163-165)

However, nothing here immediately establishes that the exact environmental stimuli that elicit fear reactions are themselves *biologically fixed*: the amygdala might be equally prone to eliciting fear in cases of learned or conditioned elicitors. In fact this is likely the rule, as the knowledge structures genetically given to the organism consist of structures providing “learning preparedness” and not actual elicitor recognition (Le Doux, 1998:236-238). Rather than being innately wired to elicit fear in response to snake stimuli, the organism is wired with a disposition to learn that snake stimuli should elicit fear. Affect program theory need not see people as being innately pre-programmed to recognise all emotional elicitors. The historical source of treating the amygdala and limbic system as the seat of fear were traditional *conditioning studies* (Le Doux, 1998:158). As noted in Ekman, the affect program theorist need not claim the link between environment and affective reaction is given, even though the nature of the automatic appraisal mechanism and affect program are themselves biologically or genetically given (Ekman, 1980). Le Doux’s discussion of “contextual conditioning” experiments also bears out this important point (Le Doux, 1998:167).

Fear activation leads to a series of stereotyped behavioural and physiological reactions:

> When the brain detects danger, it also sends messages through the nerves of the autonomic nervous system to bodily organs and adjusts the activity of these organs to match the demands of the situation. Nerves reaching the gut, heart, blood vessels, and sweat and salivary glands give rise to the taut stomach, racing heart, high blood pressure, clammy hands and feet, and dry mouths that typify fear in humans. The cardiovascular responses associated with defensive behaviour have been examined in birds, rats, rabbits, cats, dogs, monkeys, baboons, and people, to name a few of the better studied species, and the responses are controlled by similar kinds of brain networks and body chemistry in these different species. Threatening stimuli also cause the pituitary gland to release adrenocorticotropic hormone (ACTH) that results in the release of a steroid hormone from the adrenal gland. The adrenal hormone then travels back to the brain… These various bodily responses are not random activities. They each play an important role in the emotional reaction and each functions similarly in diverse groups.
>
> (Le Doux, 1998:132-133)

Other reactions, more behavioural, such as attention orientation, *freezing* behaviours, bodily and facial expressions, also often follow on from the affective reactions. In general the physiological reactions, and behavioural adjuncts, are focused on providing adaptive responses to significant classes of environmental situations. Indeed, this is why the responses exist; fear is an adaptation. For Le Doux, the *feeling* of an emotion is the cognitive registering of the feedback of these physiological responses (1998:267-303). There is a strong Jamesian theme in Le Doux's work, though feeling does not provide the *identity conditions* of the emotions.

This Jamesian feedback theme is shared by Damasio, who gives a similar story to Le Doux when he discusses what he calls the "primary emotions" (Damasio, 1994, 1999). Primary emotions are "innate" and "preorganized" (1994:133), and involve the amygdala and anterior cingulate giving perceptual and behavioural salience to important aspects of the organism's environment. The environment includes the *inner* mental environment, in cases such as emotional responses to pain (1999:71-76). Damasio cites Le Doux's work, and has carried out similar studies. Emotions are based on a "biological core" (1999:51) that Damasio sums up in a list that is worth quoting in full here, as it gives a good sketch of what I call the *affective emotions* (though I have the minor reservations dealt with shortly):

> 1. Emotions are complicated collections of chemical and neural responses, forming a pattern; all emotions have some kind of regulatory role to play, leading in one way or another to the creation of circumstances advantageous to the organism exhibiting the

> phenomena; emotions are *about* the life of an organism, its body to be precise, and their role is to assist the organism in maintaining life.
>
> 2. Notwithstanding the reality that learning and culture alter the expression of emotions and give emotions new meanings, emotions are biologically determined processes, depending on innately set brain devices, laid down by a long evolutionary history.
> 3. The devices which produce emotions occupy a fairly restricted ensemble of subcortical regions, beginning at the level of the brain stem, and moving up to the higher brain; the devices are part of a set of structures that regulate and represent body states…
> 4. All the devices can be engaged automatically, without conscious deliberation; the considerable amount of individual variation and the fact that culture plays a role in shaping some inducers [cf. elicitors] does not deny the fundamental stereotypicity, automaticity, and regulatory purposes of the emotions.
> 5. All emotions use the body as their theater (internal milieu, visceral, vestibular and musculoskeletal systems), but emotions also affect the mode of operation of numerous brain circuits: the variety of the emotional responses is responsible for profound changes in both the body landscape and the brain landscape. The collection of these changes constitutes the substrate for the neural patterns which eventually become feelings of emotion.
>
> (Damasio, 1999:51-52)

However, given Damasio's talk of primary emotions, it comes as no surprise that he postulates “secondary emotions,” which, in his view, “occur once we begin experiencing feelings and forming *systematic connections between categories of objects and situations, on the one hand, and primary emotions, on the other*” (1994:134, emphasis in original). Secondary emotions are a mix of primary emotions and higher cognitive states representing the relationship of the primary emotions to certain classes of phenomena. In this move it is probable that Damasio intends to form a framework that will capture the full range of emotions. Griffiths notes and criticises the postulation of secondary emotions as paying the vernacular category “emotion” more respect than it deserves (1997:103). In a later work, Damasio notes Griffiths’ criticism, admits the point is a valid one, yet decides to persist in his usage of the commonsense concept “emotion” because “by maintaining some continuity we will facilitate communication at this transitional stage” (Damasio, 1999:341*n*9). Damasio argues that his use of the term “emotion” is not necessarily based on strong ontological claims, but is merely for theoretical convenience.

I suspect Damasio’s claims for a single emotional kind are based on his view that the concept of *organismic homeostasis* can be used as the *causal homeostatic mechanism* that binds the emotions together into a projectable psychological kind. This is based on his claims that the emotions are about *maintaining life*, a claim exemplified in the passage quoted earlier “emotions are *about* the life of an organism, its body to be precise, and their role is to assist the organism in maintaining life" (Damasio, 1994:51). I am doubtful such a *broad* notion could be used to form a category that comes close to satisfying Griffiths' conditions for being a *kind*, as almost all of our psychology is in some way concerned with *maintaining life!*

Despite these disputes there is much to take from Damasio's work, particularly from his observations about interaction and influence of affective states on higher cognition. To this subject I now turn. I have argued higher cognitive states will sometimes play a role in the elicitation of affective states, yet it should also be noted that affect has a significant influence on cognitive states. Much of the interest I will eventually have in the emotions and fiction will be on how affective emotion contributes to our cognitive participation with fictions.

In affect theory, narrow cognitive content, the type of cognitive representations central to propositional attitude theories, will play an eliciting role only in affective reactions, and on some occasions it simply will not be involved. Affect may often be elicited by perceptual and environmental features, some of which we may not even be aware of (Le Doux, 1998:53-64; Ekman, 1980:83). Representational cognitive content does not comprise the substance of the affective emotions. Le Doux concludes as much when he says "something else is needed to turn [such] cognitive appraisals into emotions" (1998:284). What does so in cases of fear essentially involves the amygdala and its role in giving perceptual and behavioural salience to certain features of the environment, to which various behavioural and physiological subsystems of the human mind and body can then respond with relevant behaviours.

However, complex mental representations, such as beliefs, will in humans form a significant set of the information that elicits affect, and it is this feature of cognition that makes it particularly adaptive. It extends the range of situations in which we can affectively react to the environment. This increased cognitive dimension of the emotions is signalled in Damasio's theory, where the difference between "background," "universal" or "primary," and "social" emotions, is an increased complexity in the cognitive component of the emotional episodes, and an increased specificity in *intentional object* that is afforded by this increasing cognitive complexity (1994:134-139). In the *social emotions* the cognitive content becomes a knowledge structure particular to a domain of human interaction, thus guiding our judgements, decisions, and actions in that domain.

The emotional life of a bird, a pukeko, say, is restricted relative to our species' affective life. Poking a pukeko with a stick will be enough to elicit affect fear in it. Telling it there was a mean looking hombre in the bar a few minutes earlier asking if anybody had seen it, is unlikely to make our pukeko frightened. Our cognitive and linguistic abilities significantly open up our world and our options for behaviour in it, because these abilities are able to track complex and novel environmental features. Yet our advanced cognitive phenotype will only be an advance on more primitive systems if it is clued in to our affective psychology, that is, if higher cognitive representations can elicit affective reactions. And it is no surprise that they can.

A set of empirical data illustrating this are the facts Le Doux draws on in noting that in humans the amygdala, the piece of the brain that is usually cited as the seat of the affective responses, has both very direct connections to the senses and also to very complex/abstract cognitive representations. I mentioned earlier that he thinks experimental results show there is both a “low and a high road” into the affective centres in the brain (1998:161). Signals may enter the amygdala directly from the sensory thalamus, and also via sensory cortical areas in which more complex representations of *objects* are likely formed. In terms of elicitors, Le Doux thinks: “The anatomical connections suggest that it [the amygdala] can be activated by simple features, whole objects, the context in which objects occur, semantic properties of objects, images and memories of objects, and the like. Any and all of these may therefore serve as the critical trigger information for emotional arousal” (1993:112).

The types of informational or representational state that elicit affect in the brain should also be relevant to the story of fictive participation I develop later. An important point made in various places in the literature on the fictional emotions is that states other than beliefs will elicit emotional reactions. “Representation tokens in a possible world box,” or “unasserted,” “imaginatively,” and “pretencefully held” thoughts, are proposed as the proper cause of our fictional emotions. The suggestion that our emotions for fictions are elicited by imaginatively held thoughts is physiologically plausible, as the passage from Le Doux makes clear. This should come as no surprise; it is a *folk-understanding* that the imagination and emotion are clued into each other. The case of the over-imaginative child fearing *monsters under the bed* should be enough to call to our attention the role the imagination often plays in exciting the emotions. Sexual fantasies are another obvious example of *affective reactions* caused by imagined states of affairs. I will have more to say on this general topic later, especially in a section where I discuss the recent work of Shaun Nichols and Stephen Stich on pretence [3.2].

It is established that higher cognitive states, such as contextual and semantic properties of mental representations, can *cause* affective reactions, and it is also coming to be established that affective reactions influence cognitive states. It is clear in all three theorists discussed earlier – Ekman, Le Doux, and Damasio – that not only do affect programs have physiological and phenomenal effects, but they also affect inference, perception, memory and other cognitive facets of the mind. Le Doux notes the amygdala has projections to the prefrontal cortex, the sensory cortex, and to the hippocampus. Via these projections, the “amygdala can influence ongoing perceptions, mental imagery, attention, short term memory, working memory, and long term memory, as well as the various higher order thought processes that make these possible” (Le Doux, 1998:287, fig.9-5). Affect not only influences the character of our higher cognition, but enables it to function by making the important parts of the environment experientially salient. It does so, because affective emotions are, in a sense, structures that contain *information* about what bits of the world should be attended to. Affective states are *frames* that map our interests onto our environment.

Damasio similarly contends that emotion plays a central role in “enhancing” the objects in our environment so that we might respond adaptively to our environment (Damasio, 1994, 1999).

Faced with a rich *decision space* in which we need to act, emotions not only focus our attention, but also help to bias the choice among options so that efficient choices might be made:

> What dominates the mind landscape once you are faced with a decision is the rich, broad display of knowledge about the situation that is being generated by its consideration. Images corresponding to myriad options for action and myriad possible outcomes are activated and keep being brought into focus. The language counterpart of those entities and scenes, the words that narrate what your mind sees and hears, is there too, vying for the spotlight. This process is based on the continuous creation of combinations of entities and events, resulting in a richly diverse juxtaposition of images which accords with previously categorised knowledge.
>
> (Damasio, 1994:196)

"Somatic markers" help to focus attention, to bias among the various representations that are attended to, and to guide the eventual decisions that are made. They are the feedback from the bodily states attendant to affect programs. Damasio proposes that "a somatic state, negative or positive, caused by the appearance of a given representation, operates not only as a *marker for the value of what is represented, but also as a booster for continued working memory and attention*" (1994:197-198, emphasis in original). Without such a biasing, rationality would be trapped by familiar *framing* problems into an apoplexy or an unfocused wandering. With too many options and not enough reason to choose among them cognition would be and ineffectual force.[27] Indeed, Damasio thinks damage to areas crucial in incorporating the effects of affect programs into higher cognition results in such an impairment of practical rationality (1994). Affective reactions also have a profound influence on memory, to some extent influencing what is explicitly remembered, and with what vibrancy (Le Doux, 1998:206-208).

One of the most important ways in which higher cognition responds to affective reactions is via explicit thought and its expression in language. We *report* our affective episodes, attempt to explain them, justify them, puzzle at their existence, and fit them into the wider picture we form of ourselves, what Damasio calls the "autobiographical self" (1999:172-176). The autobiographical self is a developing narrative of our place in the world, built from cognitive representations of our internal and external environments. I have already paid significant attention to such cognitive actions when I discussed folk-theory. I will have more to say of this class of cognitive states in a following section on higher cognitive emotion, because it seems to me that, because the emotion concepts and the language we use to express them are *dynamic* states, it will be productive to think of emotional language as substantive emotional behaviour, typical of some higher cognitive and social emotions. For the moment though, it will do to note that our affective states and feelings do not go unnoticed; as much is obvious from the most mundane of conversations with our friends and close acquaintances, in which talk of our feelings or emotions is ubiquitous.

---

[27] The options for action, it must be remembered, at any one time are so numerous they might as well be infinite in number. Yet only a miniscule few of the behavioural options will prove suitable to the situation at hand. Most will be entirely negligible or simply injurious.

It is a further notable fact that we also seem to have other emotions about and caused by our affective states; we often regret our emotions, are angry or sad because we have them, or are guilty or ashamed we express them as we do. Sometimes the elicitors of affective states will be the cognitive appraisal of *other* affective states. Affective states seem to resonate throughout cognition, sometimes causing other affective states, and sometimes leading to what are better thought of as higher cognitive emotions, as in the case of guilt or shame, which are unlikely to be easily characterised as affect programs. That the cognitive appraisal of an affective reaction might cause another affective reaction is not in the least mysterious: part of the environment to which cognitively complex minds respond is the *internal environment* of their own minds. I will have a little more to say about the role of affective reactions in higher cognitive emotions, and higher cognitive emotional reactions to affective states, in the following section on cognitive and social emotions.

The effects that affect has on cognition will be very important when I come to talk about fictive practice. Not only does affect provide the source of the episodic feelings we experience and report as appreciators of fictions, but it also directs our attention, understanding, and participation with works of fiction. I will argue that affect is involved in our judgments and interpretative actions concerning fictions because it *frames* our interests in their fictive content. In terms of directing our understanding, works of fiction often present incredibly rich *decision spaces*. Our emotions, and the manner in which they direct our cognition, might be seen to turn the rich decision spaces provided by fictions into coherent and satisfying *aesthetic experiences*.

I need also to re-characterise the role *behaviour* and *action* play in the affective emotions; this too may have implications for the story I tell about affect and fiction. Though a great deal of the behaviour kicked off by an affective response seems primarily physiological and reaction-like, much of it will stem from the effects of affect on higher cognition, and will involve planning and deliberate action. Like the cognitive information that often elicits affective responses, much of the non-physiological behaviour that results from an emotional episode is not part of the affective response, being instead a matter of how higher cognitive functions (both conscious and unconscious) react to the disturbances and the eliciting situation. Such are Ekman's "display rules" and "coping methods" (Ekman, 1980).

Ekman points out there are four stages to a typical affective reaction. The first is an appraisal state, where perception (and in various degrees ponderous higher cognition), locate and identify an emotionally interesting stimulus. Second comes the affect proper, which, once elicited, provides a "complex, co-ordinated, and automated" response to the eliciting state (Griffiths 1997:77). Among the things comprising the responses are behavioural manifestations involving musculo-skeletal changes such as flinching, orientation toward the stimulus, and other *functional* body and facial responses. Also involved are *expressive* facial and bodily reactions, which are essentially *adaptive* responses or vestiges of previous adaptations (Darwin, 1872; Ekman, 1972, 1980; Ekman and Friesen, 1975). For example, angry facial displays and postures in humans have often been phyletically linked to responses in other animals (usually dogs) that have a more obvious and physical attachment to the environment, most likely the bearing of teeth in preparation for fighting. These traits are "exaptations" or what Darwin calls "serviceable associated habits" (Darwin, 1872/1998:78, 90, 117-118; Le Doux, 1998:111). Ekman thinks these behavioural manifestations are biologically "given" and "organised" into complex arrays (1980:81-82).

Following these are the more cognitively led display rules and coping methods, the former comprised of social prescriptions for the emotional display and the latter comprised of our planned and longer term dealings with the emotional stimulus. The classic experiment that established the existence of display rules was carried out by Ekman (Ekman, 1980; Ekman and Friesen, 1975). Ekman and colleagues found Japanese students exposed to *stress films* (of the results of industrial accidents, apparently) *masked* their emotional expressions when in the presence of a person in authority (an observer in a white coat). American students, and the Japanese students when not in the presence of any observer, did not mask their emotional expressions with more neutral facial responses. When Ekman and his colleagues looked closer they found the Japanese students in the presence of the observer still had the normal expressions, but very quickly replaced these with the passive expressions. The masking was also found to be automatic and unconscious, and particular to the Japanese as a cultural group. Ekman also postulates the existence of "personal display rules" that vary with an individual's personality and social situation (1980:88).

Coping methods are comprised of the manner in which we behaviourally respond to affective states when drawing on highly cognitive processes like reasoning and planning. Ekman uses the term to refer to "attempts to deal with the emotion felt and its source; to increase, diminish, or sustain what is occurring. Coping includes various cognitive activities as well as such organised behaviours as attacking or fleeing. As a system where learning contributes the most, coping is the most elaborated of the emotional response systems" (1980:89).

Though I have never tried this, I would expect that if you were to elicit affect fear in a pukeko by poking it with a stick, it would do one of two things. It would likely either run away, or bite the stick (or you). The familiar term *flight or fight* picks out the typical course of fear behaviour in animals.[28] Some of the behaviour animals exhibit toward emotional stimuli is part of the short-term affect reaction proper. Le Doux suggests that, like facial expressions, attention orientation or freezing is one such feature (1998:149). Some animals (notably possums) react to threats by staring at the stimulus and freezing to the spot. By itself – and this is particularly the case in today's altered environments – such behaviour might not be particularly adaptive; stare too long at the bus that is bearing down on you and your genes will likely end up smeared all over the #035 to Grey Lynn, rather than in any of your progeny. But with the environmental salience your affect reaction has established, other higher cognitive mental systems will often give rise to more adaptive behaviours. Freezing reactions in animals are precursors to other adaptive behaviours that are mediated by more complex cognitive systems (Le Doux, 1998:149).

---

[28] It has been pointed out to me that this seems an essentially negative reading of emotion. Le Doux deals with fear specifically, but the ideas presented within his and Ekman's works *are* equally amenable to the positive emotions, particularly joy.

Cognition in humans has evolved to a considerable complexity, to the extent of allowing considerable forethought and, subsequently, *planning* of our behaviours. On seeing that the #035 to Grey Lynn is bearing down on you, your cognitive systems would suggest flight. But the fact that among our cognitive abilities is *counterfactual reasoning*, it is likely you have avoided the possibility of being hit by the #035 in the first place; it has been established that *looking left, looking right, and then looking left again* is a prudent measure when crossing roads, otherwise terrible things might happen. This ability to conceive of counterfactuals, and to plan future behaviours, will later be seen as strongly related to fictive abilities [3.2, 7.1].

Where pukekos and other simpler animals have a narrow range of information to which they will respond affectively, humans have more diverse ranges of such things, including *hypothetical* situations. Subsequently, our species has a vastly improved ability to cope with the challenges of the environment. Coping can emerge in many forms of behaviour in humans: affect fear elicited by air travel might result in the rapid removal of oneself from the aircraft; the fixation on other stimuli, such as a book, a meal, the in-flight movie, or the pursuit of inane chatter; ingesting various substances aimed at quelling the affect, such as a half bottle of scotch; or in the designing of a new high-speed train aimed at circumventing the need for air travel. Again, one of the behavioural responses we have to affect programs comes in the form of language as employed in self-report, explanation, and prediction. Telling people, especially those you are on intimate terms with, about your feelings, is a common response to emotion. One of the best ways of coping with an emotion, apparently, is getting it *off your chest* in burdening someone else with it.

It can seem coping methods provide an obvious bridge to the cognitive or social emotions; Damasio holds something like this view. His secondary emotions arise "once we begin experiencing feelings and *forming systematic connections between categories of objects and situations, on the one hand, and primary emotions on the other*" (1994:134, emphasis in original). I have already expressed my doubts that such a move could be made to substantiate a single projectable category "emotion," as Damasio seems to place too much faith in the scientific veracity of the vernacular term "emotion." His secondary emotions seem to consist merely of the manner in which higher cognition deals with the primary emotions, rather than some new kind of mental phenomenon distinguishable from either affect or higher cognition. It seems undeniable that affect does have an indispensable effect on higher cognition and subsequently on behaviour, however.

Affect theorists have recast the role of feeling in emotion. I believe this interesting discussion does shed some light over the issues particular to fiction and the emotions. To revise, emotional consciousness and feelings arise when the physiological disturbances of an affective episode enter the cognitive working space of consciousness and become experientially salient. In this they act to focus the individual on the eliciting stimulus, and are subsequently framed with reference to the appraisal of the stimulus. As such there is an element of the James-Lange feedback theory in both Le Doux's and Damasio's thoughts. Damasio in particular gives a complex story of how our awareness of our feelings of emotions arise when "second-order cognitive maps" are constructed of "first-order maps" of changes in musculo-skeletal and vestibular body states, and when a person thereby realises they are having a feeling (1999:280). It is to be remembered that the feeling is not meant to individuate the various emotion states [2.2]. It is not the cognitive appraisal of a physiological perturbation that picks out the nature of the affective episode, but the neurophysiological dynamic of the affective episode itself. As such, it is quite possible to have an emotion when one is not aware of the fact; emotions are not always felt as feelings (Damasio, 1999:36-38). It is even likely that most of the animals with affective emotions never get to feel them (Le Doux, 1998:300-302).

Feelings are a manner in which emotion affects cognition, and are also utilised in producing behaviour. After the effects of the affective episode have entered the cognitive working space, the feeling is attributed to a source, and then, following Ekman, coping methods are devised or borrowed from the cultural fund that has been established during the course of human evolution. Le Doux argues that the cognitive side of many emotions occurs after the fact, when the subject identifies and links the occurrent physiological changes in themselves with features of their current environment. One can thereby begin to become aware of *the good and the bad* in the environment. Feelings are causally potent states. They influence our behaviour in as much as we attempt to avoid them when they are negative and approach them when they are positive. Those animals that are lucky enough to have feelings are at a distinct advantage over those without feelings; they can envisage the value of their relationships to eliciting stimuli in more than a metaphorical sense. Feelings are also obviously involved in fictive practice.

ii.) Borderline cases: Sexual Arousal, Startle, Laughter

The revised conception of emotion categories provides no principled reason why I should not include the *borderline* cases of emotions in this study. Startle, sexual arousal, laughter, and other distinctly physiological perturbations that do not fit the vernacular conception of emotion need not be left out because they do not sit safely inside the now dubious categorisation. Further, because understanding these things might be useful in our discussion of fiction later, I should look at them; especially when they seem to turn up similar problems to those that are more familiar from the literature in the philosophy of art.

Startle, the reflex-like reaction to sudden loud noises, sudden movements in the visual field, and the like, has many similarities to emotions, moreover to affect programs, but also several dissimilarities. Jenefer Robison, concerned with expanding the emotional lexicon in the philosophy of the arts, has written an important paper that distinguishes startle from narrowly conceived cognitivist emotions, and argues for its role in the psychology of art (1995a). Startle is elicited by certain environmental features, in this case quite simple environmental features, and leads to a series of reactions, including bodily flinching, facial and vocal expressions, changes in heart rate and other autonomic nervous system arousal, and attention orientation. In these features, startle bears a similarity to the affect programs. Yet the differences – most importantly that startle occurs in a simple reflex arc, and as such is not context sensitive as affect programs are – have led some to conclude startle is not an emotion or affect program (Ekman, Friesen, and Simons, 1985). I think startle has often been exploited as an effect in fictions, and may be part of the reason why fiction appreciators are often so willing to report they were scared or surprised by fictive material. Therefore startle has an explanatory role to play in the paradox of fictional emotions, even though strictly it may not be an emotion (but again, my complaint with this vernacular category is registered). I will discuss startle and fiction later [5.3].

Erotic fiction, both written, and electronic, arouses what we might in polite language call emotions, though the physiological and bodily manifestations of these reactions is likely to prevent most philosophers from this usage. The word “lust” is emotion-like in its meaning, however. So I do see some evidence for including sexual arousal in the paradox of fiction as traditionally conceived, and it is strange it is not more frequently mentioned in the literature on the topic. I have chosen to include sexual arousal in this study because of similarities it has to fictional emotions. If I look quite closely at the case of erotic fiction, then the paradox of fiction might be cast in an informative light. There are many parallels between the elicitation of affective emotions by fictions and the elicitation of sexual arousal by fictions. There are also many parallels in some of the issues that arise from our cognitive involvement in this erotic relationship to fiction. Indeed, the experience of pornography fictions mirrors much of the experience of the works of fiction under discussion here, even in terms of the semantic features of our descriptions of it; the main difference being, I think, our inability to think the resultant sexual arousal anything other than genuine arousal.

Laughter is a quite similar case, and I have chosen to include it here for many of the same reasons as my discussion of sexual arousal. If we reflect that laughter is often caused by fictive narratives, in the form of jokes and humorous stories, we might begin to see the parallel with fictional emotions. What exactly are we laughing at? Is it *the bishop, the Irishman, and the pneumatic tyre tube* that can really be said to find funny? After all, like Leopold Bloom and Stephen Dedalus, *they do not really exist*. I will make some preliminary remarks on this issue later, but here I will note some things about laughter that makes it like other emotional reactions as well as other reactions such as yawning. Robert Provine’s scientific study on laughter, the imaginatively titled Laughter: A Scientific Study (Provine, 2000), makes it clear that he sees laughter as like an affect program. The typical cause and historical source of laughter is seen to be social interaction, specifically in the form of tickle and sex play.

> [The] vocalisation of laughter did not arise de novo, but originated in the ritualised panting of rough-and-tumble and sex play, whereby the sound of labored breathing came to symbolize the playful state that produced it. The vowellike “ha-ha-ha”s that parse the outward breath in modern human laughter is one step removed from the archetypal huffing and puffing that signalled laughter and play in our ancient ancestors.
>
> (Provine, 2000:96-97)

Laughter in its modern state is at one further remove from its origin, in that it often does not involve rough-and-tumble or sex play, but is elicited by more complex social situations, even by jokes and comedians. Yet laughter is quite similar to emotion, and specifically to affect programs, in that it is *elicited* by various classes of environmental stimuli that might be context sensitive (comparable to *elicitors*); by a neurophysiological mechanism that is automatic and out of our conscious control (an *automatic appraisal mechanism*); and in that it involves a set of physiological reactions, such as the laughter itself, and other expressive and physiological changes that coincide with the laughter (perhaps in the form of a *program of physiological and behavioural adjuncts*). These features, and the fact laughter is often elicited in conjunction with fictions, mean it is likely that *fictional laughter* can be explained in the manner I propose the *fictional emotions* are.

In the second half of this thesis, I will introduce startle, sexual arousal, laughter, and some other *affective* states into my discussion of fiction and the affective emotions [5.1-5.5]. My remarks will necessarily be provisional and limited, but I hope they will add weight to my overall characterisation of fictive practice and the emotions, and will indicate how I think the study should develop.

iii.) Higher Cognitive and Social Emotions

A second promising line of inquiry exists in the heterogeneous literatures on the *sociality* or on the *cultural* nature of the emotions. Griffiths thinks several of the modern approaches to the emotions, particularly the cognitivist and social constructionist schools, and much *sociobiological* and other work with an evolutionary focus, are primarily concerned with the higher cognitive emotions. He is critical of the psychological ground these phenomena cover, finding they need to be carefully distinguished from stereotyped *short-term* states such as affect fear and surprise. I will speak of the *cognitive* or *social* emotions to characterise this approach. Romantic love, guilt, embarrassment, and shame would be typical instances of cognitive or social emotions.

This is not to say however, that short-term affective responses are not involved in social behaviour. It seems likely that affective emotions are often involved in social transactions, are often elicited by social cues, and have a role to play in the higher cognitive and social emotions I identify here. The cognitive or social emotions, like the secondary emotions Damasio characterises, might in fact amount to a mixed bag of affective and higher cognitive features. My later discussion of fiction and the emotions will trade on this division in its expository schema, however.

Furthermore, it is strictly misleading to call only these social emotions “cognitive emotions,” given affective emotions might equally be given that label. Griffiths (1997:24-27) thinks it a misnomer to call propositional attitude theory “cognitivism”; it implies the competing affective theory is non-cognitive. This is only true so far as cognition is conceived as the quasi-logical manipulation of mental representations, such as beliefs and desires. As I found earlier [2.2], in one sense all that the brain does is cognitive, but modern cognitive science (and its running mate, folk-psychology) gives “cognition” a particular emphasis.

First I will make a preliminary sketch of some of the features of the cognitive or social emotions, and how they are likely to differ from affective emotions. These differences will shape how social or higher cognitive emotions are involved in fictive participation. Affect programs are short-term stereotypical responses, elicited by perceptual cues and other representational states, that play out in an automatic fashion. Many of the episodes that are commonly characterised as emotions seem quite unlike these states. Many emotional states endure over long periods, contain no significant physiological arousal, and do not seem to be tied to cues in the perceptual environment in the manner affect programs are. Love, for example, may endure for many years, often in the absence of the salient physiological effects notable in affect programs. In the case of the social emotions, such as romantic love, shame, and guilt, there seems to be a significant cultural component to the emotions, as they are often deeply ingrained in our traditions, institutions, and cultural norms. What a person does when they are in love is guided by the prescriptions of the culture in which they live. Furthermore, some of these emotions have significant temporal, societal, and geographical variation, something unlikely to be strictly consistent with affect program theory.

Affect is consistent in some part with societal variation, but not *absolutely* consistent. Cultural (and even personal) variation in terms of elicitors, display rules, and coping methods, are a strong point of affect program theory. Griffiths has also argued against the assumption in the evolutionary psychology school that the human mind is “monomorphic,” that is, in every normal individual invariable (1997:122-125). Yet the depth of variation in some apparent emotions is unlikely to be reconciled with the view of emotions as mental traits established in our distant evolutionary past. Some emotions seem to be extremely local cultural effects, dependent on particular institutions. One example is *amae*: the deep respect Japanese people often exhibit for people in positions of authority. Another possible example is the *divine joy* often seen in Pentecostal churches in the USA, the symptoms of which include uncontrollable laughter and “speaking in tongues” (Provine, 2000:134). Indeed, religious practices give rise to many odd putative emotional states that threaten to evade explanation as *physical states*, let alone as affective states!

Though affect is to be strictly identified with the underlying neurophysiological systems, it seems unlikely romantic love, amae, divine joy, or guilt could be characterised successfully in this manner. In these we find a greater social complexity and a measure of social variation not found in short-term affective reactions. What seems important to these cognitively complex emotions are the types of beliefs, behaviours, and social structures that surround their instances. Attempting to define the source of the physiology and feelings involved in these complex states would tell only a very partial story. Social construction theorists seem correct in pointing out the importance of emotion concept usage to the nature of emotion, since with many social emotions subjects must be able to understand and deploy the concepts central to the emotion (see especially Harré 1986).

Though we can assume many other animals have affective responses, it is not clear they show sufficient cognitive and social sophistication to allow them to have higher cognitive emotions, or at least *the type of cognitive or social emotions we have*. Though they have affective emotions, it may be that animals besides humans do not fall in love in any interesting sense. It is very much an empirical question whether chimps, whales, lizards, or pukekos have romantic love, and it is a very interesting issue. These animals all have mating drives, and exhibit mating rituals. It is a possibility that, like the affective emotions, the traits underlying these are homologues of the traits underlying our own mating drives and rituals. If so, then these animals may fall in love in exactly the sense we fall in love. If we have a strongly social constructionist theory of love, however, we are less likely to see other species as having love unless the have very similar social structures. Given the importance of language in our social structures, this is unlikely. Also, love in different species is likely to be quite different, given the wide-ranging variables pertaining to mating in other species. If the eusocial naked mole rats fall in love, it would be a perverse form of love by human standards! Though it is romantic to think of other animals falling in love with each other – the sentiment is often expressed in the schmaltzy pictures on greeting cards – to positively conclude this way, I believe we would have to either downgrade the cognitive and cultural complexity of our own romantic states, or credit other animals with a cognitive and cultural complexity they may not have.

Cognition plays a role in affective reactions, both in eliciting them and in guiding our responses to them. Part of this cognitive involvement in affect will owe its existence to *a cultural fund of knowledge* concerning what environmental situations are emotionally relevant and how to behave in such situations. Yet the affect, though elicited in conjunction with this cultural knowledge, remains distinct from it. Ekman believes many emotional reactions may be “ponderous” in that they are not due to the automatic appraisal mechanisms associated with affect programs (1980:84). The reactions are ponderous because they involve a highly cognitive appraisal of an environmental situation, and the behavioural reaction to the situation is *less stereotyped* than the affective reaction. Ekman is a little resistant to calling these emotional reactions (1980:84).

But this I believe is where the cognitive or social emotions might begin to establish themselves. As our cognitive sophistication has increased, the cognitive component of our affective life has increased. With the transmission of customs, traditions, and the other facets of *culture*, what has been traditionally seen as our emotional natures is bound to become *cognitively and behaviourally complex*. Affective reactions *are* often likely to play a large part in cognitive emotional occurrences such as romantic love and guilt. Instances of anger, sadness, and joy (three of Ekman’s six typical affects) may play a developmental or even substantive role in their more cognitive counterparts (Griffiths 1997:121-122). But it is not necessary that they do so; some emotions will entirely lack an affective component.

Damasio thinks the social emotions combine affective states with a complex cognitive component (1999:342n). Damasio argues that, with an increased cognitive component in affective episodes, emotional episodes take on an *increasing specificity in object* (1994:136). This is consistent with what is being held here; that affect and narrow cognitivism are essentially different and yet will *combine* in the complex arrays of mental attitudes that constitute modern human psychological activity. Damasio thinks we can in principle sketch a theory of the emotions that includes all the emotions within a single framework. Yet I need not accede to this feature of Damasio's theory to find use in his observations. I will not want to assent to this aspect of his theory because it is likely there are emotions that do not at all involve the physiological arousal he and others think typical of primary or affective emotions.

The last of the four stages in a typical affect program response (of which only the second is the affect proper) involves the cognitive and rather more conscious part of the mind directing behaviour toward the object of the affective emotional response (Ekman 1980:87-90). Given that the objects of elicitation have now become more cognitively complex, we might also expect that the methods for coping with the objects will be correspondingly more complex. Standing, frozen to the spot, and staring at the punk who has just insulted your masculinity in front of your girlfriend is unlikely to be successful. Frequently encountered situations give rise to standard lines of reaction. Certain of the standard lines will be more successful than others and will more strongly establish themselves in the canon of human behaviours. Some will develop social or institutional aspects, and so it is that we find romantic love is integrated with the social institutions of marriage (and, more so of late, divorce). Since a great deal of life involves our interactions with other people, and these dealings determine much of our success and failure in the world, it is obvious such interactions will furnish affective reactions as well as the subsequent standard lines of cognitive and planned response. Talk and self-report, being essential aspects of how we comport ourselves toward others, are important parts of cognitive emotional substance. Talk is behaviour, and so, when we say we love or hate someone, these reports (unlike in the case of affective reactions) are very much part of the social emotional substance.[29]

In what follows I will briefly discuss some of the approaches that have been made to explaining these states. My discussion of the higher cognitive or social emotions involves less detail and less length than I spent on the affective emotions for two reasons: first, in the emotion literature, the affect programs are more concretely described and the research is further advanced; the cognitive emotions are still often discussed in *functional* and *task description* terms, because the theory of the cognitive emotions is far less developed than the many programs of research into the affect programs. Cognitive and social emotions might be typically seen as informational structures aimed at allowing us to track certain domains such as mating, cooperation, fighting, and status. Second, my later theory makes more of the involvement of affect in fiction, and where it discusses the higher cognitive emotions the problems are often *specific to cognitive attitudes*, such as those problems that are the target or pretence theories of fictive practice. That the more interesting problems concern affective reactions is, I think, to be expected, as the opportunities to engage in *full-blown social emotional behaviour* with fictions is limited. Though people on occasion might be tempted to *describe* themselves as in love with a fictional character, it would be difficult for them to satisfy us they were really in love with the character [6.3]. Anyway, it will be enough here to note that cognitive and social emotions are cognitively complex, that they may often involve an affective component, that they involve us in complex, often long-term, relationships and behaviour, and that they are the subjects of mental ascriptions and reports that themselves play an important role in the behavioural nexus that constitutes the emotional episode.

---

[29] But as I hope I have shown earlier in this section this does not mean vernacular emotion talk is to be seen as *transparent* in its emotional descriptions.

A promising literature on the cognitive or social emotions from an evolutionary standpoint has been developing for some time. Though this evolutionary framework is equally applicable to the primarily affective emotions – the affect program theory can be seen as a development of Darwin’s work on the expression of the emotions – it more often has a concern with how the individual *behaves* toward others, reflecting its objective to understand *social adaptations*. In the 1970s, the sociobiology of E.O. Wilson and others pressed the thesis that many behaviours were adaptations to conditions in the environments of our ancestors (Wilson, 1975). Sociobiology was the result of the developing theory and methodology in evolutionary theory: principally work that made it clear that cooperative actions not necessarily beneficial to the individual organism – the main example being altruism – might be explained as the result of a “kin selection” driven by the *interests* of the “replicator” at the base of the traits (Dawkins, 1976/1989; Williams, 1966). Much of this work was the direct result of the application of game theory to evolutionary biology (Dawkins, 1976/1989; Maynard Smith, 1964, 1976, 1982; Williams, 1966). Sociobiology was forcefully opposed by many theorists and does seem to make claims about behavioural adaptations that are quite problematic, foremost because it is difficult to determine whether a present behaviour *is* an adaptation or merely some sort of local effect of environment on the established psychological phenotype (Sterelny and Griffiths, 1999:314-315). Evolutionary psychology seeks to avoid this difficulty by postulating the psychological and physiological *mechanisms* that subserve the current behaviours, as the adaptations to our “environment of evolutionary adaptedness” (Tooby and Cosmides, 1992:69). Evolutionary psychology, understandably then, has a view of the mind as a collection of discrete or *modular* mental traits. This is consistent with affect program theory, though evolutionary psychologists seem more likely to be interested in the mechanisms underlying complex behavioural and social traits. Henry Plotkin has written an accessible introduction to the evolutionary psychology program for those daunted by the book by Barkow, Cosmides, and Tooby (Plotkin, 1997).

Emotions are central to such complex social and behavioural activities, and it is no surprise to find evolutionary psychologists have made many claims about the nature of the emotions. In evolutionary psychology, emotions are seen as “intricately structured information-sensitive regulatory adaptation[s] …the emotions appear to be designed to solve a certain category of regulatory problem that inevitably emerges in a mind full of disparate, functionally specialized mechanisms – the problem of coordinating the menagerie of mechanisms with each other and with the situation being faced” (Tooby and Cosmides 1992:99). Such a categorisation mostly corresponds to the vernacular conception, as is shown by the mechanisms Tooby and Cosmides think “emotionally and motivationally” significant: “parenting, emotional communication with infants and adults, kinship, mate choice, sexual attraction, aggression, the avoidance of danger, mate guarding, effort allocation in child care” (1992:99). Yet it is unlikely the authors think this category of states forms anything like a projectable psychological kinds.

Steven Pinker is at the popular forefront of the evolutionary psychology program and, though he has written primarily on language, he has also expressed his concern with cognitive and social emotions, and even with *aesthetic emotion* (1997). Emotions are informationally or content rich psychological states that evolved as adaptations to problems prevalent in our evolutionary past. Pinker’s discussion of love provides the flavour of his ideas on the social emotions. As with many of the evolutionary explanations of our emotions, Pinker’s discussion of love might leave the more sensitive reader a little cold:

> Marriage works a bit like leases, but our ancestors had to find some way to commit themselves before the laws existed. How can you be sure that a prospective partner won’t leave the minute it is rational to do so – say, when a 10-out-of-10 moves in next door? One answer is, don’t accept a partner who wanted you for rational reasons to begin with; look for a partner who is committed to staying with you because you are you. Committed by what? Committed by an emotion. An emotion that the person did not decide to have. An emotion that was not triggered by your objective mate-value and so will not be alienated by someone with a greater mate-value. An emotion that is guaranteed not to be a sham because it has physiological costs like tachycardia, insomnia, and anorexia. An emotion like romantic love.
>
> (Pinker, 1997:418)

Pinker’s theory involves typical evolutionary ideas, particularly in its similarity to discussions of altruism. Because love, like other social and cooperative behaviours such as altruism, is open to exploitation by cheats who will attempt to take the benefits of sexual access without paying costs of supporting the progeny long term, a mental mechanism has evolved that secures commitment to the cooperative relationship. We are *passive* to the emotion because the cooperative relationship of romantic love works only if it is something that happens to us. Pinker thinks similar explanations could be given of jealousy, shame, guilt, friendship, anger, sympathy, embarrassment, and the sense of justice or fairness.

Robert Frank is another who uses evolutionary ideas to formulate a theory of the social emotions (1988). Frank thinks that emotions are often states that, though irrational in the short term because immediately costly, pay off in the long run. Angrily refusing a deal because you feel the deal is not fair, even though you stand to profit from the deal, is one familiar example of an emotional behaviour that seems quite irrational. Because anger, love, and their results, revenge and altruism, can often be costly to the individual, they can seem irrational in the short-term. Yet when situated within a population of individuals who must often cooperate, the emotions are a source of adaptive behaviour.

Much of this work depends on the development of the methodology of game theory within evolutionary psychology. Here, actions that can be immediately disastrous for the individual can form the basis of an “evolutionary stable strategy” in a population of organisms (Sterelny and Griffiths, 1999:237-238). In game theory, refusing unfair trades when the immediate outcome is an *absolute gain* becomes an adaptive strategy when trading situations form a series of *iterated games*. The immediate loss is outweighed by future gains. Therefore, a sense of fairness, or a willingness to be vengeful and angry, though on occasions immediately unprofitable, overall are a good idea because trading partners can expect the outcome of past trading situations to have a bearing on future ones. Emotions are evolutionarily stable strategies in dealing with the social world, once again, because we are *passive* to them.

Pinker’s discussion of the “doomsday machine” of our sense of fairness, of revenge, and of anger is very similar to these ideas (Pinker, 1997:407-416). Social anger, Pinker thinks, is like an emotional bomb everyone has, and that is not under their control, but wired directly to the type of situations where anger is of use; if you wrong me, then I might just go off. The features of the doomsday machine, especially the *passivity* it derives from its modularity to higher cognition, its short-term stereotyped nature, and it qualitative dimension, make it easily characterised as an affective emotion. Affect may have an immediate role to play in this social emotion, given the non-cognitive nature of the affect program response, something that makes affect mandatory, difficult to fake, and difficult to control.

Understandably, for some theorists, these ideas are still going to be far too *naïve* and biologically reductionist. Love is not just a *propensity to form irrational attachments to sexual partners because of the expected fitness utility that such an attachment will have for one’s genes*. Culture must be given more importance. What is important in love, it will be held, are the sorts of customs we engage in. The genetic propensity to form *irrational bonds* tells us nothing about courting and dating, giving flowers, love letters, and the other trappings particular to love in Western society (or at the least the love particular to American television shows). Furthermore, the biological story does not seem to explain the cultural and personal variation many social scientists have seen as inherent in emotional behaviour. For many anthropologists and social scientists, human variation is ubiquitous. Margaret Mead’s studies of Samoan communities are seminal in this regard.[30] There does seem to be a tension between the biological and cultural aspects of the social and cultural emotions.

There are many interesting things to be said about this, but short of developing *a theory of culture* that would explain how the traditions that affix to our biological core arise and develop, I will be limited in what I can discuss. Further, it is not clear I will need such a complex theory of the developmental source of the social and cognitive emotions; I merely need *enough* to be getting on with the process of characterising their place in fiction. It will suffice to produce a rough sketch of the dynamics of the cultural and social emotions, where several features, those that assume importance in our fictive dealings, have been picked out. I already have these features before me. Of particular importance are the highly cognitive nature of the emotions, the involvement of affect in the cognitive and social emotions, their social and behavioural dimensions (that is, that they involve us in relationships), and (as I will shortly argue) the fact that we have complex theories of these emotions, and express these theories in our explanation, prediction, and characterisation of ourselves and others.

However, evolutionary psychology does have a model of culture that is not without merit, and that can account for the variation in emotional behaviour and customs without ejecting the core of the evolutionary characterisation of the emotions. One of its particular merits is that it does much to undermine the putative nature/nurture divide that seems to place a wedge between the culture and biology of the emotions.[31] Evolutionary theory does not claim that an organism receives all the information it uses to develop solely from its genes. Organismic development or ontogeny is fundamentally underdetermined by genetic information. This is a fact reflected at all levels of development, from the cellular to the behavioural (Sterelny and Griffiths, 1999). Organisms, especially complex ones such as ourselves, need social and even cultural input, that is, sources of the environmental information to which we are adapted, if we are to develop successfully. Indeed, some measure of social constructionism is absolutely consistent with the view of human nature as having a biological core.[32] Many evolutionary psychologists even argue that *culture is inherently biological*. I do not have the space to develop this argument here, but for further discussion see the books by Griffiths and Sterelny (Griffiths, 1997; Sterelny and Griffiths, 1999), and the seminal collection of essays edited by Barkow, Cosmides, and Tooby (1992).

---

[30] Though anyone who has dipped into the evolutionary psychology literature will be well aware of the attitude to Margaret Mead’s work that is often displayed there (Pinker, 1997:368-369; Tooby Cosmides, 1992:44).

[31] Griffiths argues it does not do enough (1997).

[32] Importantly, this social construction is a benign form of constructivism; it is nothing so adventurous as what we might find in Nelson Goodman’s work (1978) or in the work of others, where even stars and planets can be socially constructed (Devitt, 1991:233-258). This benign form of social constructionism tends to apply only to the socially constructed parts of reality!

In addition to social and cognitive emotional behaviours, *whatever their actual nature*, we form beliefs and theories about the motives and likely actions of other people and ourselves, we report these, and we also discuss them with others. This ability is central to our success as social beings. It will also provide one of the more important ways in which the cognitive and social emotions are involved in fictions. One of the most salient of the cognitive responses to the affect programs is theorisation and its expression in the language of self-report, prediction, and explanation. Theorisation and language plays a part in our *characterising* of our emotional life. Where the concepts play a causal role in the emotional behaviour, theorisation and self-report will play a role in *implementing* our social emotional lives.

We must then pay close attention to the current theory theory/simulation theory debate, about which I will have much more to say in the section on the higher cognitive emotions and fiction [6.2]. Briefly though, the debate concerns how it is we manage to generate beliefs about the mental states of those around us, and further, to extrapolate these mental states into predictions of probable behaviours (Davies and Stone, 1995a, 1995b). Most of us are not good at making the mental calculations needed to judge where to hit a cue ball to have the desired effect of a snooker ball dropping into a pocket. Human beings are objects that move around the world in ways far more complex than the manner in which snooker balls move around a table. Yet it is evident that my flatmate would be annoyed if I did something as innocuous as emptying half a bottle of vodka into her fish tank. The debate between theory theorists and simulation theorists concerns the best way of explaining why we seem to have such a natural access to understanding others.

Simulation theorists believe that in explaining and predicting the behaviour of those around us we run *off-line simulations* of their situations. We plug their situations into our own emotional and behavioural guidance systems, of which the behavioural adjuncts are taken off-line, for obvious reasons. The subsequent cognitive states are used to explain and predict the behaviours and mental states of the subject under question. The cognitive states are suitable to form the basis of predictive explanation as they presumably share a certain amount of *structural or informational identity* with the states we are attempting to predict. This is because the cognitive states are generated by our own *standard factory fitted human mind*.

Theory theorists see the situation as being quite different. Rather than running off-line simulations of the cognitive/affective states of others, what we do is *theorise* about the situations of others. This use of the term “theory” to describe a rather fundamental human psychological practice is bound to seem to some deceptive. Theory theorists argue that evolution and cultural history has provided us with a set of *knowledge structures* of use in explaining and predicting the behaviour of others. This theory has been minimally formalised into the folk-theory that was discussed previously [2.3].

However, recently the disagreement between simulation and theory theorists has been significantly undermined (Stich and Nichols 1997). Stich and Nichols think, "“simulation” is a hopelessly heterogeneous category of no theoretical utility. Thus there is little point in debating whether or not simulation theory is true" (1997). They think what simulation theorists see as simulation amounts to a collection of different things, some uncontroversial, some less so, and that many of the processes discussed are necessarily *hybrids* of the two theories. The debate is ill-formed because simulation is ill-defined, and Stich and Nichols, themselves originally defenders of theory theory, find themselves comfortable about being called “simulation theorists” in at least some of that terms’ vague senses. Despite this complex debate, it remains the case that human psychology has provision for some pretty spectacular predictive and explanatory actions on the parts of normal people. As will be argued later, whatever form(s) this ability is eventually explained as having, it makes up a significant part of our emotional involvement with fictions.

Some have claimed this ability to be absolutely central to fictive practice (Currie 1995a, 1996). It is perhaps thought that to emotionally engage with a fiction we need to understand the fictional emotional lives of the characters. Though this is important to the emotional experience of fictions, I have doubts as to how *central* it is. Indeed I have doubts that there is any one process that is central to fictive practice. Rather I think there are collections of attitudes that tempt us to describe fiction appreciators as *emotionally involved* in fictions. Some of these, such as affective attitudes, may not involve the simulation or complex mental calculus thought to be involved in understanding others’ psychological states.

One more note in this regard. Concepts of the types we use to explain the behaviour of those around us are not simply passive or *descriptive* states. Unlike the affective emotions, where concepts and language are primarily involved in response to the physiological episodes, concepts and language in cognitive and social emotions can be fully-fledged social *causal* states. Our relationships with others are often achieved through the expression of emotion concepts in language. Any successful theory of the social and cognitive emotions will have to take this complication into account. While the report of a feeling may not always be a good indicator of the presence of an affect program reaction, the report of a *lovin’ feelin’* may well be indicative that some type of emotional episode is occurring because the employment of the concept in a social transaction *amounts to an emotional episode*. The fact will be important to my argument here, because people often employ emotional concepts in response to fictions. If we are interested in the nature of what they are up to, then we should have an interest in how concepts function as behaviour. Needless to say, this is a complex subject.

Much of the work on the cognitive and social emotions is far less developed than the advances found in affect program theory, where the actual neurophysiological systems are being discovered, described, and mapped out. The work on the social and cognitive emotions seems confined to providing abstract task descriptions of the social and mental mechanisms that will eventually be discovered. It does provide hints and important constraints on theories that need the emotions to play a productive role, however. For the issue under discussion here, fiction and the emotions, these cognitive and social emotional theoretical posits suggest a source of the kinds of interests appreciators will have in fictions. *Social and cognitive emotions depend on rich informational structures specific to areas of adaptive interest to our social and highly cognitive species.* We should not be surprised later if we find these interests and the structures that subserve them involved in fictive participation, especially when the substance of these fictions concerns people and their relationships with one another. Our interpretative actions in fictive participation are likely to draw on the information structures that play a central role in our cognitive and social emotions.

***

Some might have lingering doubts about the fact that the real nature of emotion eludes us. Love may not be the state many think it is, and that is reflected in their observations about love, but may instead be an adaptation based on important features of evolution and population dynamics. Many will doubt we can get things this wrong. However, not only are we relatively ignorant of our internal psychological natures [2.2-2.4], but some of the structures in society that seem to guide or reinforce many of our behaviours may also be opaque to our observations; we again have reason to question our epistemic access to our own minds and behaviours. Disclaimed action emotions (Griffiths 1997:139-141) are a striking example of this. The Gururumba people of Papua New Guinea[33] have a behaviour called *being a wild pig* in which members of the society (usually stressed-out young men with the newly inherited financial burden of a wife and children) commit minor assaults and thefts. The peoples' explanation of these actions is that the individual has been bitten by a ghost and turned into a wild pig. In actuality, the behaviour is a communicative outlet for the stressed individuals, since after the behaviour has been shown, the community often re-appraises the amount the individual must commit to the society. Being a wild pig is an adaptive emotional behaviour.

Two levels of explanation exist concerning what it is to be a wild pig. There is a *manifest* level of explanation, and a *latent* level. The former introduces ghost and spirit possession, corresponding to the folk-theory I have been criticising. The second explains being a wild pig as an adaptive and socially constructed behaviour. To dissect the language of the former explanation would provide us with a wildly inaccurate theory of the behaviour. The latter explanation is furnished by an empirical study not only of the language use, but also of the behavioural, social, and internal psychological features of being a wild pig. If I am after an interesting explanation (in the scientific, rather than *amusing* sense) of being a wild pig, then the latter is certainly the method I should pursue.

---

[33] Le Doux mistakenly thinks this tribe inhabits a highland market gardening area in New Zealand. If they do they are a secretive lot; I definitely have not met any *wild pigs* at my local cafe!

There are various criticisms that might be raised concerning these ideas. Firstly, to claim we might question our epistemic access to our own minds and behaviours, might begin to make you wonder how I subsequently come out with such conclusions as "In actuality, the behaviour [being a wild pig] is basically a communicative outlet for the stressed individuals, as after the behaviour has been shown, the community often re-appraises the amount the individual must commit to the society." Are not such conclusions themselves subject to the doubt we might have in our epistemic access to our own minds and behaviours? Not at all, the types of epistemic access are different. The former type (folk-theory) is empirical only insofar as it is introspective and based on casual or tacit observance of the environment. The latter involves a fully-fledged scientific project informed by the type of empirical method science (and philosophy, as here) has established as useful. One of these points of empirical method, prevalent in Le Doux's theory, and certainly obvious since the behaviourists picked up on it, is that mental introspection and anecdote are dubious empirical methods. Another is that semantic field analysis or conceptual analysis will tell us only what people believe, or the manner in which they use language. Should such points of empirical method cast doubt on the theoretical posits of folk-theory (as they do in the case of the disclaimed action emotions), then so much the worse for folk-theory; this is one of the central points in Paul Churchland's "Eliminative Materialism and Propositional Attitudes" (1981).[34]

Second, one might propose that in such cases as *being a wild pig*, *the natives know perfectly well what the natives are doing*. In assuming otherwise, sociologists, anthropologists, and philosophers, are just being patronising to native populations. It is really an empirical matter whether the natives really know what the natives are doing, so I guess there is nothing I can do to counter this objection except to assume the sociologists and anthropologists are doing their jobs correctly.

But there is something I might say to this dissenter that fits very much with their objection. Not only are the sociologists and anthropologists being patronising, but so are the people who raise this objection! This is because not only do the natives not know what the natives are up to, but also *neither do we*. You only have to watch an episode of *The Oprah Winfrey Show* and observe the new brand of spiritual self-help psychology peddled there to discover the *truly odd* beliefs *ordinary* people can have about the workings of the mind. Griffiths' notion of disclaimed action emotions, and the gap between a person's *explanation* of what they are doing and what they are actually doing, pertains as much to modern Western behaviour as to *putatively Stone Age* behaviour. For example, modern religious institutions cannot assume their existence is based on the *manifest* or *explicit* principles that their practitioners hold. We are no different to the natives, and it is a real possibility our folk-theories are just as bad as theirs. The natives as a whole may not know what they are doing, because wherever you are, *the natives are you*.

---

[34] In the case of wild pig behaviour, as with Griffiths' *disclaimed action emotions* generally, the strong version of mental elimination, that is the *entity eliminativism* advocated by eliminativists such as the Churchlands, *is* appropriate for the theoretical entities used in the theory.

I believe that these two levels of explanation can be found across all the various emotions, affective and social/cognitive. At one level, romantic love is something special or spiritual that goes on between two people, often involving the notion of *soul mates*. On an interesting level, it might be seen as a state or process generated by a heterogeneous set of cultural and biological structures, that arose because it is good at providing children with a stable growing environment, or securing prolonged sexual access to a mate. The latter explanation – which will include descriptions of the type of concept and language use, and beliefs of the participants, *as well as* the biological core of the state – is the explanation we should be aiming at. It will not be furnished solely by conceptual analysis or other methods that do not subject the postulates of folk-psychology to the doubts we subject the postulates of explicit science to. Other methods, among them Le Doux's and Damasio's neurophysiology and an empirical psychology that sees emotions dependent on rich knowledge structures established in our evolutionary past, will be needed if we are to find out how such things are implemented in the mind.

This brings me to my conclusion to this chapter. It is unlikely that the vernacular usage of the concept "emotion" is going to go away. Indeed I find it difficult here not to use it, as it still refers to a real but heterogeneous collection of mental and cultural material. But I do think we should be more careful when we use it in other than informal contexts, especially in philosophical and scientific circumstances. I will argue later that inconsistency in usage of emotion terms, which equivocates across, among other things, the affect/social emotion categories noted here, adds to the philosophical confusion in the debate concerning fiction and the emotions [4.3]. Distinguishing between affective states and cognitive social emotions can only be a good move. Such a revision of emotion theory is likely to deny the priority of narrow cognitivism of the emotions in the characterisation of fictive practice; *a fortiori*, it will discount the *exclusive* use by aesthetic theorists such as Walton of the type of cognitive emotion theory that is currently prevalent in philosophical circles. The division that has opened in emotion theory will correspond to a division in my explanation of the role of emotion in fictive participation.

If the mind is an array of many special purpose mechanisms, as evolutionary psychologists tell us, then it is the state of that array I should be looking at when I come to characterise fictive practice. The consciousness derived from the higher cognitive abilities we gained somewhere along the evolutionary line will play an integral part in the wielding of that array in the complex social situations we now find ourselves in. Damasio, and the evolutionary psychologists think emotion is a central feature of this array of mental states (Damasio, 1994, 1999; Tooby and Cosmides, 1992). Our mentality is pervaded by emotion because of the effects of affect on enabling higher cognition to function. This conclusion is based on empirical findings; the brain structures central to successful rational and social functioning often seem to be the same as the structures central to emotion

(Damasio, 1994). Furthermore, deficits in emotion lead to deficits in our ability to participate in the social world. Rationality without emotion is not rational at all, as emotion makes salient those portions of the environment to which we need to attend. If it is true that emotion does pervade the entire array of our mentality, it should be expected that it plays a role in the psychological features of fictive practice. The task is to show how. I attempt that task in the following chapters.

3. A Theory of Fiction

3.1 The Theory of Fiction

I have developed a theory of the emotions that sees them as a heterogeneous grouping of mental and cultural phenomena. I now need to develop a theory of fiction to see how these heterogeneous states fit into fictive practice. What should a theory of fiction amount to? Moreover, given I am seeking to explain the involvement of the emotions in fiction, what do I need to know about fiction to satisfy this understanding? I will develop a theory of fiction focusing on the aspects of fictive practice especially relevant to our emotional involvement.

There are a variety of traditional philosophical problems involving fiction. I will take a look at these preliminary to developing my theory of fiction. Several extensive works in the philosophy of art on our interactions with fictions, including those of Currie, Walton, and Lamarque, deal in part with issues that *pre-date* the philosophy of the arts' concern with the paradox of fiction and the emotions. These issues are familiar from the philosophy of language and became especially important to the analytic philosophers of the early Twentieth Century. I will argue they are both historically and theoretically prior to some of the vexing issues dealt with under the rubric of the paradox of the fictional emotions. I doubt they should be the central issue here, as they threaten to *derail* our interest in the nature of the emotions that are involved in fictive practice.

A first problem concerns the apparent *reference* the language arising from fictive practice bears to fictional entities, especially where proper names are involved. Within non-fictional language, declarative sentences such as "Frank Zappa had a groovy moustache," make reference to objects in the world, in this case, Frank Zappa and his moustache. In the case of fictive language the situation seems quite different. The proper name "Frodo Baggins" cannot refer to the individual Frodo Baggins in the manner "Frank Zappa" refers to Frank Zappa, because the former surely does not exist in the manner the latter does. There seems to be nothing for the name to refer to, at least nothing that looks at all person-like (or *hobbit-like*, for that matter). Though it is most apparent there, the problem is not restricted to proper names, given that our descriptions of the states of affairs of fictional worlds are not taken to refer to states of affairs in an actual world.

This problem of reference has implications for other features of fictive language, such as for notions of *truth* and *falsity* in fictions. Ascriptions of truth and falsity do seem applicable to fictional worlds. For example, it is true of the fictional world of J. R. R. Tolkien's *The Lord of the Rings* that Frodo is a hobbit, and false that he is a goblin. Furthermore, the reference of "Frank Zappa" contributes to the truth conditions of the declarative sentence "Frank Zappa had a groovy moustache" in that if the object the proper name "Frank Zappa" refers to had a groovy moustache, we would know the sentence was true. To verify the truth of the sentence we would look at the sentence's correspondence or lack of correspondence with the world. Verifying the truth of the sentence "Frodo is a hobbit" would involve quite a different process. It would involve in some part looking at the work of fiction *The Lord of the Rings*. *Intuitively* though, it seems that whether Frodo was a hobbit was true or not would depend on states of affairs in the fictional world of Middle-Earth. The problem of course, is that Middle-Earth does not exist. It is the work of fiction that seems to provide the *objectivity* that verifies the truth or falsity of statements about the fictional *world* of Middle-Earth.

This is already to complicate our notion of truth and falsity; but it gets worse. It seems to imply that truths in fictional worlds are contingent upon the identities of fictional works. For example, in Peter Jackson's films of *The Lord of the Rings* trilogy, there are factual differences to the Tolkien books. In Jackson's films, much to the chagrin of some Tolkien fanatics, elves turn up to assist the humans at the battle of Helms Deep. Now whether it is true or not there were any elves at the battle of Helms Deep will be relative to whether we are concerned with the world represented by Tolkien's book, or Jackson's film.

This seeming lack of reference of proper names poses a problem for the *meaning* of fictive sentences because of the connections between truth and meaning. If reference contributes to meaning in virtue of contributing to the truth conditions of sentences via the roles of proper names and other referring terms, then without reference we are in danger of not having a theory of meaning applicable to fictions that explicates sentences such as "Frodo is a hobbit." Yet it is obvious we *do* need a theory of meaning in fictive language given our interest in *interpreting* fictions. Moreover it seems a truth conditional semantics would be the type of thing we were after. There is a warrant or utility in applying truth conditional semantics to fictive works and fictional worlds given that the relationship in meaning between individual fictive sentences may be governed by truth conditions. Fictional sentences are profitably subjected to many of the same inferential operations as real sentences. Truths about hobbits, and the fact Frodo is one, will allow appreciators to infer truths about Frodo. Truth and falsity play a role in our interpretative participation with works of fiction. What is needed is a theory of how semantic particles such as proper names play a role in generating content when they lack the features they are assumed to have in non-fictive language.[35]

It might be thought the meaning of fictive language depends not on what is *literally* meant, but what is *fictionally meant*. We might think "Frodo is a hobbit" does not mean what is literally implied by the sentence, but is some sort of non-literal usage of language. This is not the case.

---

[35] This is the complicated project Currie carries out to some success in The Nature of Fiction (1990). Currie thinks apparent proper names play various semantic roles in fictive language, determined by the use to which they are put. When used by authors in the production of fictions they act as "bound variables" (1990:146-155). When used in "metafictive" senses, that is, by audiences when they discuss fictions, they act as "abbreviated descriptions" (1990:158-162). Currie also sees a third use of proper names in fictive contexts where they refer to "roles" (1990:171-180). This may be related to what I will later call the "external" or "factual" reference to characters [3.5].

There is a real sense in which fictive language literally expresses content (Currie, 1990:19-20). This is shown by the fact the language found in fiction includes *both* literal and figurative content. If I claimed Frodo was a rock, I would figuratively mean he was strong and resilient, not that he was a rock. Thus when I say Frodo is a hobbit, *I literally mean Frodo is a hobbit*. I will eventually argue that important to understanding this will be the distinction between the *content* of a sentence, and the *pragmatic usage* of the sentence [3.5]. The conventional meaning of the content of a sentence may differ from what a particular utterance is used to mean, and it is the pragmatics of the usage that generates this latter sort of meaning. This shows that the term "meaning" is ambiguous, and that we must be careful to distinguish issues of content from issues of pragmatic force (Currie, 1990:6).

Ontological questions are beginning to impinge on this discussion. I have just claimed that the truth or falsity of fictional sentences seems to be contingent on the *fictional world* we take ourselves to be dealing with; but what are fictional worlds? What are fictional worlds if they literally do not contain seas, mountains, clouds, and people? Further, what are fictional worlds given that they are contingent on the identities of works of fiction, because Jackson's movies and Tolkien's books represent subtly different fictional worlds but both depict the happenings in Middle-Earth?

Also, what are the entities that inhabit these fictional worlds? In particular what are fictional characters? What or who is Frodo? He cannot really be a hobbit, as really there are no hobbits. And yet, if we take the language use of those who talk about Frodo, Anna Karenina, or Leopold Bloom at face value, these entities seem to have properties ordinary people (or hobbits) do. Frodo is less than four feet tall, lives in a hole in the ground, and smokes weed. Bloom likes eating fried kidneys, has a job as a canvasser of advertisements for the *Freeman's Journal*, and has bodily functions (of which we are made aware throughout *Ulysses*). Saying Leopold Bloom exists is clearly not a meaningless expression. Perhaps, I might suppose, he is some shadowy and ontologically strange *fictional person*. But this begs me to explain more fully what it is to be a *fictional person*. Perhaps I should just bite the bullet and also attempt to *explain away* commitment to these ontological oddities. This relates to the problem of the reference of proper names in fictive language; what does the name "Leopold Bloom" refer to, if indeed it refers to anything at all?

There is also a distinct ontological problem with fictional characters in that they seem to have a *dual nature*. In forming the judgment that Bloom thinks Stephen, despite the bad teeth, has the potential to become an all-round cultured man, I seem committed to positing the existence of *man* named Bloom. The implied ontological nature of fictional characters is different when I write in the same paragraph that the "passage draws together a great deal of the detail from the previous five hundred or so pages of *Ulysses*, and makes explicit the thematic relationships that stand between Bloom and Dedalus." Here Bloom and Dedalus are obviously considered as literary constructions; it is *thematic* and not *personal* relationships we are interested in. What does *this* imply about the ontological nature of Bloom and Dedalus? The names "Leopold Bloom" and

"Stephen Dedalus" must clearly be serving different denotative functions here. Whereas above the names seemed to pick out people, here they seem to pick out literary creations.

Furthermore, and relating to the semantic problems discussed at the beginning of this section, we talk in ways that seem on the surface to commit us to *believing* things about fictional characters: I assert that Bloom thinks Stephen, despite the bad teeth, has the potential to become an all-round cultured man. What does this assertion mean, and further, how can I say such a thing if I do not *believe* in the existence of Leopold Boom and Stephen Dedalus? And what does it mean to say I believe such a thing? The problem is compounded when we realise beliefs are not the only *mental attitudes* we seem to hold toward fictional characters; we also *observe* fictional characters, *judge* them, *predict* their motives and actions, and perhaps worse, *feel* emotions for them. How can we make these ascriptions of mental attitudes intelligible? Also, it might be said, where it seemed deeply paradoxical how we might believe things about Bloom and Dedalus *qua people*, *qua literary creations* it is a simpler matter: our beliefs concern aspects of Joyce's literary artifice, which was created and exists very much in this world.

****

But there is a great deal to fictive practice other than this semantically and ontologically suspect language. Besides this language use, fictive practice involves an imaginative stance and our interpretations, judgements, and emotional reactions that occur within this stance. Fictive practice also involves our actions, not only in the form of verbal reports, but also, in the case of role-playing games and videogames, our overt physical actions. Fictive practice involves *works* of fiction, which are created by authors and other fiction producers, and are used, somehow, by audiences to generate fun, interesting, profound, and sometimes emotional, fictive experiences. Fictive practice also involves wider meanings and interpretative practices. George Orwell's *Animal Farm* might on a naïve reading seem to be an unlikely tale of a farmyard revolution, but most of us know what it is *really about*.

It might be argued that the semantic and ontological problems have to some extent hijacked the philosophical concern with fiction, given that language is only one among many important features of fictive practice. Currie's work starts out with a focus on how the notion of truth contributes to the meaning of fictive works, and very quickly turns into a complex discussion of meaning bearing all the hallmarks of the modern philosophy of language. But given its focus on truth, it is arguable that Currie's theory bears little resemblance to how people interact with fictions, and even less that it is a theory *of the nature of fiction*. To see the relationship between emotion and fiction philosophers need to widen their concerns beyond the semantic features of the language involved in fictive practice.

Fiction essentially involves a special psychological stance. How do we participate with fictions such that what we know to really exist – books and plays and movies, and the descriptions and images therein – are transformed into the apparent objects of our attention – people and situations? We must postulate a process whereby the content within fictional works is cognitively held so that the features obvious in fictive practice can take place. What is the nature of the psychological stance that subserves this imaginative involvement? This stance must not only explain the semantically and ontologically peculiar language that arises from fictive practice. It must also explain the actions, inferences, interpretations, mental ascriptions, and emotional states that involve appreciators in works of fiction. These are the substantive features of fictive practice the following theory attempts to explicate.

There is a particular problem with the inferential processing of the representational tokens that subserve our cognitive involvement in fictions that any theory of fictive practice must cope with. Fictive sentences are profitably subject to many of the same inferential operations as real sentences are, but this glosses the fact that inferential operations are made *against the background of a significant body of knowledge*. The problem is that the body of knowledge that is made reference to in the case of fictive inferences will be very different to our actual body of knowledge; the two may even contradict each other. For example, I believe hobbits do not exist, and yet I am able to infer Frodo is an unlikely bearer of the evil *one ring*, given that hobbits are distinctly domestic creatures. This problem runs two ways. Firstly, it poses a problem for fictive sentences, mental states, and inferences. How do we know what is relevant when we come to make fictive inferences, and what can we assume about the fictional world? Second, it poses an important problem for psychology: how is the background knowledge for fictive inferences *kept separate* from our genuine knowledge, because this duality seems an ideal opportunity for inferential havoc, or *irrationality*, to take place. Our fictive abilities are made all the more remarkable because they are successfully *quarantined* from our genuinely representational cognitive states.

Fictive practice makes us emotional. We cringe with fear when the giant spiders attack the spacecraft. We sob when a bad space person kills the navigator, and realise the rest of the crew will never get back to the Earth again. We are emotionally affected by fictions, and this is something that is obvious from more than semantically suspect mental ascriptions. Appreciators of fictions sob, cheer, shake, and even *avoid* fictions because of their emotions; the question of whether we have emotions for fictions is very much an empirical question that will not be solved by *a priori* argument. This is the central problem of this thesis, but it is one among many interesting features of fictive practice, particularly of the psychology of fictive practice. As such, our emotional involvement in fictions will only be explained when situated within a wider picture of our psychological and behavioural involvement in fictions.

If we look closely, we can see that fictive practice also involves appreciators in performing *actions*. These actions not only come in the interpretative, verbal, and cognitive actions already noted, but often come in the form of overt physical actions. Furthermore, in some fictions, especially videogames, appreciators seem to *act in fictional worlds*. Videogames and role-playing games seem to be among the rare exceptions to the general rule that we cannot affect the lives of fictional characters or act in fictional worlds. In the marvellous *Grand Theft Auto 3* on the Playstation 2, the player character can get up to all sorts of nefarious actions in Liberty City, racing around the streets, mugging old ladies, and generally getting up to no good. But the example of a backwoods yokel, who jumps on to the stage to stop Brutus from plunging his dagger into Caesar's back, has become a cliché. The yokel is making some sort of mistake, forgetting there is an ontological divide between his actions and those that might affect the fictional world of the play. What sort of mistake is the yokel making? Furthermore, why are overt actions available to role-players and videogamers when they are denied to appreciators of most literary and dramatic fictions? When and why are actions available to appreciators of fictions, and when is it possible to act in a fictional world?

What are *fictional works*, and how are they related to fictional worlds? I have argued that the facts of fictional worlds are problematic: they do not seem to be based in the *states of affairs* of worlds, but in fictional works. How are fictional works involved in the cognitive stance that allows our participation with fictional people and situations? How do fictional works provide the *objectivity* that seems to ground the facts of fictional worlds? Also, what are the *identity conditions* of a fictional work? Fictional works are in some regards indeterminate; they leave many facts implicit, and it is up to the appreciator to generate these facts. Given fictional works are in this way constrained in terms of *repleteness* – that is, how much they make explicit about a fictional world – how do we generate the facts of fictional worlds relevant to the interpretation of fictive works?

How are the *wider meanings* of fictions generated? *The Lord of the Rings* is explicitly and literally about a hobbit taking an evil and magical ring to Mordor (and much else besides this), but, rightly or wrongly, it is also sometimes claimed to be about other things. Some see *The Lord of the Rings* as an allegory of the great conflicts of last century. Others have seen environmental issues, or a disparagement of modern technological culture. I see a strong thread – especially in Jackson's movies – dealing with the struggle of giving up childhood fantasies in order to deal with the world in a realistic and adult manner.

Standing behind many of these observations of fictive practice is the assumption that the language use and attitudes of appreciators of fictions are *intelligible*. We are charitably assuming fiction appreciators *are not crazy or irrational*, that is, that they do not somehow actually believe in the content of the fictions. Indeed, the discovery that in cases of fictive appreciation, appreciators were struck with some strange sort of selective aphasia would significantly clarify matters. Fiction appreciation would just become another sort of mistake, and hence in no need of theoretical reconciliation with our theories in language and psychology. I hope the reader has noted the connection here to Radford's charge of irrationality of the emotions I detailed in the introduction. Indeed, I will contend the particular problem that gave rise to the charge of irrationality, though it seemed a problem particular to the emotions, is a *species* of more general philosophical problems with fictive practice.

I am going to begin this theory of fiction, not with a treatment of the traditional problems of fiction, but with a substantive discussion of the psychology and behaviour involved in fictive practice. I hope his will allow me to avoid what I see as a skewed treatment of the nature of our emotional involvement in fiction. If we begin with the traditional philosophical problems of semantics and ontology, we are in danger of forgetting that these are problems with *fictional language*, whereas fiction exists as a complex psychological and cultural practice of which language is only one among many parts.

3.2 Cognition and Pretence

The psychology of fiction is dominated by a particular cognitive stance, namely, pretence. I will argue that most of the theorists who seek to provide a theory of fictive participation see the need to postulate some cognitive stance to explain the features notable in fictive practice. Walton invokes the notion of "make-believe" (1978, 1990), as does Currie, and so both can be termed pretence theorists, given *make-believe* and *pretence* hold strong associative ties.[36]

Others, however, are reluctant to be seen as pretence or make-believe theorists of any kind. Lamarque discusses the "fictional stance" (Lamarque, 1996; Lamarque and Olsen 1994), Carroll writes of "unasserted thoughts" (1990, 1998), Feagin of "mental shifts and slides" (1996), and Robert Yanal of a "thought theory" (1999). All resist notions of make-believe, especially concerning the emotions involved in fiction. I think Carroll, Lamarque, and Feagin all provide essentially *negative* characterisations of the psychology of fictive practice, and that these characterisations are unsatisfying. I will later argue that in as much as these theories do attempt to provide a positive framework – and they do in that they all postulate a *process* whereby holding tokens of representational states before our mind leads to our generating the features we note in fictive practice – it threatens to collapse into a commitment to pretence, or even to make-believe.

I think Walton's make-believe theory is closely tied to his answer to the semantic and ontological issues identified earlier, and that this puts off some people. However, any successful theory of fictive participation needs to propose a cognitive mechanism whereby thoughts are held before the mind and produce the kind of fictive participation described in the previous section. Nothing important hangs on the issue of whether I call the mechanism "make-believe," "pretence," or something else, other than some of the associations the terms might presently hold. For this reason I will eventually reject the purportedly *doctrinal* division between "thought theorists" and "make-believe theorists" identified by Yanal (1999) [4.2&4.4].

However, even those theories that do explicitly observe the need for the process of pretence or make-believe are likely to be sketchy in the details. Nichols and Stich set out to explicate the cognition of pretence, which they think “underdescribed” by previous theories (Nichols and Stich, 2000:115). As pretence does play a central role in our fictive participation – Nichols and Stich also think it likely to play a central role in *imagination* – I should pay careful consideration to the theory.

---

[36] Currie (1990:12-18), argues against the form of pretence theory offered by John Searle (1974). I am using the term "pretence" as a synonym of "make-believe," so on my view Currie is a pretence theorist. I do not think this should prove controversial, given that the theory of pretence I develop here does seem consistent with much of what Currie wants to say about the role of content in fictive practice.

Pretence is more primitive than fiction. Nichols and Stich discuss various examples of games of pretence in children, and hope to draw from them some of the features that are characteristic of pretence. The features they describe explain some of the aspects of fictive practice I noted in the introduction to this chapter. I will list the five aspects of pretence they think important, try to fit them into the scheme of things here, and then elaborate Nichols and Stich’s model of the cognitive architecture that subserves pretence.

A first aspect of pretence is the “starting premise” or jumping off point of the episode. Any decent theory of pretence should aim to explain how the initial premise of an episode is established in the mind, and how it becomes the basis for the following cognitive operations and pretence behaviour. Consider a childhood game of pretence in which Betty is pretending to talk to her mother on a telephone, using a banana to stand proxy for the phone. Among the starting premises of this game is that the banana is to be treated as a telephone, and that Betty is talking to her mother. The continuation of this game of pretence will involve an elaboration of these starting premises via various cognitive operations. How does one adopt these starting premises? Nichols and Stich hope to provide an answer to this question.

As far as fiction goes, we can see that the starting premises are incredibly complex: they are the content presented in fictions. The premises of fictions are often also perceptually modal representations of content, as in films. In these cases the premises are constituted in part by visual and aural representations. Furthermore, not only are there starting premises, but given that most fictions are temporally extended, there are further premises that guide the action along. Pretence involved in fiction, then, will have to take account of this increase in intricacy afforded by the complexity, representational modality, and temporal duration fictions can possess. These are aspects of the *props* that play a central role in fictive participation [3.3].

A second aspect a theory of pretence must seek to explain is the “inferential elaborations” that are carried out on the starting premises. Participants in pretence episodes make inferential elaborations on the starting premises during the course of the pretence episode. An example Nichols and Stich give is of an episode in which one of the two pretence participants is a cashier in a fast food restaurant, the other a customer. As a part of the game, the first participant is required to calculate the amount of change due to the second participant on the basis of previous events or premises in the episode, namely, how much money it is pretended was given, and how much it is pretended the fast food cost. Important for my general theory here is that this inferential elaboration often seems *structurally isomorphic* to cases where inferential elaboration occurs in our genuine dealings with the world. The principles behind the inferences, in this case arithmetic, share a structure with real world instances of such inferences. Any successful theory of pretence will have to account for this structural isomorphism.

A successful theory must also explain the third aspect Nichols and Stich note: the "non-inferential elaboration" or "embellishment" that is involved in pretence episodes. The previous events or premises of the episodes do not directly entail all of the elaboration in pretence episodes; the participants of such games initiate much of the material themselves. This may also be the case in works of fiction, where I have noted people are tempted to say things about their relation to the fiction's contents, though implies nothing about them.

A fourth aspect of pretence Nichols and Stich think particularly important is the motivation of action. Why do participants in games of pretence make the actions that are sometimes also evident in fictive practice? For example, why does Betty lift the banana up to her ear as she pretends to speak on the phone to her mother? In as much as fictions involve actions – and I will argue they often do – this question will be relevant to my argument.

The fifth and final important aspect of fiction is something I discussed briefly in the introduction to this chapter. Nichols and Stich think it notable that "the events that occurred in the context of the pretense have only a quite limited effect on the post-pretense cognitive state of the pretender" (2000:120). The cognitive states that subserve pretence are, in a word, "quarantined" from our beliefs. This is interesting for two quite different reasons. Firstly, it questions how we know what is *relevant* when we come to make fictive inferences and what we can assume to know about the fictional world, as often this will differ from, or contradict, our actual beliefs. Second, it poses an important problem of how the background knowledge for fictive inferences is *kept separate* from our genuine knowledge, because without the cognitive quarantine, inferential havoc threatens to take hold. A successful theory of pretence, and pretence as involved in fictive practice, should explain these observations. Nichols and Stich also note there are other aspects of the mind that are not so quarantined from pretence, the most notable of which is emotion. I will argue later that it is not clear this is the case; emotions are *in general* quarantined from our genuine cognitive states [5.2].

Nichols and Stich proceed to outline a theory of pretence that accounts for these issues. Central to the theory is the addition to their standard cognitive functional architecture or "boxology" of a new functional box, namely, a "possible world box."

> Like the Belief Box and the Desire Box, the Possible World Box contains representation tokens. However, the functional role of these tokens, their pattern of interaction with other components of the mind, is quite different from the functional role of either beliefs or desires. Their job is not to represent the world as it is or as we'd like it to be, but rather to represent what the world would be like given some set of assumptions that we may neither believe to be true nor want to be true. The PWB is a work space in which our cognitive system builds and temporarily stores representations of one or another possible world. We are inclined to think that the mind uses the PWB for a variety of tasks including mind

> reading, strategy testing, and empathy. Although we think the PWB is implicated in all these capacities, we suspect that the original evolutionary function of the PWB was rather to facilitate reasoning about hypothetical situations… In our theory the PWB also plays a central role in pretense. It is the workspace in which the representations that specify what is going on in a pretense episode are housed.
>
> (Nichols and Stich, 2000:122)

Pretence involves placing a token of a representational state into the functionally discrete possible world box and subjecting it to various operations. Further, the representational state that is placed in the box is in the same “code” as genuine representational states, which is to say it will be subject to many of the same operations to which genuine representational states are subject. It is hoped this model, which Nichols and Stitch set out in the diagrammatic form below, will account for important aspects they noted in pretence episodes. If it does, I think it likely such a functional architecture also plays a role in fictive practice (though perhaps not an exclusive one).

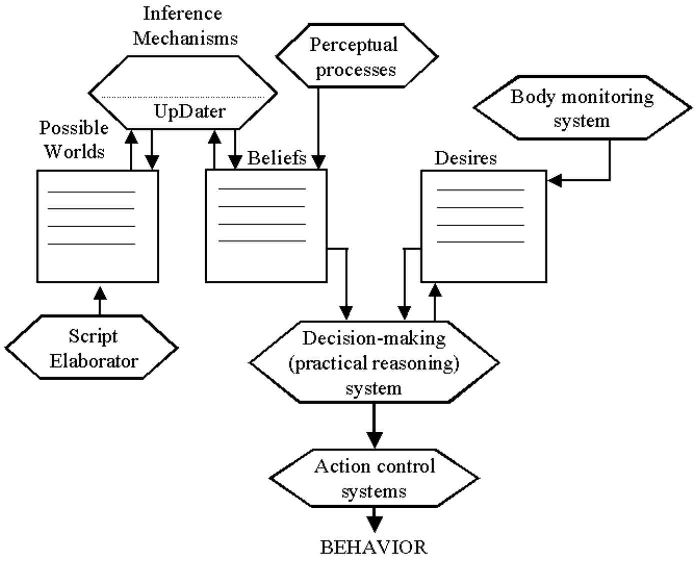

Inference Mechanisms
UpDater
Perceptual processes
Body monitoring system
Possible Worlds
Beliefs
Desires
Script Elaborator
Decision-making (practical reasoning) system
Action control systems
BEHAVIOR

Figure 3

(Nichols and Stich, 2000, fig.3)

The theory accounts for the “starting premise” noted in various pretence episodes by seeing this as involving placing a representation (or, more likely, a set of representations) in the possible world box. In pretending a banana is a telephone participators place a “this is a telephone” representational state in the possible world box. Nichols and Stich think there is a difficulty lurking here, in that reference to the banana needs to remain indeterminate enough to allow the manipulations of the pretence episode to proceed. Features of the premise that are not relevant to the pretence – such as the fact the banana is edible – need to be abstracted away so as not to play a role in the inferences that ground what follows.

Because the token representational states placed in the possible world box are in the same code as genuine representational states, they are subject to many of the same cognitive operations. This explains the second issue; our inferential elaboration in pretence episodes occurs as we subject the representations in the possible world box to inferential manipulations.

Nichols and Stich explain the issue of non-inferential elaboration or embellishment via the introduction of a new functionally discrete box, a "script elaborator," that contributes novel representational tokens to the possible world box. When Betty decides to ask her mother what is for dinner, the content is generated by the script elaborator and placed in the possible world box. The script elaborator adds to the richness and novelty of our games of pretence.

On this model, action is generated as a result of desires to act in a way that is consistent with the state of affairs represented in the possible world box being a state of affairs in an actual world. People act because "they *want to behave in a way that is similar to the way some character or object behaves in the possible world whose description is contained in the Possible World Box*" (Nichols and Stich, 2000:128, emphasis in original). Betty acts as though the representation states at the basis of her game of pretence are an accurate representation of the world, and holds the banana up to her ear because of a *desire to participate successfully in the game of pretence*. The actions are not caused by *pretended desires*, as argued by Currie (1990, 1995a). That is, participators do not act via an intention to *simulate* the desires a character or object might have in a particular role in the possible world. One important reason for this is that some roles participators play *do not have desires*. Nichols and Stich's example is of the child pretending to be a dead cat. Why does the child stay perfectly still? Not because she intends to simulate the desires a dead cat might have – dead cats do not have desires – but because dead cats generally do not move. The participator stays perfectly still because she desires to act in a manner consistent with an object in the world of the pretence or fiction.

It might still be appropriate, however, to see participators as accurately characterised as *fictionally having desires*. The Nichols and Stich theory is incomplete in that it does not deal with various semantic concerns we might have about the language that arises in fictive and pretence episodes. The *mental ascriptions* that arise via language used within a *fictive idiom* will make it true that appreciators will often have *fictional emotions* and *fictional desires* concerning fictional worlds and situations [4.1].

The subsequent actions that are generated in pretence do not get out of control – Betty does act as though the pretenceful conversation with her mother is an actual conversation – because the desires to act in accordance with the representations in the possible world box are simply weaker than those to act in accordance with the states in a belief box.

The final issue, that of cognitive quarantine, is solved by positing that the possible world box is functionally discrete from the actual belief box. Representational states do not freely switch from one box to another, so the threat of irrationality is forestalled. Tokens of representational states are often prone to find themselves in the possible world box, however, because we often *do* use our beliefs in participating works of fiction, however. The inferences we make in fictive participation are made against the background of a pre-existing store of knowledge. My actual beliefs about Dublin and Irish history are obviously useful in engaging with *Ulysses*. It is bound to be a complex story how states are put into the possible world box and it will be even more complex when the cognitive architecture of pretence is involved in fictive practice, where the starting premises are as convoluted as fictions commonly are. This is related to the issue of the generative processes by which we *fill out* the worlds of fiction, and will be discussed later [3.3].

Nichols and Stich are not proposing a story about the neural architecture that subserves these functionally discrete states; so exactly how this *boxology* is instantiated in the mind is up to empirical scientists to discover. They think it better than a fair bet that something like what they describe will turn out to be the case, however. Should this be true, then we can expect the same system to play an important, though not exclusive, role in fictive practice. I think something like the Nichols-Stich theory *is* essential to the characterisation of fictive practice, and that it constitutes a step forward in that it replaces the vague terms previously used to describe the cognitive stance basic to fictive participation with a theory that has its basis in empirical observations and a commonsense functionalist theory of the mind. It may turn out that the processes referred to in other studies of pretence or fictive practice simply *supervene* on this more concrete conception, however. This will be the case if, as I think likely, Nichols and Stich's theory can be best seen as a *complement* to the previous theories of fiction, rather than as a replacement.

3.3 Walton and Make-Believe

Kendall Walton's discussion of fiction and emotion in Mimesis as Make-Believe stands in an important position in this debate. I think there is much merit in Walton's theory, yet my reading of fictive practice will depart from his on several occasions. Walton's theory acts as a complement to the Nichols and Stich theory because it has a different focus. The theory deals with semantic and ontological issues Nichols and Stich's paper is silent on. Walton's theory also allows me to extend the cognitive theory of pretence into a *theory of fiction*, in its explanation of how fictional works are the *codification* of the premises involved in games of pretence.

Walton thinks central to our interactions with fictional worlds and characters is the practice of "make-believe." Make-believe as a concept has many associations: we speak of *worlds of make-believe*; make-believe is a common childhood game most of us can remember indulging in; and the etymology of the term also implies there are things called *make-beliefs*. Walton's notion will employ all of these associations in building what he hopes to be a convincing characterisation of fictive practice.

Walton thinks fictive practice is analogous, and perhaps genetically related, to childhood make-believe. *Play* holds an important place in his explanation of fictive practice. He uses one of the literal senses of the term “make-believe” to inform the nature of play as it exists in fictive practice. Make-believe is a common childhood game. Walton asks us to imagine a game that involves children treating stumps as bears (1990:21ff). As the children explore the forest and come across stumps, it is true in the game that they explore the forest and come across bears. The children act accordingly: they run from bears, warn others of their presence, or perhaps throw stones at them.

The actual nature of the stumps also influences the games that are played with them. Features of the stumps add to what Nichols and Stich would call the “premises” of pretence episodes. When a child comes across an extremely large stump, it is probably true within the game that they have just stumbled upon an extremely large (and ferocious) bear. This fact is something Walton calls “reflexivity” (1990:210). It will play an important role in the story he gives.

The central notion explaining this generation of fictive premises is the “prop” (1990:21). In the children’s game above, the stumps are props; that is, they stand proxy for real bears. Walton notes that children often use such props in their games of make-believe as a method of making the games more interesting and vivid, introducing an element of narrative structure that is out of their control. Corresponding to stumps in sophisticated games with fictions are fictive works. Fictions, and here Walton counts representational visual artworks as well as literary works such as novels and stories, are props that aid appreciators in playing vivid and interesting games of make-believe.

This notion of props as determiners of narrative content is how we derive a theory of fiction from a theory of pretence of the type provided by Nichols and Stich. Games of pretence are fluid or even solitary interactions where the *worlds* pretencefully interacted with might change at the whim of the participators. These worlds exist solely in the practice of the game, and any documentation that might occur is purely the informal descriptions offered by the participators in the game. Games of pretence need not be documented, but the descriptions that *are* offered are already fictional props, guiding the generation of further content. *Fiction* in a less impoverished sense, that is, the type of artefactual fictions of interest to theorists of the arts, are documentations of games of pretence that become the premises for further games in appreciators of those fictions. Fictive props exist as the *codification of the premises of a game of pretence* and provide an *objectivity* that both makes games more interesting to appreciator, and also allows the premises to *cohere into a fictional world*. Currie argues fictions “make it easier for us to weave together a pattern of complex imaginings by laying out a narrative; they give us, through the talents of their makers, access to imaginings more complex, inventive and colourful than we could hope to construct for ourselves” (Currie, 1997a:53).

In fictive practice we treat the descriptions within fictions as true in a way analogous to how children take stumps to be real bears. We *pretend* the descriptions found in a fiction describe an actual state of affairs. Just as children pretencefully applied expectations, understandings, and actions to the stumps that they might apply to real bears, appreciators of fictions pretencefully apply such things to the descriptions found within the fictions. On the page of Joyce’s *Ulysses*, “Leopold

Bloom" appears as a name linked to various fictive descriptions of properties and actions. But when we incorporate these descriptions as a prop in a game of make-believe, Bloom becomes a person in the same way a stump becomes a bear. Within a game of make-believe that we play with *Ulysses*, Leopold Bloom is a real person. He has a heart that pumps blood around his body. He gets hungry and must eat. He has thoughts going on in his head and knows those around him have thoughts going on in their heads. Attributing such properties to Bloom is a sign of successfully entering this game of make-believe.

However, this does not mean we must always be *consciously aware* the representational props warrant adopting the pretenceful attitudes involved in fictive practice. The position is consistent with the possibility that we are able to shift unconsciously into the fictive stance. Nichols and Stich argue children need not be aware of the concept of "pretence" in order to engage in pretence behaviour; they merely need to display enough understanding to be able to engage in pretence behaviour (2000). It is likely that children unconsciously switch into pretence behaviour. This also fits with casual observations of fictive participation, where it does not seem to be a case of explicitly acknowledging to oneself "now I am going to enter the fictive stance." Games of pretence or make-believe need not be consciously adopted; something that may count against using the nomenclature of "pretence" to describe our fictive activities if we too strongly identify the process with our pre-theoretical understanding of the term. Indeed, such a conscious realisation can often serve to frustrate attempts to participate with fiction, as anyone who has spent a day studying the philosophy of fiction, and then subsequently attempted to enjoy a piece of fiction, may be aware!

Corresponding to a child's running from or throwing stones at a bear, appreciators of fictions make complex psychological actions concerning fictions. They adopt fictive attitudes toward fictional characters, ponder their motives, make moral judgements, and even report doing so, actions that are in evidence in my passage concerning *Ulysses* at the beginning of this essay. These are all aspects of the complex interpretative processes that involve us in fictions.

Another substantive detail of Walton's account that needs to be discussed in order to explain the complex interpretative processes involved in fictive practice is the distinction between "game-worlds" and "work-worlds" (Walton, 1990:58ff). When we talk of the world of *Ulysses*, we are talking of the "work-world" of the novel, which in this case is a world very similar to real world Dublin, Thursday 16th June 1904. The work-world of *Ulysses* is constituted by the propositions that are fictionally expressed within *Ulysses*.[37] In the world of *Ulysses* Paddy Dignam is buried, Buck

---

[37] Walton is aware he is adopts a metaphysically controversial notion when he writes of "fictional propositions," but thinks that such worries will not affect his account of fictive practice (1990:36). Because fictional propositions are fictional in virtue of being the *apparent commitments* of language arising from a game of make-believe (or by fictive practice being *fictionally characterised*, as I shall argue) I think Walton is correct in this. If "proposition" drops out of philosophical usage, Walton's theory could simply adopt the new theory circumscribed within a game of pretence. If "propositions"

Mulligan goes for a swim, and Stephen gets rather drunk, as we discover when we read the text. But the work-world also includes those propositions legitimately derived or generated from the explicit content of the book. It will become of much interest for interpretative issues to ask what it is to generate the fictional truths of work-worlds.

Game-worlds are constituted by all the fictional facts that seem to make reference to the appreciator of the fiction. There are facts that are fictionally true of our fictive appreciation that are not fictionally true of the work-world of the fiction. It is fictionally true of the work-world of *The Lord of the Rings* that the ring obsesses Gollum, but it is not fictional of the work-world that *I observe* Gollum's obsession. That fact is fictionally true of my *game-world*. Walton thinks that when we appreciate fictions we create a fiction about our relationship to the work-world. The game-world is a fiction in which the work-world is *nested*.

Walton is broaching one of the philosophical questions that were introduced at the beginning of this section, namely, the problem of fictional worlds. Walton ultimately thinks the talk of fictional worlds open us to the *danger* of ontological profligacy (1990:57-58), but continues with the usage because of its theoretical utility. Indeed, the notion of a fictional world, and its sub-types, work-worlds and game-worlds, though ultimately not referring to anything world-like, seem an indispensable part of our fictive interpretative participation.

The notion of a *game-world* allows Walton to discuss interpretative matters in a very fruitful manner. It may be true of different people that they play different games with the same work of fiction, and so it is possible for single work-worlds to inspire multiple game-worlds. At the heart of interpretation is the attempt to formulate a game-world consistent with the facts fictional to the work-world and the external stylistic features of the work that *colour* such fictional contents. A successful interpretation will not only involve an accurate appraisal of what is fictionally true in the world of the fiction, but also the formation of appropriate *attitudes* to these fictional propositions. The work-world of Bret Easton Ellis' *American Psycho* contains much that is shocking and repulsive, but we should question whether someone has successfully interpreted *American Psycho* if their attitude consists exclusively of shock and repulsion. *American Psycho* is also an extremely black comedy and, interpreted in this way, our attitude to some of the material may be one of humour. Such an interpretation demands we take into account aspects of the artifice of the author and the novel. The meaning of works is unlikely to be determined solely by the facts we derive from our pretenceful engagement with their fictions [3.4].

As just mentioned, work-worlds are not comprised only by the fictional propositions explicitly presented within the work. Despite how replete we might think a novel such as *Ulysses* is with the minutiae of every day life, restricting its world to those things indicated by the content in the

---

were replaced by "notional worlds," (Dennett, 1987:152-155) it would become "fictional that we relate to notional worlds." I will argue there is another more serious problem with Walton's treatment of the role of propositions in fiction, however.

work would leave *Ulysses* still a radically incomplete work-world. Joyce never (to my knowledge) mentions walruses in *Ulysses*, but we are probably meant to assume the work-world of *Ulysses* is one in which walruses happily go about their business. This does not mean we need to *ocurrently* assume walruses exist in the work-world of *Ulysses* in order to understand it; such obscure inferences are likely to be entirely irrelevant to an understanding of *Ulysses*.[38]

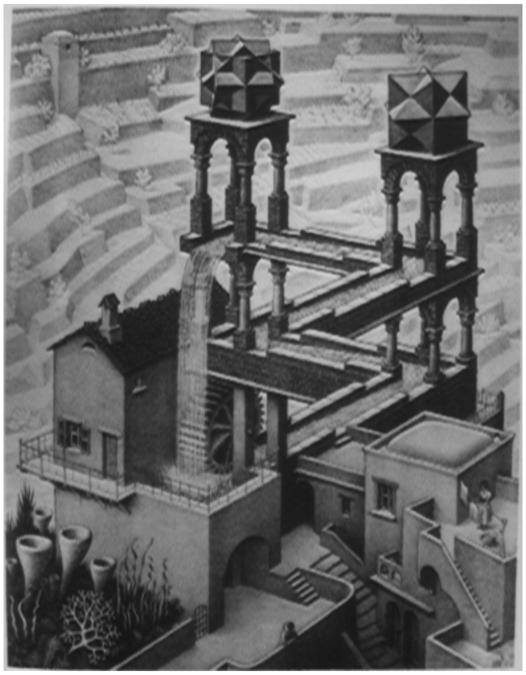
The point is that various narrative or pretence conventions give *Ulysses* an implicit generative structure that licenses our inferential and imaginative cognitive actions. This, as I argued, was part of the justification for thinking a truth conditional semantics appropriate within our interpretation and construction of fictional worlds [3.1]. Besides the explicit contents of its constitutive descriptions, we can assume a multitude of other facts *fill out* the work-worlds of fictions. It is part of our task in appreciating works of fiction to generate such facts where they might be important to the more explicit content of the works.

We must be careful, though, in what generative principles we apply. The world of *The Hobbit* or *The Lord of the Rings* may be one in which we *cannot* assume walruses happily go about their business. It certainly is not one in which NATO could be expected to sort out any renegade dragons. We must take into account that the work-world could be significantly different to our own. Furthermore, it might even be ambiguous as to which work-world is correctly attributable to a fiction. Henry James' *The Turn of the Screw* is a possible example, being interpretable as either a ghost story or a portrait of madness (Currie, 1990:66).

Also, a fictional world might be logically strange, or even logically or physically impossible. The fictional world itself might be ambiguous and the sorts of inferential processes we apply to our own world may not be applicable there. Lewis Carroll details work-worlds having features we would ordinarily think logically impossible, the famous example being that of the Cheshire Cat and its smile. Another set of famous examples are the works of Escher, especially the lithographs *Ascending and Descending*, and *Waterfall*, where what we would ordinarily think of as impossible worlds are represented. In the latter work, water flows down a channel that joins back on itself.

---

[38] It is often claimed that the worlds of fictions have the curious property of being *indeterminate*. Are there an odd or even number of hairs on Sherlock Holmes' head? Presumably it is one or the other, but which one is not made explicit in Sir Arthur Conan Doyle's works. In Currie's discussion of truth in fiction (1990), this indeterminacy plays an influential role. He eventually denies that truths in fictions are based on the facts of fictional worlds, claiming instead that it is our interpretation of a "fictional author" that grounds fictional truth (1990:73-81). I suspect there is some confusion between *relevancy* and *truth* when it is claimed that fictional worlds are indeterminate, though I do not have the space to provide an argument here.



(Escher, *Waterfall*)

The fact that walruses likely exist in the work-world of *Ulysses* will not amount to much, but such obscure possibilities will sometimes help us construct coherent readings of particular works. There also exists the possibility authors will arrange for the meaning of their works to hang on minute and not particularly salient detail. This is the case in the *who done it* genre of crime fiction, where we are primed to take account of seemingly irrelevant detail, sifting the important clues from the red herrings.

How do appreciators *generate* the truths of fictional worlds? Walton thinks there are at least two interpretative principles guiding this and suggests the form the principles will have (1990:144-169). Most of the inferences we must make in filling out the worlds of fictions are guided by the "reality principle" (1990:144-150) or the "mutual belief principle" (1990:150-161). I think participators also use what might be called a "trans-fictive" principle.

The reality principle asks us to assume that the general properties of the world of the fiction are very much like those of the real world. Under this schema, participators use their actual beliefs about the real world to back the inferences they make concerning the content of the fictional world. If it is not stated in a work of fiction that shooting creatures at close range with powerful shotguns generally kills them, we should assume this is nevertheless the way the world of the fiction works, because reality is like that. The work is not like reality in all regards – because the work *explicitly* describes facts that are only fictionally the case – but in general the work-world is assumed to be similar to reality. The reality principle is consistent with the world of a fiction being very different to our world because it may be that it is explicitly noted in the work that things are strange, for example that shooting creatures with powerful shotguns does not kill them. Such might be the case in a fictional world involving hoards of zombie undead, where, as everyone knows, *the headshot is the only true stopper*.

The reality principle is problematic because in some fictions the fictional world is different to ours, though this is not explicitly stated. In ancient Greek drama, the author thought the world was very different to how we, a modern audience, imagine it is. In this case we should adopt the mutual belief principle by imagining that the fictional world of a work is much like the original author and audience would have believed it to be. If it was a common belief at the time of the fiction that the gods played moralistic and sometimes capricious games with people's lives, though reality is not like this, we should assume the world of the work is one in which the gods played moralistic games with people's lives. Our interpretation of the work is much more successful in this case if we apply the belief sets of the original audience to interpreting the work.

There are instances of interpretative generation that neither of these principles seems to govern; for example in viewing a child's picture of a witch we fill out missing details we do not believe in, and which we assume the child does not believe in either (Walton, 1990:161). In this case we assume neither that the world of the child's fiction is like the real world, nor like what they believe about the real world. Instead, in filling out the child's picture with relevant detail, we make reference to a body of *pre-existing fictive folklore* with which the child and we are mutually acquainted. I call this the "trans-fictive" principle of fictive generation. We generate the truths of these worlds not by assuming the world is generally like the real world, nor by assuming the beliefs the author and intended audience might have had, but by making reference to a body of fictive folklore we are aware of from our other fictive dealings with fictional worlds.[39]

Some fictions, especially those that exist in fantasy and science fiction genre communities, give rise to their own interpretative guidelines that we must be acquainted with if we are to come to a successful interpretation. Consider the fictional world, moreover *fictional universe*, of *Star Trek*. There is a rich wealth of detail to the *Star Trek* fictional universe that may not be explicitly referenced in a particular episode or movie, and yet which we may need to draw on to successfully interpret the action of that episode. These interpretative activities do not depend on the reality principle, nor the mutual belief principle, but on an acquaintance with a wider body of folklore, or prior contact with a particular fictional world. Indeed, the *Star Trek* fictional universe is so detailed there are obsessively detailed fictive encyclopaedias, histories, and philosophical and scientific works written about the universe.

An important example of an actual interpretative process we might employ in generating propositions within fictions is the mental simulation or theorisation that is of current interest to both philosophers of the mind and philosophers of the arts [2.4]. Mental simulation or theorisation – or whatever processes are settled on in these discussions – will play a significant role in generating fictional facts important to the meaning of those fictions, such as the reasons and motives of characters within fictions. *Domain specific* cognitive mechanisms will play a role in generating facts about fictional worlds [6.2].[40]

---

[39] This is related to the issue of "inter-fictional carry-over" that is discussed in Lewis (1983:274) and Currie (1990:86-87).

The question of which of the three interpretative principles this form of mental interpretation is guided by is an interesting one. If the process is characterised as simulation then it seems to conform to the reality principle; however, the assumption that the fictional world is like the real world seems to derive from the fact that we are using our own real-world simulative mechanisms in the interpretation. If it is decided our theory of mind is based on a theoretical activity, this ability to understand others might also adhere to the latter two principles of generation when it is involved in fictive practice. We are able to *retool* our theory of mind to fit peculiar facts about the author's fictional world, or the world of a body of pre-existing folklore. In interpreting *Star Trek*, it might be necessary to revise or retool out theoretical commitments about behaviour causation to take into account the fact Vulcans suppress their emotions, and that Data the android does not have any emotions at all.

***

Semantic and ontological issues intrude on Walton's discussion of fictive practice, significantly complicating the issues. It is here in Walton's theory that we begin to see the shift to the distinctly philosophical problems introduced in the opening section of this chapter. The theory, initially a substantive theory of fictive practice, turns to the questions of semantics and ontology. I will discuss Walton's dealing with these problems as a preliminary to a later section where I attempt to resolve these problems in a discussion of the pragmatic features of the language arising from fictive practice [3.5].

First is the problem of the characters that, via games of make-believe, seem to become real people, and seem to be the referents of the apparent proper names in fictions. For Walton, saying Leopold Bloom is a real person or attributing properties to him such as having thoughts in his head, is not to be read as a commitment to *ontologically strange* people who somehow really exist within works of fiction; that is, it does not imply a commitment to fictional objects *qua people*. The view is far more parsimonious (not to mention sensible) than this. Furthermore, the statement is not to be read as fictionally attributing the property of *personhood* to *something else* that needs to be discovered. For Walton fictional statements are not attributions of properties, or claims of existence at all; they are pretended attributions and claims of existence. Walton thinks it is beside the point to ask what they are *really about* (1990:385ff).

---

[40] The involvement of domain specific mechanisms in fictive interpretation may show that the three principles of generation are in part *normative* in intent. They may not describe how we participate with fictions, but are criteria for discovering what is true in a fiction. This appraisal certainly seems appropriate of Currie's theory of truth in fiction, given his admission that his theory seeks to explain a "semantic ideal" (1990:175).

For Walton, games of make-believe are game-bound practices that contain tacitly understood rules that license language use that on its surface may seem metaphysically and ontologically profligate or incoherent. This metaphysical worry is resolved by seeing our fictive language use as emerging from games of make-believe. The rules of games of make-believe *qualify* the attitudes and language of participants arising from those games. These rules qualify the worlds and entities seemingly referred to in fictive practice, as *fictional*. It is only fictional there is a character called Stephen Dedalus. Fictional characters do not really exist. The same is true of fictional worlds, and the fictional entities that inhabit them. For Walton, talk of fictional worlds – and their *contents* – is "dangerous" (1990:57-58).

The solution to the apparent truth and falsity of the language involved in games of make-believe – both that situated within works of fiction, and that arising from appreciators – is related to this aversion to attributing properties to fictional worlds. Sentences regarding fictional worlds are not *true simpliciter* – or *false simpliciter* for that matter – but are truths concerning the world of the fiction. They are *fictional truths*. A *fiction operator* always qualifies attributions of facts concerning work-worlds. It is not true simpliciter that Dedalus has bad teeth; it is fictional that this is true.

The solution to the problem of fictional reference is related to the treatment of fictional truth and falsity, given the role reference has in establishing the truth or falsity of sentences. Walton replaces conditions of truth and falsity with conditions of *fictional truth*, and *fictional falsity*, thereby resisting the exotic implication that declarative sentences within fictions refer to actual individuals or states of affairs, while retaining the utility a truth conditional semantics has within fictive practice. The reference ordinary *asserted* declarative sentences have to individuals – or apparent individuals in cases of reference failure – is among the items of illocutionary force that are suspended within games of make-believe. It is replaced by *fictional reference* to *fictional states of affairs or individuals* within a game guided by the tacit rules of fictive practice.

Walton's introduction of "fictional propositions" in his theory of fictive participation is ambiguous at this point, however. This in fact is the major problem I have with Walton's theory, and why my own departs from his at crucial points. Walton denies there are propositions one pretends to entertain or assert, holding instead that fictional propositions are only the *apparent* commitments of the language and attitudes arising in our episodes of pretence:

> Insofar as statements appearing to be about fictional entities are uttered in pretense, they introduce no metaphysical mysteries. When Sally says "Tom Sawyer attended his own funeral" in pretense, it is *fictional* in her game that there is someone whom she calls Tom Sawyer and who she claims attended his own funeral. We need not suppose there is really any such person. We need not suppose even that there is someone, or some character, about whom Sally *pretends* to speak, nor that there is a proposition that she pretends to

> assert. Her utterance is fully intelligible as an act of participation if we take it to be *merely* fictional, not true, that her words express a proposition, that they express a proposition about someone whom she calls Tom Sawyer, and that she asserts such a proposition to be true.
>
> (Walton, 1990:396)

Thus Walton does not favour the option of seeing the apparent reference to fictional people, entities, or propositions as transformable into references to descriptions, but thinks it only a pretended reference. Appreciators only pretend there are propositions, characters, or entities to which they attend or refer.

However I think it must be the case there is something like a cognitive representation state that subserves the *content* of the make-believe and the subsequent inferential transformations of it that occur in fiction practice. There *is* something involved in Sally's fictional attributions some might be tempted to call a proposition. *There are and must be cognitive representational states underlying our fictive participation, otherwise our fictive participation would amount to mental magic.*

I suspect that two mistakes contribute to Walton's error here. Firstly, I think it might be that Walton is a little woolly over exactly what he thinks a proposition is. If I substitute "state or affairs" for Walton's use of the term "proposition" it is not so clear he is wrong to deny that there are propositions involved in fictions. Whether he should be read in this manner is admittedly unclear. I suspect this may be the mistake Walton is making because of the first clause of the sentence in which he denies the involvement of propositions in fiction: "We need not suppose even that there is someone, or some character, about whom Sally *pretends* to speak" (1990:396). Here Walton is denying the existence of substantive beings, and is correct in doing so.

Second, I think Walton has not allowed for the division between the *content* of a sentence and its *pragmatic force*, a distinction that was acknowledged earlier [3.1]. We do need a theory of content in fictive practice, and it is one of the most important tasks of the theory of pretence introduced earlier to supply this theory. Real propositions have an important role to play in fictive practice. Consider the case of an *extremely* naïve friend of mine, Peter. Over coffee one morning, I begin to tell Peter about a friend of a friend of mine named Gregor who awoke one day as a beetle, who was horrified at the prospect of living as a beetle, and who was eventually spurned by his family. At first Peter wonders about the truth of the story. Does this Gregor really exist? Surely people cannot turn into beetles? But soon, after being reassured of the truth of the story by another devious friend, Peter begins to believe the story. He even feels comfortable with asserting the proposition that Gregor turned into a beetle, even when other people he tells it to question the story's truth. Some months later Peter comes across a story written by Franz Kafka, and the deceit is revealed to him. Nevertheless, he enjoys the Kafka story, reporting to me how he identified with Gregor's alienation, given all the nasty tricks he is subjected to by his friends. This example shows that what we need a theory of fictive content for is to explain how throughout this case the content of the story of Gregor remained the same even though the mental attitudes toward it varied. Furthermore, it becomes obvious we need a theory of content where content is seen as independent of the pragmatic illocutionary acts that employ it. Though it is fictional that Gregor is a beetle when that utterance is employed as a part of fictive practice, we need to understand how that very real proposition might be used to utter an assertion in a counterfactual situation.

This is all related to the issue of inferential elaboration in Nichols and Stich's work. Nichols and Stich argued that the content subserving our pretenceful practice was prone to inferential manipulation by the same cognitive structures subserving genuine inferences. This was because the content underlying both types of inferences was in the same *code*. This is just to say that the content involved in the belief that Gregor turned into a beetle, might also play a role in the pretenceful premise that Gregor turned into a beetle.

Nevertheless, though there are problems with Walton's theory, his claim that statements about fictional worlds are not true or false simpliciter but fictionally true or false, *does* make an important contribution to the characterisation of fiction. Unlike a work-world that is constituted by all the fictional propositions derived from a work of fiction, a game-world consists of all the fictional propositions that are true in virtue of the game an appreciator plays with the fictive material. Not only are the facts within the work-world (those attributable to the authored work) subject to the semantic qualifications implied by fictive practice, so are those that are attributable to the attitudes of appreciators. Foremost, game-worlds involve the *relationship of the appreciator to the fiction*. It is fictionally true within the work-world of *Metamorphosis* that Gregor turned into a beetle. It would be fictionally true within a game-world that an appreciator instantiates by playing a game of make-believe, that the appreciator *observes* that Gregor turned into a beetle. Furthermore, it would be a fictional feature of an appreciator's game-world that they *reported* Gregor turned into a beetle, or *believed*, *doubted*, *ignored*, or *referred* to this fact. Not only are the contents of fictional work-worlds qualified by a fiction operator, so is the content we generate about ourselves in relation to fictional work-worlds.

Though dogged by the identified difficulty, Walton's use of the fiction operator to qualify the ascriptions of relations between fictional worlds and appreciators and the apparent attitudes appreciators have toward fictional worlds occurs for reasons related to the traditional logical, semantic, and ontological concerns familiar from the philosophy of language. It cannot be true simpliciter I *believe* Bloom likes to eat fried kidneys, no more than it is true simpliciter Bloom likes to eat fried kidneys. To assert either *unqualified* would be, as Walton says, "to tolerate mystery, and court confusion" (1990:196). We do not have "psychological attitudes towards fictional entities" (1990:196); we have *fictional psychological attitudes*. It is the great merit in his work that Walton draws our attention to the fact that not only do the traditional logical and semantic concerns place demands on our account of the content within works of fiction, they also place demands on the way we characterise our relationship to this fictive content.

These are difficult issues that I am not convinced I have solved at this point, or that I will be able to solve in this thesis. I will return to them in a later section [3.5], before explaining how the emotions are involved in fictive practice.

***

For Walton, fictive practice allows artists, through the construction of props that play a role in generating the content of games of make-believe, to take hold of our attention and guide our imagination. By constructing complex and interesting work-worlds with which we participate and construct interpretative game-worlds in which we see ourselves as related to the world of the fiction, artists provide us with rich narrative experiences. There is a long humanist tradition that says that, in doing so, artists are able to convey to us, if not truths, then new ways of seeing the world. Fiction may even play a role in our moral education, as Martha Nussbaum has suggested Greek Tragedy has the potential to (Nussbaum, 1986; 1990). Lamarque and Stein Haugom Olsen (1994) argue against the notion that fictive literature expresses truths, but supply a subtly different account of the source of its value to us. Two recent collections of essays further advance this humanistic perspective (Davies 1997; Levinson, 1998).

3.4 Fictive Practice

My characterisation of fiction appreciation at this stage is incomplete. I have set out the bare bones of a theory concerning how an attitude of pretence allows us to interact with worlds of fiction. Now I need to attend more closely to the manner in which appreciators perceive and generate the content that is central to the pretenceful participation. This section is my synthesis and elaboration on some other theories of fictive participation. The theory presented is intended both to clarify the logical, semantic, and ontological difficulties noted earlier and to provide a substantive psychological theory of processes involved in the appreciation of fictions that explains our emotional interaction with fictions.

A notion that will become crucial to understanding how appreciators participate with works of fiction is that of the *modes of access* participators have to the facts within fictional worlds. There are various modes of access to the fictional worlds presented by different works of fiction. In the most simple cases where fictional worlds are created and presented to participators, such as in a childhood game of make-believe, the content of the fictional world might be presented as an utterance qualified by the invitation *let us pretend such and such is the case*. This is both the pragmatic prompt for the beginning of an episode of pretence, and also the establishment of the *premises* that begin the episode. Following this, if the invitation to enter into the game is accepted, a set of representation tokens are placed in a possible world box, where they are then subject to the operations that constitute a game of pretence or make-believe. The mode of access is comprised by an utterance of fictive content. *Modes of access are the manner in which we acquire the* *premises that base our games of pretence, that is, the manner in which we acquire what is fictionally true of the world of the fiction we are participating with.* In novels and short stories, the modes of access are the descriptions and other language found therein. In films, visual and aural images, as well as the language use of characters and narrators, contributes to the modes of access to the fictional worlds presented there. These are the primary *representational* features of the props at the basis of these works.

I will also argue there are two broad classes of participation that contribute toward the instantiation of our fictive attitudes and generative activities, and so of the *meaning* of works of fiction: *isomorphic* participation and *representational* participation. Isomorphic participation is that class of participation that is apt to being characterised as involvement within a fictional world. That is, isomorphic participation is comprised of the participatory processes that are themselves prone to play the role as props in games of make-believe. This is because isomorphic participation is itself often a *reflexive* prop (Walton, 1990:210-215), making things true of the game-world of the fictional interaction. I will argue isomorphic participation is subserved by important psychological states and processes, interesting in their own right, and of vital importance when I come to discuss the place of the emotions in fictive practice. I suspect philosophers usually have more concern with this former type of fictive activity, and even then, with a subcategory of it, namely pretence or make-believe. But there are other appreciative states, which do not involve pretence, that can be the objects of fictional ascriptions.

Fictive practice involves things besides isomorphic participation with the *worlds of fiction*. Aside from our pretenceful interaction with fictional worlds, we also engage with the features of *fictive works*. We must involve ourselves with the *representational* conventions of fictions. This is what I will call *representational participation*. Representational participation involves engaging with descriptions and images and the other formal properties that add to the meaning and experience of fictions. But a great deal of such interaction cannot be described as an interaction with a fictional world, and this is why I choose to denote it as *representational* participation. So strictly, isomorphic participation is a subset of representational participation.

The distinction between isomorphic and representational participation is entirely based on whether the process is itself prone to playing the role of a reflexive prop in a game of pretence, but I think it is an important distinction to make. It allows us to see that engaging with fictions is comprised of more than an apparent interaction with a *fictional world*, and that much of our emotional interaction in works of fiction derives from this wider sense of participation.

However, despite the obvious importance of representational participation in generating the emotional states obvious in fictive practice, in the remainder of this thesis I will be concentrating primarily on those emotions that are easily fictionally characterised. Among my reasons for doing so is that many of the representational manners in which fictions might generate emotions in their appreciators seem general to art rather than to fictions. For example, the emotional engagement with the music that accompanies many fictions is likely to find its explanation in the literature on the emotionally expressive features of music (Davies, 1994; Kivy, 1990). Furthermore, it is isomorphic participation – that is, the participation with *apparent people and situations* – that seems to me to be the more philosophically pressing issue given the importance in this debate of the paradox of fictional emotions. The following is a discussion of the various types of participation that fit under the isomorphic and representational categories.

i.) Isomorphic Participation

For Walton, though this is not his terminology, *isomorphic participation* would be paradigmatically instantiated by the pretenceful manner in which children treat stumps as bears. In the game, those stumps *really are bears*, and the children's participation with them (in the game) amounts to an ordinary interaction with bears, though, when described from a factual point of view, this interaction must be seen as placed behind a fiction operator. In participating in this manner, children fictionally run from the bears, warn their friends of their presence, or even bravely decide to wrestle with them. I have chosen to call this class of participation *isomorphic participation* because it shares a *structural isomorphism* with our participation with real-world events. Many states of fictive participation, in particular, the emotions that arise, are *dynamically* or *functionally* isomorphic with corresponding genuine states. Alternatively, we might see isomorphic participation as somehow *parallel* to real-world participation, as its form within a game of make-believe corresponds to the form such participation has in the real world.

Another way to signal the distinction is that, in the case of isomorphic participation, but not representational participation, representational conventions particular to fictive practice do not stand between the prop and the theoretical or affective processes of deriving the meaning or significance of that prop. It is in this manner that isomorphic participation does not acknowledge the fictionality of content. There is an available description of the participation that does not make reference to the work of fiction.

Therefore isomorphic participation is readily characterised in a fictional manner as attitudes toward fictions qua people and situations or fictional worlds. This is another way of saying isomorphic participation is participation that is apt to be *reflexively incorporated into our fictional game-worlds* because it is of the appropriate nature to play a role as a prop in a game of make-believe. It is this fact that has made various features of isomorphic participation, foremost *reference*, *assertion*, and of special interest to us, *emotions*, seem paradoxical, metaphysically exotic, abnormal, or irrational. Isomorphic attitudes seem abnormal or irrational because they are structurally similar to genuine states, are superficially talked about in the same ways we discuss genuine states, and yet lack an important component of the genuine states: a belief in the actual existence of the objects to which such states seem to attend. The apparent commitments of isomorphic states when fictionally or internally characterised are quite different from the actual commitments of the appreciators of fictions.

The key to understanding much of the isomorphism that exists between fictional experiences and our engagement with the actual world are the features of pretence that were noted in the earlier section on cognition and pretence [3.2]. Stich and Nichols think the representation tokens that subserve our pretenceful abilities are in the same *code* as those employed in common garden experience. This means our fictive thoughts and sensations will be subject to the same cognitive operations our other representational states are. Fictive thoughts and our common or garden thoughts and beliefs both employ some of the same representational features or capacities of our mind.

However, this is not enough to explain the nature of isomorphism, because the isomorphism seems not to be between beliefs and thoughts in fictive games, but between beliefs and those thoughts involving us in fictional *worlds*. The argument being developed here is that these latter states are only the apparent commitment of language arising from a fictive idiom. When characterised as part of fictive practice, isomorphic participation consists of the attitudes whose descriptions Walton wants to be seen as placed behind a fiction operator; attitudes such as *morally judging a character*, *questioning their motives*, or *deciding a person's voice is soft and clearly articulated*, and, as I will argue, *fearing blobs of slime*.

I think Walton and others would see what I here call isomorphic participation as coextensive with make-believe or pretence. There are two reasons to resist this usage. Firstly, "pretence" implies a *conscious* and *deceptive* aspect that is unhelpful in this connection. If it were not so ingrained in philosophical discourse, I think "pretence" might best be dispensed with because of the unhelpful connotations the term has. These connotations, which "make-believe" shares to some extent, have given rise to misunderstandings. I think one of these misunderstandings may lie behind the objection to Walton's make-believe theory by some thought theorists; namely, that the emotions we have for fictions are not pretences (Carroll 1990, 73-74; Lamarque 1996, 127; Yanal, 1999, 52). This objection involves a mischaracterization of Walton's position [4.2-4.3]. I get the impression that one of the biggest barriers to the acceptance of a pretence theory of fiction is that the term "pretence" implies a conscious deceptive action, and that this is quite different in character to fictive practice. If we are to use the term we should keep this proviso in mind.

Second, and more importantly, there are aspects of isomorphic participation, such as various perceptual and affective reactions, that will resist characterisation as any form of pretence, conscious or unconscious [5.3]. "Isomorphic participation" and "pretence" are not coextensive concepts. Some isomorphic states (remember, those apt to be characterised fictionally) will be comprised of perceptual and affective reactions, and their adjuncts. To call these states pretenceful is to credit them with a cognitive sophistication they do not have. Walton thinks there is a "central appreciative stance" in fictive participation, namely, "make-believe" (1990:190). I am not so sure such a single attitude exists. I believe the mind approaches fiction as it approaches the rest of the world: with a variety of attitudes and stances.

However, it cannot be denied that the most prominent of the isomorphic modes of participation involves what is signalled by the terms make-believe or pretence. Pretenceful behaviour is very common. In childhood games it manifests itself in such peculiar acts as speaking into bananas and pretending to eat mud pies. In fiction, pretence involves participators making *cognitive actions* with regard to fictive props. In games of make-believe, we observe, consider, predict, explain, and judge the events of fictional worlds. These cognitive actions can be described in fictional terms; indeed, in listing them as "observe, consider, predict, explain, and judge," this is what I have just done. We find it hard *not* to speak of these attitudes in other than fictional modes. But underlying these states is a fundamental ability to *enter fictional worlds* through a game of pretence. Though they do not use this terminology, I believe other philosophers of the arts have developed theories that are consistent with the theory of pretence or make-believe offered here, and that we can learn much from those theories.

Feagin devotes a significant portion of Reading with Feeling (1996) to explaining the nature of the mental ability that allows thoughts of descriptions to be replaced by thoughts of *people and situations*. There are three primary attitudes involved in fictive appreciation (Feagin 1996:23). First is the *theoretical* activity through which we arrive at the content internal to a work. Second are our *affective* attitudes toward this content. Third is a *reflective* attitude we take toward our mental attitudes to fictions.[41] Important to the understanding of all three attitudes is the manner in which our thoughts of characters *qua sets of descriptions* are replaced by *apparent* thoughts of characters *qua people*.

Feagin thinks people have a natural ability to make perceptual and cognitive mental "shifts and slides" (1996:59ff). Our responses to duck/rabbit pictures are an example of such shifts. Presented with a single line drawing, people will alternately perceive a picture of a duck or a picture of a rabbit, but not both at the same time. Feagin explains this by the automotive metaphor of "changing gears" (1996:60). When seeing the duck, people are in a *duck seeing gear*; people seeing the rabbit are in the *rabbit seeing gear*. Likewise, I might say that seeing fictional characters as sets of descriptions involves a *representational gear*; seeing them as people involves a shift to an *isomorphic gear*.

---

[41] Walton's use of the term "reflexivity" is not synonymous with Feagin's use of the term "reflective," though they share many similarities. This is because Feagin's reflective states are a subset of Walton's reflexive states. For Walton, reflexivity occurs when a prop prompts the derivation of a piece of fictional content. For Feagin, reflectivity is the characterisation participators make of their episodes of fictive practice, and includes a notion of the "appropriateness" or "warrant" for those states (1996:23). Because these are *prompted* by a participator's own affective and cognitive states, these reflective characterisations will be reflexive in Walton's sense also. Future uses of the term reflexivity will be mainly in Walton's sense of the term because I am more interested in the characterisations participators make of their affective states than in the appropriateness of those characterisations.

Feagin specifically thinks this involves a difference in the way the thoughts are held before the mind. It involves “unasserted thoughts” (1996:78-81), a term, originally attributable to Scruton (1974), but also used by Carroll in his similar theory (1990, 1998). Unasserted thoughts can best be seen as mental content without the cognitive or propositional attitudes that were described in the section on the emotions [2.2]. Generally they will be comprised by propositional content, for example, *Imperial Stout is good for you*, which is not subject to an attitudinal qualifier, for example, *I believe that*, *I hope that*, *I doubt that*, and so forth. In appreciating fictions then, Feagin finds we hold the senses of the descriptions we find in fictions before our minds, uncommitted to the pragmatic force of the utterances, but in such a *mental gear* that the descriptions are of people and situations. This is Feagin’s explanation of the primary act of representational participation that opens up the possibility of isomorphic participation. It trades on a psychological fact about humans: we have the natural ability to shift mental perspectives on perceptual and cognitive content.

Following these shifts and slides into fictive practice, Feagin thinks we *theoretically* investigate the fictional world further, trying to discover what is true of the characters and situations therein. This fits nicely with the “inferential elaboration” Nichols and Stich think is essential to pretence practice [3.2]. The cognitive operations Nichols and Stich describe provide a substantive basis for Feagin’s notion of the *theoretical* aspect of fictive practice.

Furthermore, Feagin thinks the unasserted thoughts that capture this fictional content are the cause of *affective* reactions we subsequently and *reflectively* characterise in a fictional manner, seeing ourselves as related emotionally to fictional characters. Feagin’s particular emphasis is on sympathy and empathy, but I think the general shape of her story will be important when I explicate the broader range of our emotional relations to fictional people and situations. This also fits with the picture of fictive practice I am drawing here. I will soon argue that the cognitive states that comprise our theoretical dealings with fiction elicit affective reactions, which then become mental props that make further reflexive cognitive actions possible. Our affective states become props in a game of make-believe that make further facts true of our game-world, specifically, that we are having an emotion for a fictional character or a situation.

Lamarque has a similar theory, though much of his focus is on the nature of fictional characters. He thinks characters considered from an “external” vantage point are comprised of the senses of sets of descriptions (1996:116). If we are to gain access to the internal worlds of fictions, we are to hold these sets of descriptions before the mind in such a manner that the real world expectations we have of the characters *qua descriptions* are suspended and replaced by expectations of the characters *qua people*. This is the process Lamarque calls “fictional stance” (Lamarque, 1996; Lamarque and Olsen 1994). Lamarque’s solution to the paradox of fictional involvement stresses that the thoughts that inspire our emotions concerning fictions can be best identified with bare *content*; they need not be conjoined with *attitudinal qualifiers* of the type involved in beliefs (1996:116-117). Carroll's theory has similar commitments (1990, 1998).

This is a point of agreement between Feagin, Lamarque, and Carroll that makes their eventual accounts compatible. They all think participation with unasserted thoughts is essential to fictive appreciation. Corresponding to this in Walton's theory is the idea that we play games of make-believe using the descriptions we find within fictive props. Indeed, one might wonder how much difference in doctrine (rather than nomenclature) there is on this matter. I suspect many theories of the fictional emotions that seem to be couched in opposition to each other are broadly compatible on substantive matters.

However, I have a general worry about Feagin's, Lamarque's, and Carroll's theories. Their theories are couched in *negative* terms that may not be the most perspicuous way to characterise the mental attitudes involved. They see beliefs as standardly conjoined with an attitudinal qualifier and subsequently couch a theory of fictive participation in opposition to these states. Carroll and Feagin suggest our appreciation focuses on unasserted thoughts, and Lamarque suggests this content can be characterised as propositions stripped of the propositional attitudes that attend them in the case of beliefs. It seems obvious to me that the thoughts playing a role in fictive practice *are conjoined with their own attitudinal qualifiers*. I think a more positive characterisation of these mental states will be more in line with theory in cognitive science. This is what we find in the Nichols and Stich theory of pretence, where the content subserving pretence episodes is placed in a *possible world box*. This is just to say the content is subject to a form of attitudinal qualifier particular to pretence activities. Rather than seeing fictive appreciation as involving thoughts lacking in the *normal* attitudes, we should positively characterise them as *pretencefully held thoughts* or some such.

This positive framing allows me to fit fictive attitudes into a more natural position in our cognitive economy. The fictive states become less an anomaly and more a cognitive ability in their own right. Indeed, the capacity to *pretencefully hold thoughts* before the mind and the possible adaptive significance of this capacity is becoming a respected idea in cognitive sciences, and especially so in naturalistic and evolutionary psychology. The idea links back to Karl Popper's often-quoted suggestion that complex cognition has an adaptive value in that it "permits our hypotheses to die in our stead." I will have more to say about these ideas in the chapters that follow.

***

Isomorphic participation, then, is in part based on our psychological ability to pretencefully hold thoughts before the mind such that attitudes toward *apparent* people and situations, and the attendant images, self-reports, affective physiological reactions, along with the language use of fictive practice, are forthcoming. A distinguishing mark of isomorphic participation as I wish to use the concept in this work is *that it is apt to be construed in a fictional manner*. When I reason "Bloom thinks Stephen, despite the bad teeth, has the potential to become an all-round cultured man. Something he would both want for his son and himself, but, for the premature death of his son, and his own poetic banality… he has been denied on both counts,” the *reasoning* here does not need to acknowledge the representational conventions of fiction, though I have acknowledged the fictionality of *Ulysses* in coming to adopt a pretenceful stance. The reasoning is parallel to a real world form of such reasoning. Much of the theoretical activity participators enter into in appreciating fictions is isomorphic.

The most fundamental manners in which we derive the *premises* Nichols and Stich think central to pretence is via verbal and written utterances. Thus the simplest isomorphic mode of access to many fictional worlds is the understanding of verbal and written descriptions of fictional events. The fundamental *mode of access* into linguistic fiction is *description*. If Bloom is described as holding his hand up, it is fictional he does so. The content become fictionalised by being held before the mind as part of the fictive stance, and subsequently it becomes appropriate to describe appreciators as hearing or reading of some state of affairs in a fictional world. This theoretical activity provides us with images and understandings about the nature of the fictional world. *It thus becomes fictionally true that we are reading or hearing an account of a fictional world.* Currie thinks that central to this process is the understanding that a “fictional author,” the source of the material we are reading, acts as an interlocutor in our access to the fictional world (1990). Yet we do not have to assent to this theoretical postulate to see that our reading or listening to fictive descriptions makes it entirely appropriate to characterise ourselves as reading or listening to the account of some state of affairs.

Perhaps simple verbal stories are historically prior to the more subtle and complex forms of fictions with which we are now familiar. Language holds a widespread role in propagating fictions, and language would necessarily play a vital role in the historical development of our fictive abilities. We *can* envisage the solipsistic creation of personal fictions, and so fictive ability might be seen to reside centrally on the cognitive ability to envisage counterfactual situations and the like, an ability seen as an original function of Nichols and Stich’s *possible world box*. But this solipsistic model of fictive pretence leaves out much we think characteristic of fictive practice. Foremost, it leaves out the relationship between author and audience. It also leaves out other features characteristic of communicative media. With the advent of language, linguistic fictive *props* became available and this proved to be a watershed in the development of fictive practice. Both Walton and Currie think props are essential in creating a sense of *fictive objectivity,* by taking the responsibility for fictive content away from the participators and investing it in a *fictional world*, and of *narrative richness*, in that features of the props reflexively contribute to the content of the game [3.3]. Language is ideally placed to provide fictive objectivity and narrative richness. In creating fictions, authors construct props that deliver an engaging content because we do not know what is going to happen next in the narratives and because the narratives are often complex and novel fictions. In fact, narrative indeterminacy and narrative richness are among the prime virtues in fiction. The narrative practices within fiction have developed to enrich our cognitive experience by augmenting these features of linguistic props.

In some cases it might be especially appropriate to see the reading of a text as an act of isomorphic participation. In reading epistolary novels, it might be fictional that we are reading an exchange of letters. In reading a fictional history, or diary, it seems entirely appropriate to fictionally describe ourselves as reading histories and diaries. Here the inferential elaborations that play an important role in pretence behaviours share a particularly strong isomorphism with common or garden interpretative processes.

Purely linguistic fictions are culturally ubiquitous and, despite the onslaught of modern electronic media capable of non-linguistic representation, they remain so. Furthermore, technologies such as radio, television, film, and the Internet are often merely used to distribute linguistic fictions. But it is obvious *non-linguistic representation* does also play an important role in generating isomorphic participation with fictions. Walton proposes that representational visual artworks, such as paintings and sculptures, are often fictions. Such representations are props designed to be incorporated into games of make-believe. The representational conventions developed throughout the course of representational art mean artworks have a narrative structure that provides material from which fictional worlds can be constructed. If a coloured blotch depicting St. Sebastian represents a hand being held up, it is fictional that St. Sebastian holds his hand up, and it is fictional that appreciators observe St. Sebastian holding his hand up. That the hand-shaped bit of marble connected to the arm-shaped bit of marble is held above the torso-shaped bit of marble makes it fictional that David is holding his hand up. Our observing all these things makes it fictional that we are so observing.

Drama has arguably provided the world with some of the finest artistic creations ever. A combination of spoken language and action contributes to the manner in which dramatic and comic stage works represent fictional worlds. The presence of real people in the performance of plays opens up new opportunities for convincing isomorphic representation of fictional truths. Actors can use their bodies and movements to represent fictional states of affairs.

Not only can Hamlet be represented as holding his hand up, by an actor merely holding *his* hand up and our perceiving this, the external expressive features humans have – for example, facial expressions, voice, body posture, and movement – might allow a more direct perceptual, even *affective* contact with the fictional world of the fiction [5.3]. In these cases, where we note that the actor is smiling or we have an affective reaction to the actor’s facial expression or vocal tone or to other classes of elicitors, there is no representational convention standing between us and the *meaning* of the event. The actor is able to represent a fictional state of affairs by the natural mode of expression we employ in our everyday lives to represent our emotional states. These are acts of isomorphic participation, apt to be characterised in a fictive idiom. It becomes fictional that we observe Hamlet’s uncertain and pensive look or fictional that we pity him for it.

The Twentieth Century saw the rise of an art form combining a variety of linguistic, visual, aural, and natural modes of fictive representation. Films, and subsequently television, have developed to be the prevalent modern form of fiction. Much of its popularity can be explained by the fact film does combine isomorphic modes of access in a rather realistic fashion. Reading novels requires that we supplement from our knowledge, and *under our own steam*, various truths about the fictional world we are observing. Written fictions demand more of our inferential, non-inferential, and even *imagistic* elaborations. It is not obvious from *Dubliners* or *Ulysses* that Dublin has a persistent bustling sound or that the streets are grey. Yet films can make such matters explicit. A film can provide visual and aural images of all manner of fictional worlds.[42]

With the advent of computer generated imagery and other special effects, even the very different or remote worlds of *Star Wars*, *Gladiator*, or *The Lord of the Rings* can be created. Such representational techniques are all too easily characterised fictionally. As we gaze on in amazement at the CGI images of the Colosseum, it is fictional we see the Colosseum; if we thrill at the (brilliantly depicted) pod-race in *Star Wars: The Phantom Menace*, it is fictional that we do so; as we tremble before the CGI images of the terrifying Balrog in *The Fellowship of the Ring*, it is fictional that we have the *proverbial* scared out of us by the frightening demon.

Some have wanted to make these isomorphic modes of access more explicit. Though I suspect this is a modern myth, and it would certainly prove again Hollywood's taste for the banal, I remember reading that film audiences in the United States in the 1950s were subjected to *smell-o-vision*. Audiences were invited to *smell the stench of the beast from beneath the sea!!* Despite reservations we might have about *smelling the stench of the beast from beneath the sea!!* it can be seen that such a device would merely be an additional isomorphic *mode of access* to the fictional worlds of these B-grade movies. Numerous other modes are conceivable: Taste-o-vision! *Taste the beast from beneath the sea!!* Hurt-o-vision! *Feel the pain as you are beaten to within an inch of death by the beast from beneath the sea!!*

In modern videogames Microsoft has a type of game controller that responds with physical feedback and Sony Playstation 2 comes equipped with an *analogue* controller that vibrates in correspondence with various actions and events in the essentially fictive games. For example, if you are playing a videogame that fictionally involves driving oversized monster trucks around various theme circuits, running off the road will cause a little motor in the joystick, analogue controller, or steering wheel to jerk the controller out of your hand, simulating (albeit roughly) an actual experience of running off the road in an oversized monster truck. These are a further isomorphic mode of access to the content of a fictional world.

---

[42] However, written fiction does have the advantage of being able to express complex and concrete ideas without resorting to a clumsy narrator. This is why we should not expect to see the Spielberg's filmic interpretation of *Ulysses* anytime soon (or why we should positively dread such an interpretation). This also may be why filmic interpretations of existential novels are rare. What could be less exciting than watching a Frenchman sitting in a café swirling a glass of red wine or staring at a tree root, while the narrator observes solemnly "It was at this moment that Pierre recognised the dread implications of his interminable imprisonment in the ineluctable state of *being for others*…"?

Videogames are a new and important development of fictive practice, as becomes especially evident with computer games that have a strong narrative aspect. Steps have been made to explain the aesthetics of videogames (Herz, 1997; Poole, 2000) but there is still much scope for developing the theory of videogames. A genre of videogames, first-person shooters, primarily involve the participant using a variety of hilariously lethal weapons to shoot zombies, mummies, and other undead, as well as turkeys, aliens, mutant hybrids, Elvis Presley impersonators, nazi soldiers, *and any thing else that moves*, to the accompaniment of a techno or thrash metal soundtrack. Many of the things you are shooting at shoot back, attempt to dodge your bullets, and even yell abuse at you. Yet some of the fictions (and more and more of them now the hardware is becoming more sophisticated) contain an element of narrative development, such as solving puzzles, formulating plans, discovering identities, *unearthing the terrible secret that lies deep within the bowels of an alien ship*, in which the player is partially in control. Microsoft's *Age of Empires* series involves the player in building empires and fighting battles with various ancient civilisations. Ask a participant in such a game what is happening as they play and they may tell you "the Britons are attacking me with manganols and bowmen" or some such. This is indicative of a cognitive engagement in an episode of pretence, and display obvious aspects of fictive interaction in a game-world; as such it sits comfortably within the characterisation of fictive practice here.

However, videogames also stimulate their own novel interests. Videogame fictions open up enormous avenues of possibility and this is largely because the fictive props become *responsive* to the audience's participation. For most of their history, fictions have been notably passive objects, ignorant of the games (game-worlds) appreciators play with them. There have been exceptions; such as cooperative role playing games like *Dungeons and Dragons*, and pantomime, but these can often seem closer to games of pretence than to instances of fictive practice. The content of these fictive games is often due to the real-time input of other people. The props at the basis of traditional fictions have been notably *concrete*, lending themselves to an interpretative participation. The modal representational props at the basis of videogames are different in that they are *malleable*. The fictive media employed by videogames and the growing computing power of modern hardware means fictions can now be made that change according to the actions appreciators make in them. Where the focus in most traditional fictions is on interpretation, the focus with videogames is on problem solving, planning, and achieving goals, that is, acting in a fictional world. So our modes of access to the fictional worlds of videogames are not only interpretative, they also allow participators to act in those worlds.

The holy grail of *virtual reality* would be a fictive experience in which we were provided with all the isomorphic modes of access to a fictional world we commonly have to the real world. Such *saturation* in terms of isomorphic modes of access would leave very little up to the imagination, apart from the ways we use imagination in real life. Such an access would be *wholly* apt to be interpreted in a fictional manner.

ii.) Representational Participation

Representational participation is the second of the characteristic modes of fictive participation. There are various other processes besides the cornerstone isomorphic participation of fictive practice that we must use if we are to participate with fictions successfully. Participating with fictions does not merely involve isomorphic participation – even less does it merely involve pretence or make-believe. This is especially so when the fictions we are concerned with are artworks that contain not a little of their expressive effect in features not directly attributable to the fictional world presented. Participation in fictions is comprised of the construction and interpretation by appreciators, of work-worlds and game-worlds, in line with the tacit rules that guide fictive practice. Isomorphic participation is comprised by the apparent manners in which we interact with fictional worlds that share a structure with our interaction with this, the real world. An alternative, negative, way to characterise isomorphic participation, is to note that it is comprised by attitudes that *do not* involve an *apparent* realisation of the fictional or representational status of the fictive material. Representational participation, on the other hand, is participation the characterisation of which does, and must, make reference to the fictional or representational status of the work of fiction. Representational participation, unlike isomorphic participation, is not apt to be characterised fictionally or to be reflexively incorporated into *game-worlds*, because any attempt to characterise it *must make reference to the representational features of the fiction*. Where isomorphic participation has as its *apparent* object fictional people and situations, representational participation has as its object the world of artifice. It concerns descriptions, metaphors, symbolism, alliteration, textures, brushstrokes, and other representational practices and conventions. Representational participation is constituted by various conventional modes of access not only to the factual content of a fiction, but also to its *expressive* content.

There are a variety of conventional modes of access to works of fiction that artists use. The most important instance of representational participation is the understanding that works of fiction license the pretence-bound games of make-believe that comprise fictive practice and that is later characterised as our fictional relationship to people and situations. This representational participation is not the isomorphic participation of treating stumps as bears, but coming to understand the representational convention that stumps are to be so treated. It amounts to the recognition of fictive material via the various cues that identify fictive content, perhaps the clichéd cue being the preface *Once upon a time*. Before an appreciator is able to participate with fictions, they must make the understanding that some description or representation, or the sense of a set of descriptions or representations, or as Walton would say, *prop*, is to be treated as a fiction. This understanding is not isomorphic because it directly acknowledges the fictionality of the material.

Otherwise, entering into the fictive stance would amount to a mistake; we must be cognisant (though perhaps not *conscious*) of the fact that fictive props warrant our adopting the fictive stance toward them. Furthermore, such recognition of fictive status is not likely to be fictionally characterised as an interaction with a fictional world – the mark of isomorphic participation. The exceptions to this will be *nested* fictional worlds where part of the content of a fictional world is the participation with a further fiction. The famous examples of nested fictions are *Hamlet*, and *The Thousand and One Nights* stories.

Feagin thinks a particularly important feature of literary fictions is that they are temporally extended (1996). All experience takes up some time, but literary fictions take up a significant time period, substantial enough so that a reader's ideas and judgements can develop. This allows a considerable flexibility in the expressiveness of literary fictions, both because it means complex narratives can be established and because our own attitudes to this complex substance can develop. This temporal extension is a feature of the medium that adds to its representational form and, though it sometimes can be fictionally interpreted, for the most part the time periods taken to experience a fiction will be significantly different to those within the work-world of the fiction. *Ulysses* again provides an example, as, though it is set in single day, it is unlikely that anyone has ever read it in a single day!

There are also representational conventions that influence our participation with visual fictions such as the paintings of St. Sebastian and sculptures of David mentioned earlier. It would be an act of representational participation to see the properties of marble *qua marble*, as adding to the meaning or expressive qualities of a sculpted figure. If the qualities of the marble add to impression of *resoluteness* or *strength* of the subject, then this judgement is based on representational participation.

In films and television, we participate with the musical score and generate factual and expressive content about the work-world in a way that does not involve pretence; in most cases we are not to pretend the music exists within the work-world. Darth Vader does not hire a special orchestra to play ominous sounding music every time he enters a room.[43] Other forms of representational participation abound. We understand a flashback sequence – sometimes signified by chin-rubbing, a pensive look on the face of the character having the flashback, and the familiar superimposed wavy lines – portrays an earlier period. In Paul Thomas Anderson's movie *Magnolia*, Anderson exploits the visual anomaly of after-image to add to the dramatic effect of the movie. Jason Robards' character, Earl Partridge, is dying of cancer, and in several cut away scenes a dark blank screen replaces Robards' face, leaving a perceptual after-image of his profile. We might interpret this as adding to the meaning of the movie: we come to realise, although his death is immanent, Earl Partridge's life will continue to affect the world, particularly the lives of his son and widow; the after-image is symbolic of this.

---

[43] Walton points out a now familiar joke that might be played with this fact (1990:172*n*). In Mel Brooks' *Blazing Saddles*, there is a scene in which an orchestra that exists in the fictional world is playing music that would normally exist as a score to the action on the screen.

Another interesting manner in which fictional truths might be generated via representational participation is through the use of poetic language. Joyce's stylistic use of language within *Ulysses* often *colours* the fictional content of the work in much the same way as the music of a movie colours the fictional content. Though here, as just mentioned, we must be careful in our identification of what is part of the fictional world of the work and what is not. The soliloquy that forms the final chapter to *Ulysses* is unpunctuated, heaping up thoughts directly upon one another, often with no discernible transition point. It is certainly an important and spectacular stylistic – that is, conventional – move by Joyce. But it also represents features of the fictional world of *Ulysses*. It represents the thoughts of Molly Bloom as she contemplates, rather luridly, her relations to her husband and other men. In the text, the clauses pile up upon one another partly because Molly's thoughts do so too. There is also another notorious passage in *Ulysses* that bears this out:

> …and she let him, and she saw that he saw and then it went so high it went out of sight a moment and she was trembling in every limb from being bent so far back that he had a full view high up above her knee where no-one ever not even on the swing or wading and she wasn't ashamed and he wasn't either to look in that immodest way like that because he couldn't resist the sight of the wondrous revealment half offered like those skirtdancers behaving so immodest before gentlemen looking and he keep on looking, looking. She would fain have cried to him chokingly, held out her snowy slender arms to him to come, to feel his lips laid on her white brow, the cry of a young girl's love, a little strangled cry, wrung from her, that cry that has wrung through the ages. And then a rocket sprang and bang shot blind blank and O! then the Roman candle burst and it was like a sigh of O! and everyone cried O! O! in raptures and it gushed out of it a stream of rain gold hair threads and they shed and ah! they were all greeny dewy stars falling with golden, O so lovely! O so soft, sweet, soft!

The language, as florid as one could find, reflects the lewd content at this point in the novel. Joyce's representational technique enriches the fictional content of the work; the division between fictional and representational features of a work here seems ambiguous. But this does not mean the division does not exist. It is obvious elsewhere in *Ulysses*. In one chapter, the plot is punctuated by newspaper headlines that comment on the action. Here it is obvious we are not meant to take it that the fictional discussion that is occurring in the offices of the *Freeman's Journal* is somehow strangely punctuated by headlines. Another chapter is written almost entirely in clichés. Such features are part of the representational form of the fiction or game.

It is interesting that there are various examples where it is ambiguous whether isomorphic or representational participation is in order and also, whether what we observe is apt to be fictionally characterised. These cases amount to an ambiguity in what is fictionally true of the work-world or game-world. For example, is it true in the world of Woody Allen's *Deconstructing Harry* that Mel, the character acted by Robin Williams, is blurry and out of focus?[44] When the characters in the penultimate act of *Magnolia* sing along to Aimee Man's song, is it fictionally true that they are singing the song? They are fictively portrayed as singing, but one of the characters happens to be on his deathbed and another is unconscious, states in which people do not usually sing! In Martin Amis' novel *Money*, when the main character John Self meets novelist Martin Amis, is it fictionally true of Amis that he is the writer of a novel *Money*? And what is fictionally true of *us*, when regarding Martin Amis, John asks us "Do *you* know his stuff at all?" How are we meant to interpret this? It is not clear what attitude we are meant to apply, as content we would normally expect to be fictional is *prodding us* with its fictionality. These are cases of props that are drawing attention to their status *qua fictive props*. Such features are especially salient because they seem to be *breaches of fictive practice*. These features are symptomatic of much modern art fictions, though we might also note similar techniques were available to Joyce's modernist writing and to Laurence Sterne in whatever style it was *Tristram Shandy* was composed in.

Representational features and fictional content combine to form a whole out of which wider meanings and expressive content might arise. For example, in a scene in the movie *Gladiator*, the happenings in the fictional world – a battle between the armies of Marcus Aurelius and the Barbarian horde – is in its final stages depicted in silent slow motion, with a particularly lyrical treatment of the main theme of the score suffusing the soundtrack; *war is hell*, apparently. Interpreting fiction therefore involves both our representational participation with the expressive and technical features of fictions, and our isomorphic participation with fictional content.

3.5 Fictional and Factual Idioms

I have detailed the substantive nature of fictive practice via a discussion of the Nichols and Stich cognitive theory, and Walton's make-believe theory. There are still very important semantic and ontological issues at stake here, however. The Nichols-Stich theory does give a firm grounding in the psychology and practice of pretence and of fiction, but I also need to clear up the semantic and ontological issues that arise in fictional practice. I touched on these in my discussion of Walton's theory of fiction, but I now need to develop a more sustained characterisation of the philosophical problems of fiction.

---

[44] In fact, this is not quite correct. Mel is a character in one of the central character's own fictions. But the point still holds and we can note that Allen's own character later becomes blurry. Woody Allen is quite aware of the incongruity of giving a *fictional person* the properties a *fictive representation* might have. Indeed, in this we can see what Allen's play with the metaphor "out of focus" amounts to: it is the attribution of representational properties to a fictionally characterised person. The factual and fictional modes of characterisation are being purposefully mixed up for a novel comic effect.

I introduced problems fiction poses for semantics and ontology at the start of this chapter. There is the problem that fictive language seems to refer to entities and situations that do not exist, and are known not to exist, by users of that language. Another semantic concern deals with the apparent truth and falsity of fictive statements. The claim that Leopold Bloom is a canvasser of advertisements for the *Freeman's Journal* seems true, but not in a manner that some of our ordinary factual claims are true. There does seem to be some warrant or utility in applying a truth conditional semantics to works of fiction. The reflexive characterisation of our games of pretence also gives rise to *mental ascriptions* that are semantically odd. What can appreciators mean when they say they *observe* that Frodo is particularly impervious to the effects of the one ring, if they do not even believe that Frodo and the ring exist?

The ontological problems concern the reference to worlds, entities, and characters we suspect do not exist. It is natural to speak of fictional worlds. What are fictional worlds that are not literally constituted by mountains, clouds, and seas? A similar problem arises again concerning the nature of fictional *characters* that inhabit those fictional worlds. Just what are the hobbits, elves, orcs, and dwarves in *The Lord of the Rings*? If these things are meant to have the properties real hobbits, elves, orcs, and dwarves have (!), then they certainly do not exist.

This section is devoted to clearing up these philosophical issues, at least in a provisional manner. To discuss them fully would take me far out of the way of my intention to explain the involvement of the emotions in fictive practice. The main argument in this section is that we need to be careful when inferring pragmatic implications from language concerning fictions, to make sure whether that language arises within a *fictive idiom* or within a *factual idiom*. If we do not realise a particular statement is made within a fictive idiom, we might be tempted to conclude the statement has *assertoric force* or *ontological commitments* it does not. This might mean we will not be clear what semantic or ontological intentions a particular utterance commits to, which could be the source of much apparent metaphysical profligacy and confusion. I believe that by making clear the idiomatic import of the utterances that are used to characterise fictive practice, much of the semantic and ontological mystery identified above, will be resolved.

How then, do games of pretence or make-believe alter the idiomatic import of utterances? Games of make-believe are convention bound practices that contain tacitly understood rules that license language use that on its surface may seem irrational, paradoxical, or at the least non-standard. Fictive idioms become intelligible when we understand they are located within a game of pretence of the type described here. Items of language located in fictive idioms play a role as props in generating the premises of games of pretence. The difference between factual and fictive idioms is not located in the structure or surface features of fictional language, but in the *pragmatic rules* that ground fictive practice. These tacit rules both guide the *adoption* of games of make-believe by participants, and *qualify* the language and ascriptions of attitudes of the participants arising from those games.[45] The tacit rules enable us to treat descriptions as props in our fictional appreciation, so that we come to *observe* that Frodo is impervious to the effects of the one ring, *judge* that this might be related to his uncle Bilbo's similar imperviousness, and even *worry* that, as with Bilbo and Gollum, this resistance might begin to wane, causing Frodo to fall under the ring's influence.

Among the *qualifications* the tacit rules engender, is the suspension of the *real-world assertoric* force of sentences that would normally be seen as assertions when located in a factual idiom. Saying that Dedalus has bad teeth, in the context of a game of make-believe, is neither an actual assertion that Dedalus has bad teeth, nor of the obviously implicated fact that Dedalus has teeth of any sort. Among the tacit rules of games of make-believe or pretence is that the implications the entities portrayed in fiction exist, are suspended, even though many of the other logical implications of the *content* within the fictions continue to hold.

Of course, strictly the logical entailment of content is not dependent upon the propositional attitude that is taken toward that content. The content that Dedalus has bad teeth has the same logical implications if it denied, hoped, doubted, or believed. This is why the content found in fictions is in a sense *literal* and why we can literally assert things *about a fiction* [3.1]. The content at the basis of our representation of the actual world shares a *structural isomorphism*, moreover *identity*, with the content at the basis of these generative aspects of pretence. Fictional and factual content share a syntactic identity. In the mind this structural or syntactic identity may arise because of an identity in the manner in which the representation states are *coded* (Nichols and Stich, 2000). The same content that found its way into a possible world box as a part of an episode of pretence might find its way into a belief box. This would happen if it became clear that what was thought to be a fiction turned out to be a representation of reality. In language, our attitude toward this content does not change its logical import; it changes the utterances' *pragmatic context*, that is, what the content is used to express by the utterer.

This is what is meant by the distinction between *utterance meaning*, and *utterer's meaning*, that is commonly referred to in philosophy. The former amounts to the content the sentence has denuded of any particular pragmatic context or intention to mean on the behalf of a speaker. The latter amounts to the meaning the content has when used in a communicatory act, and thus when placed within a pragmatic context. Deriving an utterer's meaning involves the recognition of the pragmatic intentions of the utterer, that is, understanding what they meant.[46] Utterance meaning may not coincide with utterer's meaning because the utterer might have misspoken, intending to mean something different from what the content of their utterance expressed. Or the utterance might mean far more than the speaker realises. The pragmatic context of factual utterances involves assertion about states of affairs in the actual world, among other things, whereas the pragmatic contexts of fictional utterances amount to an invitation to imaginatively participate with the content expressed. In utterances arising within fictive practice an utterer's meaning is further distanced from utterance meaning because of the pragmatic role of language in games of make-believe.

---

[45] This corresponds to the division in Currie's work between the "attitude" and the "operator" sense of the term "make-believe" (1990:72).

[46] The primary sources for this type of theory of meaning are John Austin's distinction between locutionary and illocutionary force (Warnock, 1980), Herbert Paul Grice's Studies in the Way of Words (1989), and John Searle's Speech Acts: An Essay in the Philosophy of Language (1969).

Here I have the provisional solution to the apparent uses of the notions of *truth* and *falsity* in fictive practice. Statements that Dedalus has bad teeth are not true or false simpliciter; they are true or false *of the fictional world*. This is just to say they are *fictionally true or false*. It is not true simpliciter that Dedalus has bad teeth; it is fictionally true because that sentence is true only within the work-world of *Ulysses*. A *fiction operator* qualifies the truth of the sentence. Though it is not literally true that Frodo is on his way to Mordor, it is obviously false that he is on his way to the corner store for milk; there does seem to be some warrant or utility in applying a truth conditional semantics to works of fiction. There is such a utility because our adoption of an attitude of pretence toward the premises in the fictions *The Lord of the Rings* and *Ulysses* allows us to make logical inferences concerning their content, and subsequently, whether a particular judgement is fictionally true or fictionally false of their worlds.[47]

This semantic characterisation of the contentful states within fictions arises because these states are situated within props, that is, works of fiction, that only derive their *utterer's meaning* via playing a role in games of make-believe. The distinction between fictive truth and truth simpliciter is a distinction based on the pragmatic features of language. Works of fiction, and the language that arises from their audiences, are utterances located within the pragmatic loci of games of make-believe. The tacitly understood rules or pragmatics of games of make-believe govern the generation and semantic status of the content within these props and utterances.

What does this semantic and pragmatic characterisation of fictive language use tell us about the *worlds* of fictions – including the game-worlds and work-worlds – that I have been talking of so freely? We find it natural to talk of *fictional worlds*, and I have already used the term many times in this work, but on reflection, like the talk of fictional entities such as Frodo Baggins and Leopold Bloom, talk of fictional worlds is odd. Not many people take seriously the actual existence of fictional worlds that have the *properties real worlds do*; we do not think it would be possible to literally go to Springfield and have a Duff with Homer Simpson at Moe's bar. It is safe to assume the term "fictional world" is not used as a literal designation that these fictional worlds exist, but is used because it is a term that has a certain amount of utility in our dealings with fictions. Walton thinks apparent references to a fictional world are framed within the *pragmatic rider* "imagine this is the case." Such utterances are made within a fictive idiom, so when we appear to describe the contents a fictional world, we are only pretending to refer to such a world while *inviting* the listener to imagine such a world exists. Nichols and Stich would say we are inviting the other participants of our game of pretence to place the representation tokens we have generated or elaborated from the premise of the game, within their possible world box, and thus to join in our elaboration of the game. But Walton's spin on the situation is obviously concerned with semantic and ontological issues Nichols and Stich are not. Any talk of *fictional worlds* is located securely within fictive idioms, and so no strong factual or ontological conclusions are to be drawn from the expression. This is the reason why Walton thinks that talking about the ontology of fictional worlds is ultimately deceptive (1990:57-58).

---

[47] I have signalled this is a *provisional* theory of the semantic status of fictive language, because a theory of truth in fiction must do more than distinguish between content and pragmatic force, and thus between truth simpliciter and fictional truth. It must also characterise the particular *correspondence* conditions for fictive utterances that make them fictionally true or false. A feature of fictive language that makes this particularly difficult is the apparent lack or reference of the apparent proper names in fictive utterances. It is for this reason that the treatment of proper names plays such a central role in Currie's theory of fictional truth (1990:127-181).

Walton thinks that statements about fictional worlds and entities are not to be read as fictionally attributing the property of world-hood to *something else* that needs to be discovered (1990:396). Fictional statements are not attributions of properties, or claims of existence, at all. They are pretended attributions and claims of existence. For Walton, it is beside the point to ask *what* they are really about (1990:385ff). I think this is the case even though fictional language may literally express content. Fictional worlds and other fictional entities are only the *apparent* commitments of our fictive language use, and if we are tempted to think such things exist in fiction practice, it is only because we have confused language arising from fictional and factual idioms. One of the reasons why the philosophy of fiction can be so opaque, is the constant moving between factual and fictive idioms in ordinary language, something we normally do without explicitly noting the transition.

Another lesson that comes out of this is that the descriptions we offer of our relationships or psychological attitudes toward fictional worlds are also properly characterised as being *fictional*. Our *interactions* with these apparent worlds, whether they are apparent psychological or apparent physical interactions, are only fictional interactions. It is never factually the case that *you understand Frodo is impervious to the powers of the ring*; it is only fictionally the case that you understand this, even though you may literally generate that content. In participating with fictions we *fictionally reflexively* generate propositions about ourselves and about our fictional relationships to the work-world, such as that *we feel sympathy for Frodo's plight*. People are prone to explain their psychological involvement with fictions by attributing mental states to themselves, the apparent intentional object of which is a fictional character or situation. This contributes to the reflexive construction of a game-world, which involves seeing oneself as relating to a fictional character. This process of the attribution of mental states is paradoxical; appreciators are quite aware that the intentional objects of these psychological states do not really exist. As such the ascriptions cannot be true simpliciter. This is what we might call the problem of fictional mental ascriptions.

This problem of the semantics of fictional mental ascriptions is therefore a *species* of the more familiar semantic problems I have already dealt with. But here the situation is more complex because the fiction operator is *reflexive* in that it applies to our attributions of our own attitudes toward fictional worlds. That is, the problem is with our apparent characterisations of our own *propositional attitudes*, which are the currency of common-sense cognitivist psychology. It is no more a simple truth that *I believe* that Dedalus has bad teeth, than it is that Dedalus has bad teeth. It is only *fictionally true* that I believe, observe, or judge this state of affairs to be true.[48] It is a truth located within a game-world I instantiate by pretending I am attending to states of affairs within the work-world of *Ulysses*, and *also* by pretending I am attending to my own attitudes to this work-world. Thus a fiction operator also qualifies any ascriptions of psychological attitudes we have toward fictions *qua worlds* rather than *qua artifice*. The mental ascriptions we make concerning fictions are made within fictive idioms.

A further issue I need to clear up concerns the ontology of fictional characters. An important observation is that the concept of a *fictional character* can be analysed in two ways. Again, there are two *linguistic idioms* in which the concept character plays a role in the explanation of fiction practice. In the first everyday (and notoriously paradoxical) manner of referring to fictional characters, we treat them as ordinary people (or ordinary zombies, beetles, pigs, hobbits, cave trolls, insane computers called Hal). It is in a fictive idiom we say of the fictional world of the videogame *System Shock 2* that the marine operative is trapped on the Starship Von Braun that is itself stranded in deep space and overrun with mutant hybrids. The characters here are people (though some of them also happen to be mutants), and such apparent reference to them arises, on this theory, because we pretend there are people or mutants to which we refer when we express content in the pragmatic context of a game of make-believe.

Yet we may also speak of fictional characters in a less paradoxical manner, that is, in a *factual idiom*. Philosopher of the arts Peter Lamarque calls this the “external” reference to characters in his valuable discussions of the nature of fictional characters (1983:84; 1996:120). Lamarque also introduces the corresponding notion of an “internal” analysis of character in the same sections; however, I will shortly argue this is not the most profitable way to approach the issue. Regarding an episode in James Joyce’s novel *Ulysses*, it is the mode of *external* analysis of character that allows the following observations in Jeri Johnson’s annotations to Joyce’s work:

---

[48] Of course it is true simpliciter that I believe that *in the fiction* Dedalus has bad teeth. Yet this ascription is not metaphysically presumptuous because it does not make reference to people and teeth because one of its clauses is explicitly denoted as belonging within a fictive idiom. The language refers not to people and teeth, but to the very real aspects of fictive practice.

> [The episode] divides neatly between Gerty's narrative and Bloom's, though the two stand in asymmetrical relation to one another. This is no relation of parity. In Bloom's half of the episode, the narrative returns (after the narrative disruptions of Wandering Rocks, Sirens, and Cyclops) to interior monologue. So, Bloom is thus accorded (whatever his vocabulary) the illusion of an integral subjectivity and narrative control. Not so Miss MacDowell. Gerty has no voice of her own. Her narrative proceeds in the third person as free indirect discourse, the lexicon provided courtesy of Madame Vera Verity, Miss Cummins, the litany of Our Lady of Loreto, the Lady's Pictorial, with a passing glance at Walker's pronouncing dictionary.
>
> (Johnson, 1993:899-900)

Here the discussion of the characters of Bloom and Gerty admits an awareness of Joyce's artifice in writing *Ulysses*. Characters regarded or described in such a manner have properties people do not: Bloom has the illusion of an integral subjectivity and narrative control. They also have precedents or archetypes, in that one can talk of the *femme' fatal character* or the *girl next-door character*. We might say of the Stephen character in *Ulysses* that he is based on aspects of Joyce's own juvenile period. These characters are *constructions*, more specifically according to Lamarque, the "senses of sets of descriptions" (1996:116). Thus Lamarque can be seen as a proponent of a tradition that goes back to Bertrand Russell, of fictional entities or particulars conceived of as descriptions. Factually then, characters are features of the *props* at the basis of fictions.[49]

The use of fictional and factual idioms in the case of characters is invariably a mixed bag: whatever Gerty is as a feature of Joyce's artifice it certainly cannot be a thing that has no voice of *her* own. Senses of sets of descriptions are not the type of things that have gender roles of the type *people* have. Again, fictional and factual idioms are switched with such rapidity and ease in ordinary language it becomes difficult to distinguish between the *actual* and *apparent* commitments of the language use.

Furthermore, we think nothing of comparing Bloom to a real acquaintance. Not only do we talk of our relationship to fictional characters, we compare other people to fictional characters. We might think our friend Bob is just like Bloom. Like the mental ascriptions we make in fictive practice, such comparisons need not be seen as irrational or incoherent. Such comparisons are, I think, to be explained by making their pragmatic implications apparent. Comparing a real person to a fictional character is not an assertion that the person bears a similarity to the fictional character, as that would be odd: people do not much resemble sets of descriptions. Instead, the situation is like the use of irony. A piece of content that is not asserted to be true is used to mean something quite different than what the content literally means. It is fictional that Bloom is avuncular and peaceable; comparing Bob to Bloom is an actual assertion that Bob has the qualities that it is fictional we associate with Bloom.

---

[49] Currie, in his discussion of proper names, characterises three manners in which the apparent proper names of characters are used. Currie's first two uses – the author's use (1990:146-155) and the metafictive use (1990:158-162) – correspond to the fictional reference to characters here. When Currie discusses "roles" (1990:171-180), I think he is using "character" in an external or factual sense. Currie's discussion is very focussed on explaining how these apparent proper names function as semantic particles in fictional content.

Lamarque thinks this theory of characters resolves the problem of reference to fictional entities. He holds that when we seem to refer to characters in the *internal* sense (*qua people*) such reference is always analysable into references to *external* features of the fiction:

> My suggestion then is this. When names of fictional characters appear in the context of either an author's use or an informed reader's, that is, when they are used from an external perspective, *the internal references to names, i.e. to the ordinary people in the story, get transformed into indirect references, i.e. to the senses of the names themselves.* In short, proper names of fictional characters, as used by an author or an informed reader, refer only to senses, not to persons or particulars of any kind. The names lose the internal references they have when used by other fictional characters. What that means is that we, of course, cannot refer to fictional characters in the way that other characters can. From our perspective, *we* can refer only to the *senses* of the names *they* use to refer to *persons*.
>
> (Lamarque, 1983:85, emphasis in original)

I have various worries about this aspect of the theory. Though I agree to some extent with the intent of the internal/external distinction, I think Lamarque is too literal in seeing an aspect of fictions that is "internal." In Walton's work, referring to the internal or fictional realm of fiction is *just a manner of speaking*, namely make-believe verbal participation. There really is no internal aspect or ontology of fictions; we just pretend there is. I do not see how "internal references to names, that is, to the ordinary people in the story, get transformed into indirect references, that is, to the senses of the names themselves," as there was no reference to start with. There is nothing to be *transformed*. There was only a *make-believe reference*, and this is no sort of reference.[50]

I think this conclusion is further supported by the fact that our references to characters in the external sense play a very different role in the sentences we use them in, one that is usually not consistent with what should be seen as the pretended usage. Sentences about internal characters (people) cannot be paraphrased or transformed into sentences about external characters (senses of sets of descriptions). "Bloom likes to eat fried kidneys" paraphrased into the external sense in order to avoid reference to non-existent entities, might come out as "The sense of the sets of descriptions Joyce labelled “Bloom” likes to eat fried kidneys,” which is plainly ridiculous. Lamarque would argue that I also need to *translate* the second half of the sentence, which would give us “The sense of the sets of descriptions Joyce labelled “Bloom” is attributed via description the further properties of liking to eat fried kidneys.” I think this loses the sense of the original sentence completely and that readers certainly do not mean *this* when they utter such sentences about the work-world of *Ulysses*. Most people are simply not aware that a character is externally analysable as the *sense of a set of descriptions* and I strongly resist the idea that readers mean anything about these things at all. They *pretend* to mean things about Bloom and his liking for fried kidneys, and do so by expressing content, and not by *covertly* referring to the sense of sets of descriptions. I side with Walton in observing that *realists* concerning fictional reference, such as Lamarque, are mistaken in thinking there was ever any internal reference that needed explaining, or indeed, explaining away (1990:390). There was not: there was make-believe reference.

---

[50] Though again, we will now need to provide a theory of how the proper name, now stripped of its reference to a person, contributes to the meaning of the content within fiction (Currie, 1990).

But this does not discount that people *can* refer to externally analysed characters, speaking about them as items of fictive artifice, roles, or archetypes. Yet this reference is not transformed out of some other *internal* reference, it is merely a stand-alone reference to the aspects of the fictive props that are the basis of our pretended references, inferences, and judgements. Lamarque’s introduction of the *external* sense of the concept *character* is an important theoretical understanding, but his general distinction between *internal* and *external* usage of fictional names seems to me to get things wrong. I prefer to distinguish between the usage of fictional and factual idioms.

Furthermore, I see no reason why Lamarque cannot accept the possibility that we fictionally refer to characters in the manner other characters can. Lamarque denies we can refer to internal characters such as Frodo Baggins or Leopold Bloom in the manner other such internal characters do (1983:85). But *internal characters do not refer to each other*, *we only pretend they do*; internal characters do not really exist! Internal characters only exist as described or referred to in fictive idioms. Thus because we *can* fictionally ascribe properties to ourselves – such as when children, while speaking into a banana, will describe themselves as talking on the phone – it seems to me we *can* be described as referring to internal characters in the same sense as fictional characters are described as referring to each other. In both cases we are only pretending any reference is taking place, and the reference is only the apparent commitment of language use within a fictive idiom.

Lamarque resists the aspect of Walton’s theory that refers to our fictional relationships to fictional worlds, but I do not think the two theories should be incompatible on this score; Walton (in Lamarque’s terminology) is merely noting that our *relations to fictional characters* can also be characterised or analysed in an internal or fictional sense. Lamarque though, tends to see the appropriateness of internal fictional characterisation stopping at the surface of the fictional *work-world*; he does not seem to have much time for *game-worlds*.

***

A prime virtue in the distinction between fictive and factual idioms is that it allows me to resolve cases in which our intuitions about the nature of fictions seem to pull in incompatible directions. The distinction allows me to explain how one could both want Romeo and Juliet to survive – fictionally, it is not nice for *people* to die so tragically – and why we would be a little disappointed with a performance in which they did not both snuff it at the end of the play – factually, the tragic effect of the *plot* would be spoiled if they did not both snuff it. The former statement is bound within a game of pretence. The later factual vantage point takes into account the contribution the characters make to a well-written and emotionally effective plot.

The distinction also explains the confusing features of the passage of comment on *Ulysses* with which I began this thesis. There I said both "Bloom thinks Stephen, despite the bad teeth, has the potential to become an all-round cultured man," and "the passage draws together a great deal of the detail from the previous five hundred or so pages of *Ulysses*, [making] explicit the *thematic relationships* that stand between Bloom and Dedalus" [1.1]. The difference comes from the fact that the former statement was made firmly within fictive practice and therefore concerning the fictional features of the work, while the latter comment took into account the artifice of the fiction, and therefore was a comment on the factual or representational features of the characters. Understanding the nature of the division obviously proves vital in solving this paradox of fiction.

This fits very well with the details of Walton's theory, and I can use the terminology to elaborate on a point he makes. It is a *very important* conclusion concerning these matters, as it will help to bear out the actual relationship between Walton and Lamarque's theories on fiction and the emotions, reconciling what has been thought to be a significant doctrinal disagreement. Following Lamarque's terminology, Walton might say that *internal* responses to fictional characters *qua people* – that is, responses within a game of make-believe – are qualified by being seen as *fictional* responses. In framing ourselves alongside fictional characters and psychologically participating with them, we are entering game-worlds. It cannot be true simpliciter that I believe anything about Bloom and Stephen *qua people*, because they do not exist, and I know they do not. It is fictionally true that I believe such things. Any beliefs or psychological attitudes I have toward fictional characters *qua people* must be fictional attitudes, that is, attitudes within a game of make-believe. *Internal attitudes* are essentially qualified by a fiction operator, no matter whether they are putative beliefs, observations, judgements, or even *emotions* [4.1]. The fiction operator Walton applies to the content within fictional worlds, and our attitudes toward it, are properly seen as a response to the traditional logical, semantic, and ontological problems detailed at the beginning of this chapter. This fact will assume increasing importance when I come to talk specifically about the emotions.

As I suspected, the surface of sentences dealing with fictional people *does not* give the full *pragmatic sense* of those sentences. And fictional sentences are not alone in this given the meaning of *utterances* is underdetermined by their content (Grice, 1989). When I write, "Bloom thinks Stephen, despite the bad teeth, has the potential to become an all-round cultured man," I cannot be taken to be asserting this as a fact about our world. It is tacitly intended to be taken as a report of the content I am involved with within a game of pretence. This is really what I mean in my utterance, as my meaning inevitably makes reference to the pragmatic circumstances of the utterance.

When talking of fictional people, “Bloom thinks Stephen, despite the bad teeth, has the potential to become an all-round cultured man,” will suffice to communicate my proper meaning. I do not need to add the pragmatic qualifications “I am pretending that,” and “In the world of *Ulysses*,” because the language is located within a commonly understood pragmatic framework each of us has learned to use during our enculturation. The pragmatic cues are obvious from the social context of the utterance. A central feature of fictive props is that they *cue* our pretenceful engagement. Just how participators identify these cues is another interesting question that cognitive and behavioural scientists will have concerning fictive participation. But some of these cues are already obvious: when the house lights go down and the pictures begin to flicker on the screen, little cognitive switches are being thrown as the audience members engage in fictive practice.

The use of sentences concerning fictional people is similar to the use of irony. Ironic, sarcastic, or metaphoric utterances, like those concerned in fictive practice, involve a particular linguistic idiom. Ironic utterances also have the quality of being non-assertoric of the literal content. In saying “Bill Clinton is a puritan role model” I am not taken to be asserting this content. What I am asserting is quite different to the literal meaning of the content. Likewise the possible utterance that “George W. is an intellectual colossus.” The type of use to which I am putting these sentences changes the meaning of my utterance in each case. The incongruity of my utterance with the evidence of Clinton’s exploits, or George Bush’s track record on being able to name the leaders of foreign countries, or his many speaking gaffs, is likely to make you suspicious. You would conclude that, rather than intending to assert Bill Clinton is a puritan role model, I was being facetious. Likewise you would not conclude I was making existence claims if I said, “Bloom thinks Stephen, despite the bad teeth, has the potential to become an all-round cultured man.” That I was talking about this long and complicated Modernist novel by Joyce, or that my utterance was part of a critical discussion of fiction, might tip you off to the fact that the utterance was located within a fictive idiom.

The ordinary language that is used to describe our interactions with fictions, both by participators of those fictions, and by theorists when they discuss fictions, is beguiling. It tempts us to talk of the *worlds* of fictions, the *truths* of these worlds, of the *entities* and *characters* that inhabit fictional worlds, and of our *relationships* to these worlds. Any theory of fiction must account for these issues. I deal with them by seeing these pervasive utterances as located in a fictive idiom that like the ironic idiom, does not commit speakers to the assertoric force usually attendant to factual idioms, and that allows participators to make attributions and ascriptions that on their surface seem paradoxical. The semantic and ontological problems of fictional interaction are resolved by introducing the notion of a game-bound participation with the content found within works of fiction.

3.6 Real and Fictional Actions

There is one final topic to discuss in this chapter: action. How is action involved in fictive practice? Given that emotion and action are close bedfellows – Damasio sees emotions as those cognitive states that enable us to act so as to maintain homeostasis [2.4] – this issue will become especially pressing in the discussion of the role of emotion in fiction. There are several issues here. Firstly, is physical action available within our interaction with fictions? Second, can we literally interact with a *fictional world*? Third, because I will argue the answer to this second question is *no*, why are we *not* able to literally interact with fictional worlds? Fourth, what provides motivation for the action that is involved in fictive practice?

First, can we *physically* interact with fictions? The lack of a place for action in our experience in fictions has often been held to be a fact that marks audiences' experiences of fictions as somehow different to their normal experiences. This is especially the case when it is the *emotions* audiences have for fictions that are the subject of debate. I think the idea that our lack of action in fictional contexts marks this experience as peculiar can be a red herring in the debate, only serving to confuse. This is because action obviously *is* available to the audiences of fictions. In childhood games it manifests itself in such peculiar acts as speaking into bananas, wrestling with stumps, and pretending to eat mud pies. Action is a *primary* part of pretence behaviour in children. In pantomime, seeing the crocodile sneaking across the back of the stage invites us to shout to the policeman "He's behind you!!," "Look out! Alligator!!," or a simple incoherent "Arrrgghh!" Role-playing games such as *Dungeons and Dragons* also involve participants in overt actions. When practiced by older folk, particularly in my experience those people who hang around university bars, such role-playing games often give rise to the re-enactment of battles in which participants dress up in medieval garb and thrust fake swords and halberds at each other. In some sort of really convincing *virtual reality* (for example, the holodeck of *Star Trek* fame) a full range of physical fictional interactions is conceivable – in as much as anything from the show *Star Trek* is conceivable. It is also conceivable that live action drama could be performed that made a full range of physical actions possible on the part of the audience. Indeed, even now I expect there are performance groups preparing for the forthcoming season of *Shakespearean Cage-fighting*. In videogames, with force feedback controllers and other such physical modes of control, we are made to struggle *physically* with the monster truck we are fictionally driving, or the skateboard we are attempting to ride, because of our physical interaction with the controls of the game.

I suspect it is a *contingent fact* about the fictional games we are used to playing, and that have subsequently become the focus of the philosophy of fiction, that they do not allow us *physical* interaction. Another way to say this is that the modes of access novels, plays, and films make available to us, do not typically include overtly physical modes of access. In traditional literary fictions pretence involves participators making *cognitive actions* with regard to fictional props. In games of make-believe we observe, consider, predict, explain, and judge the events of fictional works. In principle then, there is nothing stopping participators from performing actions in their dealing with fictions. It is just that the opportunity for action has been circumscribed in the traditional form of fictions, and we have tended to overlook many of the actions appreciators do in fact perform in their participation with fictions. The image of fictional participation as a passive and interpretative spectator sport ignores the activities that go into our games of pretence.

Here is why overt physical action may have dropped out of fictional practice, and thus out of the characterisation of fictional practice that is common in the aesthetics literature. In childhood games of pretence, props are both more *malleable*, and also more apt to be physically interacted with than the props involved in literary fictions are. Childhood games of pretence, the props are often other children. When Johnny runs amok in his cardboard car, the screaming children who run out of his way are props in the game of make-believe that is being played. Action is central to the games children play with pretence. In literary, visual, and cinematic fictions, the props are concrete, and thus not apt to be physically interacted with. They simply are not able to *respond* to our input. Thus in the history of fiction, which has primarily been constituted by interactions with oral story telling, poems, novels, short stories, drama, and more recently, movies and television, the audience has become a passive observer. The role of action in fiction has been subordinated to interpretation.

Yet with fictive videogames, action is being reintroduced because the props in videogames, that is the digital representations of fictional worlds, are responsive and malleable, and are quickly becoming more so as both the technology and game design become more sophisticated, and as the control surfaces of games become more overtly physical. Videogames invite the audience to move from being a passive interpreter, toward being an active participant in the fiction. This shift, where the participant is seen as actively playing a game, may be one reason why videogames have the juvenile image they do. Play may be seen as not a *proper* adult or aesthetic activity; indeed, some theorists might want to resist characterising the interaction with literary fictions as a form of *play*. The art objects that have previously been the concern of aesthetics, and the *aesthetic attitude* that is sometimes seen as being important to their perception, are centrally passive and interpretative (Beardsley, 1978; Stolnitz, 1969).

However, though it is pretty clear fictions can involve us in actions, is it correct to describe participants as literally interacting with, or acting in, fictional worlds? Can participators *literally* physically interact with fictional worlds? The answer to this question is a resounding *no*; because of the semantic features of our descriptions of fictional practice, we must realise that agency within fictional worlds such as the world of the videogame *Red Faction*, is only *fictional agency*. It is not literally true that players can act in the world of *Red Faction* because, literally, there is *no world of Red Faction to interact with*. The *world* of *Red Faction* is only the apparent commitment of language used within a fictional idiom [3.5]. I need to make clear the difference between *real physical interaction in fictional worlds*, *real physical interaction with fictive props*, and *fictional physical interaction in fictional worlds*. The former is *never* possible though the latter cases may be when licensed by the type of fiction involved. As already seen, we can envisage fictions where our fictional physical interaction is warranted; where children wrestle bears, role-players thrusts swords at each other, and videogames players furiously operate the pedals and steering wheel attempting to keep their car on the road. These involve real physical interaction with fictive props, which is later fictionally characterised as action in a fictional world.

Lastly, what is the motivation for the actions that are involved in fictive practice, and that are later characterised as fictional motivations and actions? Nichols and Stich think this an important problem with action and fictional practice [3.2] (2000). The two options they envisage are that participators have desires to act in accordance with the game of pretence, or that participators have "make-desires" of the type argued for by Currie (1990; 1995a). Nichols and Stich argued that because some fictional agents do not have desires – for example, dead cats – it is not *make-desires* that motivate our action, but desires to act in accordance with the game of pretence. I am quite comfortable with this theory, as the alternative, make-desire theory, puts too much importance on the notion that we project into the characters in a fiction or game of pretence, and that we *simulate* their mental states on our own architecture. I am not convinced that mental simulation – even should it survive the current debate over its existence (Currie, 1995b; Davies and Stone, 1995; Stich and Nichols, 1997) – is as centrally involved in pretence as some think.

4. Fiction and the Emotions

4.1 Fictional Emotions

How should I fit the emotional response to fictions into this framework? I have set out the basics of the story Walton and Lamarque tell; filling out and evaluating their approaches will be important for my eventual conclusions. To summarize:

Walton provides a substantive characterisation of fictive practice that suggests, and is integrated with, his solution to the problem of fictive semantics, that is, to the manner in which he explains away the seeming reference to fictional worlds, truths, or characters. It is not true simpliciter that Bloom's wife is having an affair; it is *fictionally true* because that sentence is true only within the *work-world* of *Ulysses*. A *fiction operator* qualifies the truth of the sentence. Furthermore, the fictional propositions we generate concerning the work-world of *Ulysses* will be truth conditional in a way that satisfies our intuition that *Bloom is a cuckold* is true while *Bloom's wife is faithful* is not true.

This is also employed as the solution to the closely related problem of the semantics of fictional mental ascriptions, that is, the question of what it means to say: *we observe the blades of Lionel Cosgrove's lawnmower whir, as he battles the zombies*. We have strong intuitions that such a sentence, like an assertion of the content within a fiction, cannot be *literally true*. The sentence is located firmly within a fictive idiom because of its role in fictive practice. As such, Walton will say such an ascription is only a fictional ascription, and that the mental state seemingly referred to in the ascription – in this case an *observation* – is also only fictional. The problem of the semantics of fictional mental ascriptions is *a species* of the familiar semantic problems. Here the situation is more complicated, though, because the fiction operator is reflexive in that it applies to our ascriptions of our own *attitudes* toward fictional states of affairs. Yet this should not lead to the conclusion that the "fictional" qualifications of attitudes and of propositions have different motivations. These *attitudes* and *propositions* function in a similar way in the appreciation of fictions. Both are *fictional facts* drawn from our theoretical activities and from which we generate the *truths* involved in fictive appreciation. Both seem intuitively to be subject to a truth conditional semantics, and therefore being placed behind a fiction operator. It is no more a literal truth that I *believe* Bloom's wife is having an affair, than it is that Bloom's wife is having an affair. It is only fictionally true that I believe, observe, see, or judge this state of affairs.[51] It is a truth located within a *game-world* I instantiate by attending to the fictional propositions within the work-world of *Ulysses* and also by attending to my own attitudes to this work-world. A fiction operator also qualifies any *attitudes* we have toward fictions *qua worlds* rather than *qua artifice*.

---

[51] I proliferate the verbs here because I think when we talk about our attitudes to fictions, "observe," "see," and "know" will be the more frequently used and more familiar mental ascriptions. When talking about fictions – and indeed in everyday dealings – people hardly ever talk about their beliefs.

It is in this context and for *this reason* that Walton refers to the emotions forthcoming from the appreciation of fictions as "fictional emotions" (1978, 1990). It is not true simpliciter that I fear for Lionel Cosgrove as he battles the Zombies with his lawnmower, it is fictionally true I do so. The emotions are fictional because they are only the apparent commitment of ascriptions couched within fictive idioms. The ascriptions are made within fictive idioms because they originate from fictive practice, and express the elaborative premises of a game of pretence or make-believe. It is not literally true that I feel sad for Bloom, it is fictionally true I do so, for *just the same reasons* it is not literally true that I believe, observe, see, or judge that Bloom's wife is having an affair. The emotions are fictional because their ascriptions are seen as qualified by a fiction operator, due to the fact that they are pretended ascriptions situated within a game of make-believe.

Therefore, and despite the important conclusions Nichols and Stich make concerning the causation of the action involved in pretence [3.2] it is often appropriate to fictionally ascribe *desires* to appreciators of fictions and games of pretence given their literal actions and cognitive attitudes. And participants in games of make-believe or pretence often make such ascriptions when it is necessary to explain their emotional reactions, attitudes, or behaviours.

Here is one of the points I have been working up to: Walton's motivation in calling the emotions we have toward fictions, *fictional emotions*, is to resolve the *semantic* difficulties these mental ascriptions present in virtue of sharing in the wider semantic difficulties all ascriptions of intentional mental attitudes and states have when they seem to refer to fictional entities and situations.

Indeed when attributions of *causal* mental states are framed in reference to fictional people and situations, this problem is again evident, though these cases are less discussed than the apparent intentional attitudes we exhibit toward fictions. If asked what caused their startle while watching Stanley Kubrick's *The Shining*, a participator might respond that it was the sudden appearance of the face of the crazed killer, even though this is not literally true. The startle was caused by the *representations* of the movie [5.3]. As well as fictional attributions of intentional and semantic attitudes toward fictions, appreciators may also make fictional attributions of causal interactions with fictions.

Not only does Walton think it is only fictional there are propositions within fictional worlds; he also thinks any *ascription of intentional attitudes* toward these propositions must be fictional. Walton successfully draws our attention to the reflexive language we use to characterise our involvement with these works, arguing it is also fictional in being located within a game of make-believe. While we might want to quibble with some of the details of Walton's position – arguing, rightly I think, that there are genuine propositions involved in fictive practice – it is essentially fictional that Charles fears the slime.

Furthermore, because Walton's characterisation of the fictional emotions is made for semantic and logical reasons, it is consistent with many of the substantive claims we find thought theorists such as Lamarque making. Indeed, I suspect that abstracted away from its semantic focus and sometimes difficult terminology, Walton's *make-believe* view is quite consistent with a *thought theory* such as Lamarque's [4.2].

This semantic point should not undermine the status of the emotions caused by fictions as *real emotions*. Fiction is one of the most important and ubiquitous ways in which we experience our emotions. Without the emotions we have in fictive circumstances – even though it is only fictional that we have emotions *for* fictional circumstances – our emotional lives would be pale imitations of what they are. Our ability to become emotionally engaged in works of fiction such as *Anna Karenina*, *The Lord of the Rings*, *Magnolia*, *Ulysses*, and *The Days of Our Lives* is something to be treasured.

***

Walton's view of fiction and the emotions amounts to more than this semantic proviso. He also argues that make-beliefs play a substantive role in causing affective reactions, which though not fully-fledged emotions, are often good evidence for the existence of emotions. Walton says horror movie appreciator Charles' "quasi-fear" – the feelings and physiological states that typically accompany emotions – is caused by his "realizing that fictionally the slime threatens him" (1990:245). Just what exactly this causation amounts to is something that remains mysterious due to Walton's unwillingness to deal in psychology, and his denial of the existence of propositions in fictive practice. Yet it seems clear that make-beliefs play a role in generating the affective dimension of our appreciation of fictions.

Here, though, I think we should be careful to keep the logical and substantive issues distinct, something Walton and others do not always do. It is *make-beliefs* that cause the affective reactions, not *fictional beliefs*. Make-beliefs are substantive cognitive states that play a role in generating the cognitive operations involved in fictive practice. Fictional beliefs are not a type of substantive state at all, but are only the apparent commitments of language used within a fictive idiom. As such, fictional beliefs cannot *cause* affect, anymore than fictional reference can be *caused by* fictional entities; Walton argues that fictional reference was no type of reference, but the apparent commitment of an existentially liberal and non-committal manner of speaking [3.3] (Walton, 1990:396). If we are sufficiently clear, however, we can see that *make-believing* in fictive content plays a substantive role in generating affective reactions, or what Walton calls "quasi-emotions" (1990). We hold the senses of the content we find within fictions before our attention via an attitude of pretence or make-believe, and affective reactions, that is, quasi-emotions, are produced.

Furthermore, Walton thinks such physiological disturbances are reflexively incorporated with the content of our games of make-believe and characterised in a fictive idiom, so it might be fictional that we have such physiological disturbances caused by our perception of the situations in fictional worlds. Appreciators note the affect within them, and taking a point of reference from their fictive activities or games of make-believe, fictionally realise they are afraid. For Walton, the mental states elicited or caused by fictions provide reflexive props from which we can draw further fictional truths. In this case, the generation of fictional truths corresponds to the process that makes self-attribution of the emotions possible in ordinary situations, and as such, is an *isomorphic* process of fictive appreciation. Self-attribution plays an important, though often ambiguous, role in emotional episodes [2.2]. Self-attributions of mental states such as emotions in the case of fictions are *parallel* to ordinary instances of emotion attribution.[52]

Make-beliefs play a *substantive and causal* role in producing the physiological aspects of the emotions, that is, the quasi-emotions or affect. And such causation is apt to be described in a fictive idiom and to become a feature of our fictional relationship to the states of affairs in a fictional world. This substantive contention, as far as I can see, is a necessary feature of any theory that hopes to clarify the interaction of fiction and the emotions. It is odd that *thought theorists* have generally not seen this aspect in Walton's theory, most likely because they were distracted by the term "fictional emotion," which seems so obviously *theoretically and substantively radical*. But if that term is understood as a response to familiar logical, semantic, and ontological concerns, it is *far* from being theoretically radical. Indeed, I think it is something any thought theorist should accept. I will shortly note that Lamarque does seem to acknowledge this need.

***

Gregory Currie is another philosopher whose account can be given a similar reading to Walton's, though it also has problems. Currie's The Nature of Fiction (1990) is a work focused from a logical, semantic, and ontological vantage point. After dealing with questions concerning the mode of being of fictions and fictional characters, and the nature of reference to such things, Currie presents a chapter on the emotional paradox of fiction. Here he coins the term "quasi-emotions" (1990:183-

---

[52] The causation of the fictional emotions is clouded by this contention. Le Doux denies such cognitive self-attribution should be seen as definitional of emotional substance [2.2 & 2.4]. Instead he thinks it is the causation of what Walton calls the "quasi-emotions" that comprises the emotionally interesting stuff. Still, we can stick with an altered version of Walton's claim. If Le Doux is correct that subjects often confuse their mental states, it will be a reflexive fact about appreciators of fictions that they sometimes do so. Artists and other fiction makers might already play on this propensity; certain fictions may frustrate our attempts to produce consistent attributions of emotional states by depicting ambiguous emotional situations.

184) which are the emotional states caused by fictions. Currie notes Walton’s use of the same term as similar to his, and further, that Walton’s make-believe solution to the paradox of fiction is similar to his own. I am not sure we should assent to this. I suspect Currie’s comparison rests on a vagueness concerning what the paradox of the fictional emotions is about.

Currie’s quasi-emotions run parallel to common or garden emotions and are apt to be fictionally characterised as emotions for fictional characters and situations. As in all aspects of make-believe, though, there are significant differences in the afferent and efferent profile of the emotions. For example, what goes in is a make-belief, and what comes out lacks an isomorphic behavioural component. These factors account for the two incongruous facts Walton originally discussed in his important early work “Fearing Fictions” (1978): that we do not literally believe in fictional people and situations we seem emotional about and the that we have no impulse to act in accordance with the apparent emotions. Currie thinks it is merely a matter of “labelling” whether we assent to calling these states emotions or not (1990:212).

Currie’s account obviously also lends itself to a substantive reading of make-beliefs as causes of affective reactions. He suggests the following architecture for make-believe emotions:

$$\begin{tikzcd}
\text{Believing } P\text{/Desiring } Q \arrow[d, Rightarrow] & \text{Make Believing } P\text{/Desiring } Q \arrow[d, Rightarrow] \\
\text{Feeling } F & \text{Feeling } F \text{ (or } F^{*}\text{)}
\end{tikzcd}$$

(Currie 1990:197)

Here the functional mental states that provide the cognitive mental substance of the emotions (feelings) in the left hand column are replaced, in the cases of emotions directed at fictional situations, by the “make-beliefs” or “make-desires” represented in the right hand column. Corresponding to feeling *F* (be it sadness, joy, anger, and so on) in the normal emotional interaction is a functionally equivalent feeling *F* – though properly feeling *F*,* because it is a quasi-feeling, that is, caused by a make-belief. The emotions we have for fictions are caused by the cognitive states underlying make-believe.

Currie, in comparing his own theory to Walton’s, argues that “Instead of saying that it is fictional that you pity Anna or fictional that I fear for the governess, we might, without distorting Walton’s position, say these things are make-believe” (1990:209). I think Currie comes very close to blurring the line between the logical and the substantive issues in this way of setting things out. Currie’s suggestion *does* distort Walton’s theory. Walton’s emotions are characterised as fictional because their ascriptions are located within fictive idioms. Currie’s make-believe emotions, given this diagram, are make-believe because they are *caused by* make-beliefs and make-desires. This is a substantive fact, and seems to imply that make-believe emotions are make-believe because *substantively different* from common or garden emotion (though the difference may not be a bar to legitimately *labelling* them emotions). For Walton, *fictional emotions are not substantively anything*; we have fictional emotions only in virtue of a manner of speaking that is made intelligible by games of make-believe. A comparison between Currie's and Walton's theories on this point does not seem apt given their different motivations.

Currie *does* seem to recognise these differing motivations, but this has not worked its way through the entirety of his theory, especially his characterisation of Walton's position. Currie makes the distinction between the notion of make-believe as an "attitude," and the notion of it as an "operator" (1990:210). I think the former sense refers to the substantive cognitive states in fictive practice, and the latter sense refers to the semantic qualifications attached to language arising from fictive practice. Concerning the paradox of fiction and emotion, Currie's concern is with the former attitudinal sense of make-believe, and Walton's is with the latter operator sense. This distinction has been quite central to my thesis here, but has not shown up as a distinction in the notions of make-believe, because I have called "attitude make-believe*"* make-believe *simpliciter*, and have spoken of fictive idioms or operators when referring to what Currie calls "operator make-believe." I hope that in doing so I have avoided the confusion Currie exhibits on this topic.

Indeed, Currie seems to think the emotions he is describing do not need to be qualified by the operator sense of make-believe (1990:210-211). I disagree with Currie's evaluation of his own position. Currie's semantic concern in the book has been to explain how sentences might be fictionally true or false, and a major part of this was explaining how the proper names of characters in a "metafictive" sense – the sense in which audiences of fictions use those terms – act as abbreviated descriptions. So "Anna Karenina," when involved in the sentence "I pity Anna Karenina" acts as an abbreviated description. When I use that name as a semantic item I am literally referring to an abbreviated description, something that allows the generation of content that guides games of make-believe. However, it must still be make-believe – operator sense, that is, *fictional* – that I pity "her" (Currie, 1990:212). Again, abbreviated descriptions do not have the gender people do. Currie seems to have forgotten his original purpose, that of characterising how notions of truth and falsity function *within* fictions. It is an inescapable pragmatic fact that it is fictional that we pity Anna Karenina; even though it is true simpliciter that we have attitudinal make-beliefs involving abbreviated descriptions.

We do have substantively real affective reactions caused by make-beliefs, unasserted thoughts, or pretencefully held thoughts. These states are not the same kind of *things* as fictional beliefs but are only apt to be fictionally characterised as such fictional beliefs. If Currie's diagram is taken to be a substantive characterisation of the production of *affect* by games of make-believe, then it becomes sensible. But then it has nothing to do with the status of fictional emotions. Currie is expressing a form of thought theory, and has different motivations to Walton, subjects to which I now turn.

4.2 Thought Theory and Genuine Emotions

Lamarque confronts Walton on the issue of whether the emotions involved in fiction are fictional or genuine. Yanal characterises this confrontation as that between a “thought theorist” and a “make-believe theorist” (1999). Lamarque thinks there needs to be an important refocusing of questions concerning the apparent fictional emotions. He suggests the crucial question is of the form “What objects are our emotions directed at?” (1996:114), and thinks we should ask, “What are we responding *to* when we fear Othello and pity Desdemona?” (1996:114). Lamarque purportedly rejects Walton’s thesis by concluding “it is in the real world that we psychologically interact with fictional characters” (1996:115). The causal object of the emotions is not a cognitive object that exists within a fictional world, that is, a *make-belief*, but is an *external* quality of the fiction. We are made afraid by thoughts. He makes the observation that “the thought contents derived from fictions do not have to be believed to be feared” (1996:117). What we are made emotional *by*, are, in Carroll’s and Feagin’s terminology, *unasserted thoughts*. Since the causal object of the emotions is not a game-bound make-belief, but a real world unasserted thought, the subsequent emotion is real.

Lamarque makes it clear it is characters referred to in the *external sense* [3.5] we are responding to “when we fear Othello and pity Desdemona” (1996:114); that is, characters *qua senses of sets of descriptions*, rather than *qua people*, cause our emotions. To quote Lamarque, “The fear and pity we feel for fictions are in fact directed at thoughts in our minds” (1996:116). This allows him to avoid what most people see as the metaphysical absurdity in claiming that characters in the former internal sense are causally efficacious. It also means that the resultant emotions are genuine emotions, albeit emotions directed at – or more likely, I will argue, *caused by* – descriptions rather than people. And again there is no reason to doubt this is possible; it is a feature of both folk-understanding and empirical science that imagination can cause emotional episodes (Le Doux, 1993:112; Nichols and Stich, 2000). Thought theorists understand that cognitive states other than beliefs can play the required cognitive role in emotions.

This, then, seems to me the motivation thought theorists have for holding, contrary to Walton’s conclusion, that the emotions we have concerning fictional characters are genuine and not fictional. Yet in as much as this is the case it can also be seen that the two positions are not contrary. The observation that cognitive states based on the imaginative apprehension of the sense of sets of descriptions does nothing to blunt Walton’s conclusion that it cannot be true simpliciter that I feel a little sad Bloom’s wife is having and affair with Blazes Boylan. In fact, this is something Lamarque explicitly accepts. In a postscript to a chapter in which Lamarque sets out his solution to the paradox of fiction, he casually remarks “it can not be literally true that Charles fears the slime, for the reasons Walton gives” (1996:125). Lamarque still shares the logical intuitions that gave rise to Walton’s postulation of fictional emotions, *despite* his conclusion that emotion states are legitimately *caused* by cognitive states other than beliefs. The conclusion to be drawn from this is that Walton characterises the emotions involved in fiction as fictional for quite different reasons to those Lamarque has for characterising them as genuine. *The emotions are both fictional and* *genuine*, though under different descriptions, one located in a fictional idiom, the other in a factual idiom. What disagreement there is between Walton and Lamarque in the literature is attributable to their unawareness of their differing foci on semantic and substantive issues and to their differing terminologies. I suspect Yanal's division between "thought theorists" and "make-believe theorists" involves a terminological difference that is elevated to a doctrinal one, along with a failure to recognise Walton's semantic focus.

This conclusion is further supported by the surprising fact that the force of the thought theorist's argument is something Walton implicitly accepts [4.1]. Walton thinks it natural that his familiar fictional moviegoer Charles "experiences quasi fear as a result of realising that fictionally the slime threatens him" (1990:245). This amounts to the claim that Charles experiences the physiological symptoms of fear because he imagines being in a situation of danger, and the claim is part of Walton's substantive characterisation of fictive practice. That his opponents have not seen that Walton accepts this is due to Walton's often opaque and seemingly *radical* use of terminology, especially the terms "fictional emotion" and "quasi-emotion", which as I will argue in the next section are consistently misinterpreted.

Further evidence for my conclusion is that Walton writes in a footnote that "evolutionary pressures may be responsible for our being organisms of a kind susceptible to quasi emotions in situations in which they might enrich our psychological participation in games of make-believe" (1990:245*n*). Not only does Walton think that thoughts other than beliefs will naturally give rise to the affective component of emotions (quasi-emotions), but that they may do so for evolutionary reasons. Any faculty that allowed us to rehearse social and environmental situations by constructing narratives would certainly need to be tied into our affective capacities if it were to have any motivational efficacy and, thus, selective advantage.

I think that different motivations lead to the contrary conclusions of the two schools of thought. This is most forcefully shown by the fact that Walton and Lamarque *both implicitly recognise and accept the force of the arguments for the opposing views*. I believe another major contributor to the divide between the thought theorists and the make-believe theorists is the vagaries in the use of emotion terms. To this subject I now turn.

4.3 The Usage of Emotion Terms

A valuable exercise in clarifying matters here will be to set out the usage of emotion terms among some of the philosophers involved in this debate. Walton sides with a rudimentary cognitivism of the emotions in taking them to be essentially comprised in part by beliefs (1990:197). This is a characterisation that should be familiar by now; a person's fear, the theory goes, in part involves the belief that they are in danger. Cognitivists do not discount the fact that emotions also involve physiological and qualitative dimensions. But Walton seems to reserve the term "emotion" for states that involve belief; in his view emotions essentially involve an *intentional attitude of belief toward objects* (or *apparent objects*, as in cases where we are not aware our attitudes fail to obtain a referent). This places a wedge between Walton's theory and those of the thought theorists. Walton's reservation of "emotion" for fully intentional states abets the common mischaracterization of his position. Since emotions essentially involve the intentional attitude of belief, the having of emotions toward fictional characters and situations shares in any difficulties that having beliefs about fictional entities has. For Walton, appreciators do not have emotions concerning fictional characters because they are quite aware the objects that would satisfy their having the full-blown intentional attitudes do not exist in the case of fictions. It is fictional, not true, that appreciators have emotions *for* characters.

Yet thought theorists are likely to jump in with the observation that cognitive states other than beliefs will satisfy the cognitive demands of emotion states; *unasserted thoughts* will inspire emotional reactions such as sweaty palms, raised pulse rates, nausea. But this is already to alter the conception of "emotion." And this is what we find in the work of thought theorists. Emotions are not characterised as the intentional attitudes Walton has in mind – they are not emotions *for* characters – but something less, such as physiological reactions to perceived stimuli or to cognitive contents. Thought theorists conclude that fictions can cause genuine affective states, and it is in this sense the emotions that are inspired by fictions are genuine.

Thought theorists, however, have not noticed that Walton is willing to agree that fictions naturally cause genuine affect. This is because they have been distracted by his seemingly radical terminology. The thought theorist's use of the term "emotion" in this connection is almost synonymous with Walton's term "quasi-emotion." For Walton, quasi-emotions are the physiological and qualitative baggage that typically accompanies a full-blown emotion (which centrally involves intentional attitudes, moreover, beliefs). He says of his familiar example Charles, who is watching a horror movie, "His muscles are tensed, he clutches his chair, his pulse quickens, his adrenaline flows. Let us call this physiological-psychological state *quasi-fear*" (Walton, 1990:196). Furthermore, "It is partly the fact that he experiences quasi-fear, the fact that he feels his heart pounding, his muscles tensed, and so on, that makes it fictional that he is afraid; it would not be appropriate to describe him as afraid if he were not in some such state" (Walton, 1990:243).

Unfortunately, many philosophers have concluded either that quasi-emotions were the features of the fictional emotions that make them fictional, or more strongly, that "quasi-emotion" is *synonymous* with "fictional emotion." But neither of these views is the case. For Walton, *real emotions, and those involved in fictive appreciation, both typically have a component of quasi-emotions*. Furthermore, as I have been arguing, the quasi component we fictionally characterise as fictional emotions is *not* what makes them fictional emotions; rather it is the apparent episode's situation within a game of make-believe and its apparent focus on a fictional situation.

Other unwarranted interpretations of the meanings of the qualifications "quasi" and "fictional" are also common. Yanal groups together a set of descriptions Walton would never assent to, when he writes "What I'm questioning is whether the mere fact that it is a game of make-believe that generates the emotion is sufficient to establish that the emotion generated is odd, abnormal, quasi, or make-believe" (Yanal, 1999:52). Here Yanal makes two characteristic mistakes. Firstly he conflates "quasi" and "fictional" emotions. Second, he assumes the "fictional" qualification arises from the fact that the fictional emotions are *different in kind*, or as he says, "odd," "abnormal," or "make-believe," presumably in the sense of pretence as a *conscious deception*.

Lamarque makes the same mistake (1996:127), as does Carroll when he writes of Walton's theory "Purportedly, when we recoil with apparent emotion to *The Exorcist*, we are only pretending to be horrified. But I, at least, recall being genuinely horrified by the film. I don't think I was pretending…" (Carroll, 1990:73-74). Neither Lamarque nor Carroll seem to recognise that the "fictional" qualification is derived from a semantic concern, and is not something you could deny on the basis of your past experience of an event and of how genuinely horrifying it might have felt. Walton notes "We do indeed get 'caught up' in stories; we frequently become 'emotionally involved' when we read novels or watch plays or films. But to construe this involvement as consisting of *our having psychological attitudes toward fictional entities* is to tolerate mystery and court confusion" (Walton, 1990:196; emphasis added[53]). To paraphrase: despite the fact that we become affectively engaged in fictions, to say we had "psychological attitudes toward fictional entities" would be to ignore the fact that fictional entities such as Leopold Bloom do not exist and that we know they do not. Again, *and with emphasis*, to quote putative thought theorist Lamarque, it "can not be literally true that Charles fears the slime" (1996:125).

One of the reasons why the paradox of fictional emotions exists, I suspect, is that, as I explained earlier [2.3], emotions, like other mental states, verge on unobservable posits of a folk-psychology, so that it is an option to simply deny they are involved in fiction. It also makes it seem natural to think that calling the emotions make-believe or fictional is a substantive claim about the psychology of fiction. This is why it seems odd to claim the sexual arousal caused by fictional pornography is fictional, even though in Walton's sense it is. This is not an explanation of the philosophical paradox – the paradox is based on a semantic issue – it is an explanation of the longevity of the paradox of fiction, and how the substantive paradox might be confused for the philosophical paradox.

I think philosophers of the arts should attempt to revise usage of the emotion terms, though this will lead into the complex issues in the philosophy of mind and science I discuss in the second chapter. "Quasi-emotion" is plainly a misleading, not to mention ugly, phrase. One cannot really blame various theorists for misconstruing it to mean *unreal*, *pseudo*, or *abnormal* emotion. I think "affective emotion," or "affective component of the emotions" are more accurate terms. Firstly, they serve to bear out the similarities rather than differences thought theorists share with make-believe theorists: both parties accept that some thoughts will naturally cause affect. Second, this terminology fits well with some current trends in the philosophy of the emotions [2.4]. Ekman's work on the affect programs (Ekman, 1980), Le Doux's and Damasio's work on the neurophysiology of the emotions (Le Doux, 1998; Damasio, 1994, 1999), and Griffiths' theoretical observations (Griffiths 1997) all tend toward taking seriously a revised conception of the emotions as involving a separable independently characterised species called *affect* (or in the case of Damasio, "primary emotions"). Affect is seen as involving short-term patterns of autonomic nervous system arousal, endocrinal changes, expressive vocal changes, and expressive facial changes that are "complex, coordinated, and automated" (Griffiths, 1997:77). They may be elicited by simple perceptual cues and need not be seen as fully intentional states. This certainly fits well with what both Walton and Lamarque are willing to say about the more basic reactions appreciators of fictions have.

---

[53] Walton places "emotionally involved" in scare quotes, further indicating his reservations in using the terminology in this *liberal* manner.

Philosophers are less likely to come to any consistency in the use of the term "emotion" in the debate. Thought theorists seem to think that, already at the stage where affect enters the picture, one is having an emotion. Walton reserves the appellation for fully-fledged intentional attitudes. But his cognitivism is certainly too narrow and this proves problematic for his theory. There are many emotions that do not involve an intentional attitude of any kind. Griffiths (1997) argues that the vernacular usage of "emotion" is not fine-grained and projectable enough to pick out any scientifically interesting kind [2.3]. He suggests replacing the term by three others: "affect," "higher cognitive emotions," and "disclaimed action emotions" [2.4]. Although it is unlikely the folk will adopt this taxonomy, the philosophical and scientific study of the relation of the emotions to fiction may be one area where such a move would be beneficial.

4.4 The Paradox Resolved

There are several conclusions I think I can draw out of this and the last chapter. Firstly, the emotions – fictionally characterised as a subset of putative intentional attitudes toward fictional people and situations – are indeed fictional emotions. It cannot be *true simpliciter* that I feel a little sad for Leopold Bloom. It is *fictionally true* that I do so. Second though, the emotions – genuinely characterised as affective states or quasi-emotions caused by the cognitive states necessarily subserving pretence – are genuine because these things are legitimately caused by cognitive states other than beliefs.

A third conclusion therefore, is that there are *two problems* conflated in this debate, with only one specifically a problem concerning the emotions. The first is the semantic problem of fictional ascriptions of mental states or intentional attitudes such as beliefs, and thus is derivatively a problem of the emotions when they are characterised as involving such intentional attitudes. This is in fact the paradoxical issue, but it is not a paradox of the emotions; it applies to any of the manners in which we, or our language, seem to have intentional or semantic attitudes – or even causal responses – toward fictional entities *qua people and situations*, be the attitude referring, believing, judging, or pitying. It is a species of a traditional philosophical problem demanding that the *apparent* intentional and semantic attitudes to fictional characters and situations be *explained away*. The second problem concerns how atypical things such as fictional situations and characters can *cause* emotions. What are fictional entities that they could cause emotions? The solution to this problem does borrow from the solution to the ontological problem of fictional entities in clearing up the ontology of fictions, and thus providing a *coherent* source of causation. The emotions are caused by the cognitive states subserving the appreciation of fictions. But once this philosophical matter of ontology is solved the problem is one best dealt with by cognitive scientists.

Thus I urge that there are at least two types of question to be answered concerning our problematic interactions with fictions. They are:

- ❑ How is it possible that we frame ourselves with respect to fictional people and situations seemingly ontologically severed from us, such that apparent interactions – such as reference, beliefs, observations, reports, attributions of properties, self-reports, emotions – may take place or be explained away; that is, are to be made intelligible?
- ❑ How is it that we emotionally react to objects not normally considered the proper objects of emotional elicitation and understanding, that is, fictions?

There is nothing that is specific to emotions about the first type of question. It is a broad concern with all attitudes to fictions *qua* the fictional worlds of people and situations. It has, however, seemed to many there *is* something distinctly emotional about the first type of question. This is evinced by the fact that it is commonly blurred with substantive emotional issues. There are good reasons why this happens. Emotions, for *narrow* cognitivists, have traditionally involved beliefs, and the paradox of emotion has traditionally been couched in a manner in which this belief component was put to the forefront [1.2]. Beliefs are psychological attitudes that do seem to demand that we bridge the apparent ontological gap. Sentences like “Dedalus, motherless, and a wanderer, had that day conversed on Shakespeare’s *Hamlet* in the National Library” on the surface seem to assert existential commitments that cannot be literally held. The problem of fictive emotionality is paradoxical only in so far as it trades on *this* paradox.

This paradox is solved by introducing a cognitive stance particular to pretence, and the notion of fictive linguistic idioms, and by making it clear that sentences in fictive idioms are not ontologically profligate because the pragmatic context of fictive sentences does not involve assertion. There is no ontological gap that needs to be bridged, because *literally there is no fictional ontology*. What there are is fictive representations subjected to an attitude of pretence. The cause of our fictive emotionality is our imagination. That anyone ever held to narrow cognitivism, presently staggers me. It seems so obvious to me that imaginings, dreams, fantasies, counterfactuals, unasserted thoughts, whatever you want to call them, will provide a significant set of the cognitive content that elicits emotions. To point out this latter fact is to go a long way to solving the problem of emotional interaction. Lamarque’s solution follows these lines by noting, “the thought contents derived from fictions do not have to be believed to be feared” (1996:117). Subsequently, the emotions are real emotions, but, *pace* Lamarque, this does not count against Walton's conclusion that the emotions are *fictional emotions*; this is correct in so far as the emotions are pretencefully considered as relationships to fictional worlds.

4.5 A Cognitive Theory of Fiction and the Emotions

Identifying and solving the philosophical problem within the traditional paradox of fiction is a prerequisite to providing a characterisation of the involvement of the emotions in fiction. Once we have cleared up the logical matters and made these questions distinct from the substantive concerns already outlined, we can set about characterising the involvement of emotion in fictions. The following chapters of this thesis are devoted to this project. I finish this chapter with preliminary remarks on the substantive issues.

What does our emotional participation with fictions involve? Walton suggests there is one underlying structure: pretence-led games of make-believe. I am not sure this will provide the complete story; hence, my reluctance to see "pretence" and "isomorphic participation" as coextensive terms [3.4]. I intend to conclude this inquiry by giving a detailed description of interactions with fictions that fall under the vernacular heading "emotional," plus of some states that do not so comfortably fit within the vernacular usage, yet which will add to our understanding of this subject. Both episodes of pretence and other psychological relationships and processes will be important to this story.

The vernacular emotions, as I argued in sections 2.1-2.4, may be a variety of mental phenomena. I settled on a practical division between the affective emotions and the social and higher cognitive emotions. If this turns out to be an accurate characterisation, then our theories of emotional involvement in fiction will have to acknowledge this distinction.

It will be important to make clear the particular emotional type that has our interest. It certainly will not be helpful if, noting an *affective* reaction toward a movie, we ask what were the higher cognitive representations or presuppositions of this reaction, as affective reactions do not necessarily fit with this highly cognitive stereotype. Indeed, I think this sort of confusion has occurred in the past. An example comes from a recent criticism of the movie *The Lord of the Rings* I read in which, noting the abundant horror content of this movie, the reviewer asked his son what the scariest part of the movie was. The son responded it was a scene in which Bilbo, again tempted by the ring, attempts to seize it from around his nephew Frodo's neck. The reviewer went on to write that his son was frightened because the scene represented the danger inherent in even the very familiar. The scene scared me too, but I doubt it was because of a highly cognitive appraisal such as that described by the reviewer, and I doubt the reviewer has described his son's reaction accurately. Instead I think I was simply *startled* by the scene. The scene employs a trick familiar to observant moviegoers, in which a sudden burst of stimuli interrupts a quiet passage, causing the appreciator to be startled. As such, this example involves the engagement of our affective responses rather than our more complex cognitive understanding.

Also, if we are discussing *feelings*, we must be aware that “feeling” and “emotion” should not be seen as synonyms. Currie’s solution to the problem may open itself to difficulties due to this inaccuracy (1990:197). Currie strongly identifies feelings and emotions, but having a feeling is not sufficient for having an emotion, and one might have emotions without feelings.

It will be necessary to identify the particular *focus* of the emotion; for example, if we are talking about a case of *fearing for* a character, or simply *fearing* the situation the character is in. The difference here is that the first is a case of sympathy or empathy, and the latter may just be a response to the situation itself. I think both are common experiences. But where the former seems to involve an emotional *relationship* the latter might simply be characterised as a *response*.

Emotional participation will involve the inferential and non-inferential forming of *images* and *understandings* concerning the situations within fictions, and holding these before the mind so that emotional reactions will arise. Reading fictions we often get pictures in our heads of what characters look and sound like. Stephen Dedalus for me will always look like a youthful Joyce. Bloom is on the plump side, perhaps with a large moustache. Often these images are provided within fictions, fictions being replete with descriptions and images. Both images and theorisation will form the *eliciting* conditions of emotions. I should look carefully at the manner of elicitation of affect and at the type of phenomena that will elicit affective emotions.

Some emotions will be readily characterised in a fictional manner – that is, as emotions *for* fictional people and situations – because they are isomorphic to real world instances; others will be more obviously focused on representational features of fictions. In the case where the elicitation of the emotions or the emotional theorisation runs parallel to a real world form, it will be readily characterised in a fictional manner. There will be aspects of our experience of fiction that will not be so prone and these also need to be discussed, however, as they will form an important contribution to our emotional experience of actual fictions, especially when these are considered from the point of view of the philosophy of *art*. Much of the art in fictions derives from their representational features, or from the interplay between these and the fictionally characterised content [3.4]. I should make reference to the many ways in which fictions make us emotional via our participation with the representational features of works of fiction.

Another *substantial* question I might want to answer concerns our abilities in emotional interactions with fictional characters to understand their mental lives or the emotional significance of the situations in which they are placed. This aspect of our emotional participation in fiction has often been the prevalent focus in studies in this topic, though I am sure there are other things to be said about emotion and fiction, particularly concerning the elicitation of affect. The discussion of our fictive emotional theorisation is likely to be parallel to theories of how we achieve such an emotional understanding in real life, but it will also have to account for how this understanding can apply to folk we know to be fictional. The theorisation in this case will run parallel to the psychological participation we have in ordinary understandings. Thus, interestingly in this case, the form of emotional processing involved is likely to be isomorphic to common or garden cases. This is because the form of generation in action in our dealings with fictions and their emotional terrain is likely to be of the sort that provides this understanding in everyday emotional behaviour; that is, mental simulation or theorisation. I will have more to say on this when I come to discuss the higher cognitive emotions [6.2]. But, for now, I will contend that, as far as our fictional emotional understanding goes, make-believe or pretence provides a fictionally characterised object for our attention such as a *person* or *situation*. Our emotional psychology provides the emotion and emotional understanding.

Combining two of the substantive questions identified – concerning our affective emotions and our emotional understanding and theorisation – I might ask how our understanding of fictional folk and their emotional situations makes us feel sad or happy, that is, gives rise to affective reactions. Again, in attempting to answer this question, my focus will be on substance rather than ontological or semantic matters.

The two types of question identified earlier can now be augmented with a third:

- ☐ How is it possible that we frame ourselves with respect to fictional people and situations seemingly ontologically severed from us, such that apparent interactions – such as reference, beliefs, observations, reports, attributions of properties, self-reports, emotions – may take place or be explained away; that is, are to be made intelligible?
- ☐ How is it that we emotionally react to objects not normally considered the proper objects of emotional elicitation and understanding, that is, fictions?
- ☐ How is it that we understand the emotional lives of characters in fictions or engage in the higher cognitive, or social, emotional situations within fictions?

The answer to the first type of question I provided, at least in a provisional fashion, in this and the previous chapter. That of the second type will be explained in drawing on the resources of cognitive science and cognitive neuroscience, evolutionary psychology, and neuropsychology. Its details will be located in a discussion of the nature of both affective and cognitive emotions. An interesting mix of the answers to questions 1 and 2 will answer question 3. All three questions (and indeed many others) will need to be answered if I am to understand the nature of the fictive emotions and also the nature of make-believe and fictive practice.

The emotions involved in fictive practice, though they are emotions arising from the imagination, might turn out to constitute a greater part of our emotionality. This will be the case especially if we live safe lives, getting our thrills vicariously from pretending things. The imagination, where it plays a role in pretence, fantasy, and fiction, affords us the safe ability to indulge in an exciting emotional life. Despite the fact that we should qualify some of the emotion ascriptions we make while participating in fictive practice as fictional, the emotions arising in episodes of imagination are very real. It might be feared that to explain our emotions as I am attempting to do here, will be to undermine their reality and worth. This is not the case. Where we are prompted to fictionally report our emotions for fictional characters and situations, we are indulging in an emotionally vibrant and important practice.

5. Fiction and the Affective Emotions

5.1 Affect and Fiction

Looking at the roles affective emotions and reactions play in fictive practice will go a long way toward elaborating on the simple model established in the previous chapter. It will also explain much that is of interest to the philosopher of the arts, including the phenomenal character of our fictive experiences, in that it contributes to the explication of one of the most puzzling features of our fictive experiences: why they can seem so real. This is because the affective reactions *are real* and make real demands on our attention and other higher cognitive functions. Real affective states are involved in fictive practice.

I suspect some philosophers of the arts have a disregard for psychology. I think if we are going to understand the nature of our emotional interactions with fictional folk – that is, my queasiness while watching *Saving Private Ryan* – we are going to need to understand the mental mechanisms involved. To understand why Spielberg's film is so emotionally gruelling, and how a host of other of our affective emotions toward fictional people and situations are achieved, we are going to have to look closely at what people like Le Doux and Ekman are saying about the emotional mind.

Admittedly, the following observations and examples are made with only the most cursory of empirical backing. They are mostly due to casual observations of fictive practice in others and myself. As with most philosophical endeavours, the examples are not the data of a serious science, since anecdote and thought experiment play a crucial role. It may turn out that the psychological characterisations I discuss here are off the mark. Indeed, as with most anecdotal observations, it may turn out that the states I propose do not exist at all; one of the options when confronted by the paradox of fictional emotion is to deny that we have emotions for fictions at all. I think my observations are testable and that they are no more vague and speculative than some other theories of the fictional emotions. Because they make only tentative suggestions at what the solution to the problems of the fictional emotions might look like, it really is not the end of the world if my theories turn out to be less than accurate. In providing a plausible model of our emotional fictive practices, and one that fits with much current science, such a theory will have gone some way to *demoting a philosophical paradox into a scientific problem*. Such studies open the way for further empirical research, which will elaborate tentative suggestions into working and testable hypotheses.

What then is the scope I intend this explanation to have? Rather crudely, the typical example of a fiction that trades on our affective capacities will be Spielberg's *Saving Private Ryan* rather than Joyce's *Ulysses*. Though this is certainly too crude, because, though the former is notable for the visceral or affective reactions it inspires and the latter for the type of cognitive and social emotional engagement that it gives rise to, both will arouse instances of both emotions. But I think my meaning is clear. The first thirty minutes of *Saving Private Ryan* are powerful not because they drive us to a highly cognitive emotional appraisal of the horrors of war, but because they have us cowering in our seats, wincing, looking, and yet not really wanting to, at the carnage and terror on the screen. I think there is good evidence to suppose that much of this wincing and cowering is based on the depictions eliciting affective responses and that the *horrified attention* such depictions also give rise to may be similarly explained. Affective reactions are short-term, stereotyped reactions, that often give rise to feelings, and that are often elicited by specific cues. Where we find these features in fictive practice, I think we are likely to see the involvement of affective emotions.

There are two general ways in which affect will be important to the story I will tell about our emotional interaction with fictions. Most obviously of importance for the case of affective emotions elicited by fictions is the manner in which imagination and affect interact. As we imaginatively fill out fictions with images and understandings, the cognitive states at the basis of these imaginings will elicit affective emotions in us that will add to the fictive experience and that we will reflexively incorporate into our games of make-believe. Earlier, I found both Lamarque and Feagin reserving a special place for the imagination in the causation of the emotions we have concerning fictions. I also noted that Walton proposes something quite similar, though his term for affect is the highly unhelpful “quasi-emotion.” Nichols and Stich think that, because the representational states that subserve imagination are written in the same “code” as our beliefs, fictive thoughts will naturally inspire emotions. My account in the following section will be consistent with, and should supplement, these accounts of the causation of affect by pretencefully held thoughts. A greater part of fictive emotionality will then involve imagining states of affairs and images, and having these mental contents elicit affective emotions and, subsequently, feelings in us.

Those mental contents that prove important to human life will likely form the greater part of the imagined facts and images that elicit affective emotions in us when we appreciate fictions. We should expect fictions to be filled with opportunities for imagining things about love and social conflict (moreover, about *mating* and *fighting*). The object of our fears and desires should be reflected in the content we imaginatively hold before our attention, as should the novel, the unlikely, the tragic, and the disgusting.

Second though, I will identify an affective stance we might take toward fictions that does not depend on a cognitive engagement with fiction via pretence-bound games of make-believe. There are ranges of simpler affective responses that are involved in fictive participation that do not depend for their effect on pretence-bound engagement. The nature of the interaction will be elaborated by paying close attention to particular features of affect appraisal mechanisms [5.3]. Startle responses are familiar from the affect literature, and though startle is not often characterised as an emotion, we can see this affective reaction does play a role in certain fictive experiences: the sudden looming figure of a zombie at the window causes most of the audience to sit back in their seats.

Some shriek and most feel their pulses racing.[54] As noted earlier, this effect is exploited in a common filmic technique. Similar are the various types of uncomfortable, almost nauseas, feelings that can be induced by disrupting, upsetting, or limiting our sensory capacities. Shaky, unfocused, or spinning visual images in movies can often induce a mild nausea[55] and, when stereo sound is available, multiple and indistinct sound sources can make the listener feel distinctly uneasy. Horror and thriller movies often employ a technique of constraining the visual field in a first person shot to induce a discomfort that is based around our affective discomfort in having our visual field compromised. The appearance of a tight first person sequence in a horror movie is a sure sign the protagonist is about to be attacked by the psycho in the hockey mask. While not all of these effects exploit affective *systems* – it seems likely the cases of shaky or spinning images *disrupt* the normal workings of our experiential psychology – it does seem clear that these types of reaction do add to our experience of fictions and that the theories of fictive involvement just discussed do not pay much attention to them.

Vocal and facial displays are also likely to engender our affective engagement in fictions. There is currently a great deal of interest in the communicative and expressive features of facial displays, especially from an evolutionary focus. The psychological mechanisms behind these displays are automatic and mandatory. In fictions where we have relatively *veridical* access to these displays – in drama, movies, and television for example – such reactions might contribute toward our fictive involvement. Fictions might also affectively engage us when they veridically present the objects of our fears – some unfortunate people are so afraid of spiders that even pictures of spiders induce fear in them – or when they present the objects of our desires – much pornography or more mildly *erotic drama* engages the attention of appreciators in a naïve affective manner. Material that induces disgust will also often be used in fictions.

Damasio’s physiology of consciousness suggests emotion is central to the “enhancement” of the objects to which we attend in our psychological dealings with the world in that it acts to focus the attention and make salient the *important* aspects of the environment [2.4] (1994:173; 1999:193). For Damasio, emotion, or at least what I have been calling affect, pervades all of our higher cognitive activities. Given this, it should be central to the cognition involved in fiction because of its involvement in cognition in general. Our affective engagement with fiction would thus be isomorphic with our affective engagement in the world, because in both connections affect plays a significant and identical role in *enhancing* the cognitive content to which we attend. Could such a theory be adapted toward the psychology of fiction, we would be close indeed to the source of *emotional pretencefully held thoughts*.

---

[54] I think the ability to place startle alongside the other affective emotions may be one of the positives of the approach of this account. Earlier treatments of the fictional emotions often left out such responses, perhaps because they did not fit with the propositional-attitude characterisation of the emotions. I can count it here because the unitary conception of the emotions has been undermined. One of the benefits in undermining the vernacular concept might well be that new generalisations can be made.

[55] The visual technique employed in Hitchcock’s *Vertigo* is of this kind.

This gives the reader an idea of the scope I envisage a treatment of fiction and affect to have. The scope is limited and there is still room for emotions with a larger cognitive component, and less short-termed stereotypicity, to play a significant role in our fictive experiences. Many theorists seem to have a general resistance to theories of fictive practice that see the emotions as *reactions* or as *instinctual*. Such accounts *are* crude and incapable of capturing the cognitive richness of the full range of our emotional experiences with fictions. Yet I think that allied to this is a humanist worry that reducing our responses to instincts seems to sell short our moral and intellectual engagement in fictions, something that is prized by many theorists. Perhaps this is why philosophers of the arts seem often to have downplayed the physiological or affective side of our engagement in fictions. My account is not subject to this worry, both in that it does not see affective engagement as the *complete* story, and in that it sees our affective engagement with fiction as providing substance for our further cognitive engagements, just as it does in our interaction with the real world. In as much as this array of affective involvement in fictions proves to be the case, fictions that employ such representations, will involve affective reactions in *engendering* fictive practice, and elucidating these affective reactions will add to our understanding of fictive practice.

5.2 Affect and the Imagination

The imagination as involved in fictive practice causes affective reactions. As mentioned earlier, Feagin thinks three attitudes essentially comprise the aesthetic appreciation of fictions (1996:23). Without the presence of these three aspects, rather than *appreciating* fictions we would merely *react* to them. The following sections warrant a little revision of her views because of the shared themes; her talk of *theoretical* activity disguises the fact that it is foremost a theory of the involvement of *imagination* and *emotion* in fiction. The first, theoretical component, allows us to derive from fictions the content to which we are to adopt a fictional attitude. Nichols and Stich call this the inferential and non-inferential elaboration of the premises involved in episodes of pretence. Engaging with fictions demands we theorise from the obvious *fictional facts* within such works to discover their implications and perhaps deeper truths of fictional worlds. Much of this theorisation is isomorphic, in that it can be described using fictional terms, as I have done here in describing the process as one of finding out the deeper truths of fictional *worlds*.

Second, the content we derive from fictions through this theoretical appreciation will be the cause in us of *affective* responses. Once I have reasoned from the text of *Ulysses* that Leopold Bloom is in the position he is, such pretencefully held thoughts might elicit my affective sadness, or, if I realise his equanimity in coping with his immediate situation, they will elicit my affective joy. These affective reactions to pretencefully held thoughts are also isomorphic in the sense designated earlier [3.4].

A further *reflective* cognitive component responds to these affective attitudes. It is at this point I realise I am sympathetic with Bloom's situation and admire his attitude toward it, and I think about the warrant in my being so affected. Because of the isomorphism inherent in my theoretical manipulations and affective reactions to pretencefully held thoughts, I am able to characterise these aspects of my psychology that draw me into the fiction. I reflexively describe aspects of my own psychology as being fictional of a game-world. This process fits well both with the story I provided in Chapter Three concerning the nature of fictive practice and with things Le Doux and others say about affect and our cognitive attitudes toward it. Le Doux and others argue that we attribute to ourselves emotional states on the basis of the physiological feedback we notice in our consciousness and the subsequent salient features of the environment we are in. Such attributions are not absolute evidence for a particular emotion having occurred (1998:66-67) and, *pace* Schacter and Singer and propositional attitude theorists, are not to be seen as the substance of the affective emotions: the attributions are cognitive mental actions, not emotional states.

Feagin's theory amounts to a story about our imaginative involvement with fictive material via which we generate further fictive material, our affective reactions to the imaginative states, and a further imaginative involvement with such psychological states, reflectively characterising them in terms of our fictive involvement. Admittedly the story may be vague – the nature of the *mental shifts and slides* involved in our theoretical activity is *underdescribed* – but I think the general shape of the theory is correct. What remains for me to do is to fill in the details, explicating exactly how our imaginative and theoretical involvement in fictions gives rise to affective emotional responses.

Of supreme importance in understanding our affective responses to fictions then, is the realisation that thoughts, concepts, and images are also a legitimate source of affective arousal. Beliefs are not the only cognitive states that will contribute to affective arousal and any rudimentary cognitivist model of the emotions that contends the contrary should rightly cause us to reject that model. Beside beliefs, images, counterfactuals, fantasies, and even simple perceptual features constitute the eliciting conditions for affective states. Physiological evidence for this is provided by Le Doux's contention: "The anatomical connections suggest that [the amygdala] can be activated by simple features, whole objects, the context in which objects occur, semantic properties of objects, images and memories of objects, and the like" (1993:112). The physiology of the brain allows that, besides beliefs, other complex and not so complex mental representations will elicit affect. It has also been previously argued that affect can be elicited non-cognitively [2.2].

One way to read make-believe or pretence was to see it as a process of generating fictional propositions about fictional worlds to which authors and artists provide various modes of access via fictive props. "Propositions" may not be an accurate term to characterise the cognitive states thus generated, since much of the generation is comprised by the forming of images and associations, expressive features and responses, not truth-evaluable content. But the products of these processes – truths, images, understandings, associations, and expressive features – though they do not refer to or represent (fictionally characterised) things that exist in this world, are, because of the structure of the imagination, wired up to produce real affective responses. This, as Lamarque notes, explains a lot about our fictive emotionality; that is, what it is factually we are made emotional *by*. We are indeed made emotional by thoughts and other such cognitive states; centrally, those involved in pretence.

***

How, specifically, does imagination cause affect? What does the process of imaginatively elicited affect amount to under the explanative schema set out in the Nichols and Stich cognitive theory of pretence that has been adopted here? The development of this theory is intended to be one of the most important parts of this thesis, characterising as it does, the substantive causation of emotion by imagination. Elaborating on the Nichols and Stich theory of pretence I can develop a story of how the fear and other emotions we experience while participating with fictions is generated. I think the following profile can be extrapolated to cases where appreciators have the affective reactions of fear, disgust, surprise, anger, and even laughter or sexual arousal while imaginatively engaged with fictions.

A representation token placed in a *possible world box* will not have an identical dynamic profile in terms of behaviour motivation and cognitive and inferential implications, as when a similar token is placed in a belief box; the two states inhabit different functional roles in the mind. However the two states will display a measure of functional equivalence because of an isomorphism in the *coding* of the states. Nichols and Stich would hold that the content contributing toward the admittedly unlikely belief that the mutants have taken over the spaceship is inscribed in the same manner or code as the content that contributes toward the pretenceful representation that the mutants have taken over the ship. The functional differences between the two states do not derive from the fact they are *written down* differently, but that they are written into different functional boxes. This is the feature of pretence cognition Nichols and Stich use to explain the inferential elaboration that occurs in games of pretence. The cognitive states at the basis of representation of the content we find in the fictive videogame *System Shock 2* are prone to inferential manipulation by the cognitive mechanisms that subserve inferential manipulation of beliefs and desires, because they share an isomorphism with these genuine representations. This is why we can infer from the fact that the mutants have taken over the spaceship, that most of the crew are dead, and indeed, *mutantized*!

Things become a great deal more complicated when I add in the fact that, like many other forms of fictions, videogames are modal fictive props. Their contents are not just sentences from which we derive the content of these fictional worlds, as in literature. The content is also derived from visual and aural perceptual experiences. It is not just the *content* (in a specific philosophical sense) of the fictional world that might cause people to have emotions, but the *look, sound, even feel* of those worlds. That is to say the fictional worlds of modal fictions have *qualia* and other perceptual features. These states will also have their cognitive significance because of their isomorphism to factual representation states that facilitate our interaction with the actual world. These qualities will have their psychological significance in our fictive practice, because of the psychological significance they have in our common or garden experience.

The isomorphism our pretence cognition holds to our common or garden cognition also explains why make-beliefs or pretence states can cause emotional reactions. The representation tokens involved in pretence, such as those subserving the fictional content that the mutants have taken over the ship and killed the entire crew are also *available* to our affective or emotional systems. The notable affective reactions and feelings caused in games of pretence and in fictional practice arise in part because tokens of pretence states are syntactic isomorphs with belief states, or of the perceptual experiences under concern.

However, the theory needs more than a syntactic identity in the representational states subserving fictive cognition to explain how affective reactions are caused. One thing that is needed is the *functional profile* generated for a representational inscription by placing it in a possible world box. It is not merely enough that the representational states involved in pretence be *available* to our affective systems, since many representation states with an identity in terms of content will not have the same affective potential. Beliefs, desires, doubts, and intentions may all append to the same content, but will have different emotional significance. Some attitudes will simply lack an emotional significance when conjoined with content that in other cases may prove emotionally affecting. Compare the difference between a belief that you have just won an enormous amount of money in a lottery, and a desire to have just won an enormous amount in a lottery. An important *prerequisite* of our emotional engagement with fictions is that the coded states subserving our pretence representations need their functional role in our mind to be established as *relevant* to our affective systems. This relevance allows our emotional involvement in fictions. I suspect that just why our pretenceful representations are emotionally relevant, owes to the role of pretenceful representation in planning and forethought [7.1].

However, in the process of engaging with fictive representations we also need to be *sympathetic* to these representations if emotions are to arise. Our sympathy with fictions *causes* our emotional involvement. This is one of the real points of the story Colin Radford used to introduce the paradox of fiction in his influential article on the topic (1975) [1.2]. Radford described how being *deceived* by a story and later recognising the deception, and being *told* a fictional story, can cause the fictional story to have a different emotional effect on us even though in each case the stories share an identical content. This is because the sympathetic attitude differs toward the content in the each case, even though in both cases the content is recognised as fictional. And there are many other attitudes we take toward the content of fictions *qua fiction*, including a *critical attitude*, where emotional reactions are not forthcoming.

This sympathy may be a matter of *paying attention* to the relevant parts of the fiction, but may also in part be constituted by being *indulgent* concerning the representations. I might then want to further subdivide our fictive practice into those states in which we are *wholeheartedly and sympathetically* interacting with the fiction, and cases when we are merely critically participating with the fiction. In the latter cases, though we are able to derive all the information from the fiction, we are not affectively engaged with the fiction.[56] This also shows that, despite what Nichols and Stich contend, our pretence representations generally *are* quarantined from our emotions [3.2]. Emotions are not an inevitable outcome of holding fictional thoughts before the mind. There are many instances in which we can engage with thoughts, knowing they are only fictional, in which emotional reactions are not forthcoming.

One very crude option in solving the paradox of fictional emotions is to claim the emotion is not quarantined from pretence states, perhaps because affective states are modular to the higher cognitive appreciations of the fictive material. It might be contended that, though we do not literally believe in the existence of fictional people and situations, emotions are nevertheless caused because the affective systems at the basis of these emotions *are not aware of the fictional status of the eliciting material*. Perhaps this is because these affective states are *modular* or *informationally encapsulated* from higher cognition. This is unsatisfactory as a solution because in *most* cases where fiction arouses emotion, appreciators must realise the fictional status of the experience before the emotion is forthcoming.[57] I say *in most cases* because I believe such ideas are important in explaining issues like startle, disgust, arousal, laughter, and other affective reactions often naively elicited by fictions. I discuss these states shortly, but argue they explain which participators become emotional concerning fictions only in a limited range of the cases. It is only when sympathetically participating with the fictional material that appreciators experience the full range of the emotions I have been discussing here.

One reason I might be tempted to say our pretence cognitions are not quarantined from our emotions is that our emotions often seem to persist long after the fictional participation has ended. This seems to show that our fictive emotions are prone to *spilling over* into everyday life. When I was a child I was scared for days after being allowed to watch the first half of the horror movie *The Omen*. A reason for this may be that children *do* often mistake fictions for fact. Another reason is that we simply sometimes cannot stop thinking about a particularly forceful fiction, though often this may be just because that fiction caused very strong and persisting emotions in us. I suspect this persisting emotion also sometimes occurs because our affective emotions are *trigger-happy*. Recent research shows that parts of the subcortical brain are biased toward provoking emotional responses, especially when elicited by species specific cues, despite the evidence of the more complex representations of the eliciting situation housed in the neocortex, and that emotions can be irrationally clued to various stimuli and persist despite our cognitive appraisals [2.2]. When clued to specific eliciting stimuli this is the cause of phobias. When caused by perceptions of inner bodily or mental sates where the *residue* of the internal state may elicit further affective states, panic disorders are the result (Le Doux, 1997:252-263). This may explain the persisting emotional states fictions can arouse, especially in children, but it does not suffice account for the majority of our emotional reactions to fictions, where I have argued a sympathetic attitude contingent on our understanding the fictionality of the material is necessary. Another related rationale for our fictive emotions persisting well into our everyday life may be that the emotions we have for imagination are everyday emotions, and may be involved in guiding our everyday behaviour [7.1]. The persisting emotion may simply be emphasising the significance of a noteworthy fiction.

---

[56] Cf. Feagin, 1996:23ff [3.4].

[57] Or more correctly, appreciators must realise the fictional status of the experience before *the right type* of emotion is forthcoming. Consider the case of two people, both playing a videogame such as Microsoft's *Flight Simulator*; one realises the simulation for what it is, while the other thinks the events portrayed are real. Stalling the plane on take-off is likely to arouse quite different emotions in either player: in the former, mild annoyance, in the latter, unconditional terror.

Recent research has also shown that the affective systems identified earlier have a distinct *afferent* and *efferent* profile. On the afferent side being an *appraisal mechanism* focused on producing affective reactions cued to specific *elicitors*. On the efferent side is a program of physiological and neurological effects, plus, at least in our species, a feeling (Damasio, 1994, 1999; Ekman, 1980; Le Doux, 1998). The same research shows these elicitors can be constituted by highly complex cognitive states representing complex situations or objects, semantic states, and abstract relations between objects. Here I get to the bones of the theory. My claim is that the representational states subserving pretence and fictive practice are *just these highly complex cognitive states representing complex situations or objects, semantic states, and abstract relations between objects*. If affective reactions could be elicited on the basis of these representation tokens while placed in the functional states of a *possible world box*, then this would be enough to satisfy that pretence thoughts could cause emotions. Indeed this is what is needed because it is patently clear that pretence thoughts *can* cause affective emotions. This is the solution to the first substantive problem of how fictions cause emotions: the complex representation states subserving pretence act as the elicitors of affective emotional states.

In feeling sad and fearful as we read or watch *The Lord of the Rings*, the representational states that subserve the contents that Gollum is obsessed by the ring, is struggling with this obsession, and is also a threat to Frodo, which is drawn from the representations in the book or the movie, act as an *elicitor* of an affective reaction. One of the common adjuncts of the affective state is a feeling. Real emotion is elicited, and thus we really feel sad or fearful, even though we are quite aware Gollum and Frodo do not exist. The affective states are caused because the representation tokens that instantiate our games of pretence are cognitively and emotionally significant, *qua imaginings*. This is just to say our imagination plays an important part in our emotional life. Literal worlds are not the only causes of our emotions; our involvement in the *worlds of imagination* is also emotional. Our imaginations described literally so as to lose the vague and ontological profligate reference to imagined worlds, are just another aspect of the real world, of course [3.3&3.5].

A similar story can be told for the fear I felt when I watch the first thirty or so minutes of *Saving Private Ryan* or what Walton's fictional movie spectator *Charles* feels when confronted by the fictional slime (or should that be the *fictional* fictional slime?). Though again, in these cases we need to take account that, because the *modes of representational access* are also visual and auditory in many modern fictions – as well as many older ones such as live drama – the sensory modalities processing the inscriptions subserving the representation of the fictions are going to be different. Nevertheless, if we substitute a visual or auditory image for the propositional content, the explanation of the previous paragraphs can be adapted toward these modes of access to fictional worlds. Again, the imaginative episodes would be able to cause affective reactions because they are coded to be available to our affective systems, because of the functional salience given them by placing them in the position of *pretencefully held thoughts*, and because we are sympathetic to them. Once we begin to incorporate these reactions *reflexively* into our games of make-believe, things get more complicated, however [5.4].

***

What types of imaginative content will elicit affective reactions? There are many types of pretencefully held thought content that will prove affectively engaging, corresponding, perhaps, to all the forms of affective emotion. Disgusting fictions – recall the infamous *vomit sandwich* – will engage our affective emotions. The objects of our fears, when fictively presented, will elicit affective emotions. So too will the objects of our desires represented in the sexual fantasies. The popularity of erotic fiction will trade in a large part on its ability to arouse in its readers affective reactions. It does this is because the pretencefully held thought content is of a type significant to our affective systems. The saddening and surprising may elicit our affective emotions when presented within fictions. Even anger may be elicited, for example in the case of first-person shooter videogames, where much of the appreciator's engagement with the fiction is tempered by a persistent anger at being killed by yet another grotesque shambling zombie.

Humorous fiction might also prove telling. Jokes and humorous stories, though often fictional, do inspire what might be called affective reactions in us. While laughter has not traditionally been called an emotion, we can see it is a psychological state that introduces many of the same issues for the philosophy of fiction. Indeed, I might frame a version of the paradox of fiction in terms of humorous fiction: how can I find the story of the *nun, the Irishman, and the octopus* funny, when I do not believe the nun, the Irishman, and the octopus exist? That this form of the paradox is hardly ever noted might be due to the fact that jokes involve a very *circumscribed* form of fictive practice, in which many of the forms of fictive appreciation that seem to make the paradox obvious, such as reflexive self-attribution of mental states and relationships, are not warranted. I think a characterisation of what is going on in the appreciation of fictive jokes and humorous stories may share many of the features of the theory I have just developed concerning fictions in general. Fictive jokes are engaged with in something like a pretence-bound cognitive attitude of make-believe, in which the cognitive states produced are apt to cause the psychological reactions that constitute laughter and the other humorous reactions. Laughter, like the affective emotions, likely involves innate dispositional mental states and is elicited, often unconsciously, by certain ranges of stimuli (Provine, 2000). Pretencefully held stimuli, or naïvely perceived stimuli, are likely to cause the reaction of laughter in much the same way pretencefully held content inspires affective emotional reactions.

I argue in the next chapter that people also affectively react to the higher cognitive emotional situations around them and to those found in fictions. It is at this point already obvious why people and their higher cognitive practices and beliefs are important to our success or failure in conducting our lives. Affective reactions to fictive cognitive emotional environments trade on this ability. Currie, in discussing the notion of “secondary imaginings,” thinks as much when he finds that part of the attitude of make-believe is the forming of understandings of what the fictional folk in representational artworks are feeling and thinking about the situation with which they are in by thinking ourselves into their situation (1995b; 1997a:54-56). An explanation of how we understand and react to such features in real life will provide an answer to how we achieve it concerning fictions and there is certainly evidence to show that affect programs are involved in making these things experientially salient. The affect programs may play a constitutive role in the higher cognitive emotions and will often be elicited by representations of social and higher cognitive emotional situations [2.4].

5.3 Naïve Affective Reactions

Lamarque, Feagin, Walton, and Nichols and Stich, identify a cognitive stance as the basis of our involvement with fiction and games of pretence. I would like, though, to identify within fictive practice a set of affective responses that, while being *naïve* or *causal*, nevertheless contribute much to our emotional interaction and enjoyment of certain types and aspects of fiction. In an often quoted paragraph of Darwin’s, I have the seed of an explanation for at least some of my *Saving Private Ryan* queasiness.

> I put my face close to the thick glass plate in front of a puff-adder in the Zoological Gardens, but with firm determination of not starting back if the snake struck at me; but, as soon as the blow was struck, my resolution went for nothing and I jumped a yard or two backwards with frightening rapidity. My will and reason were powerless against the imagination of danger which had never been experienced.
>
> (Darwin, 1872/1999:43-44)

Fodor’s example of eye poking is another instance (Fodor 1983:71-72). It is a familiar experience: sometimes behavioural forces just tend to take over, though Darwin’s identification of this faculty as an “imaginative” one may be off the mark. Both Le Doux (1998:112, 175) and Walton (1990:241) cite this passage from Darwin in their works, but for different reasons. Walton cites it in connection with the imagination and affect, Le Doux in connection with involuntary affect and its behavioural and expressive adjuncts. It seems to me Walton makes a little more than he should of Darwin’s use of the term “imagination.” Instead, what is at work is the *quick and dirty* nature of the various affective systems (Le Doux 1998:161-168 [2.2 & 2.4]). Affective reactions in the human brain are such that quick and conservative is best. This makes obvious adaptive sense and may be either vestigial of evolutionarily earlier responses or still playing an active and useful part in what gets us around in the world. And remember, these quick and dirty responses are still what get most animals around in the world; higher cognition and consciousness are the exception rather than the rule in the mental life of animals, and a rather small exception at that! (Le Doux, 1998:300-302).

I think the explanatory force of involuntary or naïve affective reactions is clear for a limited range of fictive interactions. There is the well-known phenomenon of *startled* audiences at horror movies suddenly cowering back from the screen as a smiling man with an axe bursts through the door. Gore – for example the notorious *lawnmower scene* of Peter Jackson’s *Brain Dead*) may elicit those uncomfortably icky feelings in a *naïve* manner. Being *repelled* by the disturbing images in the first thirty minutes of *Saving Private Ryan* (or much of the content of *American Psycho*) might also be understood in this way.

This naïve causality might be at work in bringing about startle, surprise, and disgust affective states, but also in more *positive* fictive responses in the form of our affective appraisal of other people. The literature on affect provides evidence that the perception of species-specific vocal and facial displays can cause affective responses in subjects (Le Doux, 1998:254). The idea that we appraise the people we meet unconsciously has become quite familiar, but Le Doux thinks such reactions may often be affective and, as such, of a form in which the cognitive content may not be of the type available to consciousness and in which the reaction runs automatically. Damasio provides evidence of this in the case of the extreme short-term amnesiac Dave, who, though unable to recognise people, betrayed preferences and avoidance behaviour to favourable or unfavourable *acquaintances* (1999:43-47). The emotional attachment was registered by simpler unconscious perceptual cues. Evidence of this is also provided in the form of *associative prosopagnosia*, where sufferers are unable to recognise particular faces, despite being conscious of the features of the faces, which suggests face recognition is unconscious and achieved by specialised mechanisms (Damasio, 1999:162-167). So here we have a level where we may interact emotionally (affectively) with fictional folk without involving cognitively complex beliefs and desires, and certainly without consciousness of these – a way in which a smile almost is infectious. Yawning, though needless to say of little aesthetic import, is a rather more obvious example of how facial displays can cause involuntary reactions in perceivers. Laughter is another example, and obviously of more importance in the case of the appreciation of fictions. Canned laughter plays an important role in engendering responses to sitcoms.

Following these ideas, I think it is proper to conclude that, within fictions, there are content or expressive features that are not dependent in a significant way on the *representational features particular to those fictions*. As such, in the case of fictions, they amount to *mistakes* on the part of our psychological architecture, though serendipitous mistakes in that they add to our fictive involvement. Feagin agrees when she argues “certain categories of phenomena, such as facial movements, do constitute significant subsets of the available perceptual evidence and that human beings do respond to them even without contextual factors necessary for perceiving them as significant” (1996:181). This is information our psychological systems might glean from fictions without responding to, or even being aware of, the conventions that define game-bound fictive practice as it is characterised by people such as Walton and Lamarque. This, I think, is also a way to understand how we respond to fictions despite or in ignorance of the conventional status of those fictions. And this is why they should be seen as mistakes; the systems responsible *think* (perhaps metaphorically) the fictions to be real. Darwin’s affective systems *thought* or *believed* the striking snake environmentally salient, despite the context in which the snake was placed; the *context* seemed obvious only to Darwin’s higher cognitive systems. I might conclude, then, that there is a possibility certain types of perceptual experiences – for example, facial and vocal displays such as frowns, yawns, laughs, and screams; sudden movements or ambiguous shapes; disgusting materials; and perhaps even common predators such as snakes – will elicit emotion despite the particular context – in this case fictive – in which they exist.

A critic might question the similarity between the case of Darwin’s affectively reacting to a snake behind a plate glass window and the case of someone’s naïvely reacting affectively to a fiction such as a movie. It might be suggested that the two cases differ because, in the latter, it is not just a plate glass window that stands between the source of elicitation and the subject but a representational convention necessary to our cognitive understanding of fiction. Not only does fiction demand our acknowledgement of the fictionality of fictive content, its representational conventions often alter the nature of our perceptual experience of fictionally characterised contents. In the terminology offered earlier, the *mode of access* we have to fictive content might rely on a matter of representational convention [3.4]. Snakes behind plate glass windows look pretty much like normal snakes. Facial features on movie screens, though looking like normal facial features in a manner of speaking, also look very *different* from facial features. The images on movie screens are flat and usually large. I do not know many people with flat, twenty-foot high facial features. How then can I respond to them without finding these contextual features necessary in my understanding? The fact that the face is big and flat is *very* obvious.

There are several things I can say to counter this. First, there are cases of fictions where the substance exactly resembles real people doing real things. Live drama gets much of its force from this fact. We forget (at least I do) that the history of drama is predominantly a history of live drama, up until the technological advances of the last century. If we affectively react to the perceptual features of real people involuntarily, then providing actors get things right, drama confronts us with real such people to whom the mode of access provides perceptual substance to which we will react in *ignorance* of the fictive convention. More precisely, it is that part of us that does the reacting – involuntary, mandatory, and perhaps *modular*, affective mechanisms – that is *ignorant*.

It should also be realised that this problem with the representational conventions of films may only count against one perceptual medium: vision. Though human faces on movie screens might be flat and eight foot tall, speech and other vocal responses such as screams and laughs, and environmental sounds such as animal noises, sound *exactly* like what they are. Some of the representation involved in fictions may simply replicate aural features of our ordinary environment.[58]

However, the really important thing is that the media employed in representing fictions are not specific to fictions. This is something Lamarque shows when he argues there is nothing intrinsically semantic, syntactic, or structural to fictions that makes them fictions (Lamarque and Olsen, 1994). I have argued the same point here, though with the focus placed on propositional content [3.1&3.5]. Rather, fiction is comprised of a type of prescribed social transaction; the fictive status is derivable from the pragmatics surrounding the utterance. The representational techniques used to portray fictions (such as movies or novelistic forms) might equally be used to portray true stories, and often are. What I would need to provide next for this story to make sense is a manner in which, once representational conventions such as film and written stories are grasped, naïve affective reactions might be inspired nevertheless. The images we see on the television news are a source of affective arousal, especially when they are as sensational as they commonly are. To show that such images might be the cause of naïvely elicited affective reactions would be enough to show that in fictions similar reactions might occur. I might satisfactorily reconcile the understanding we need to have of representational convention with Feagin's contention that "human beings do respond to [fictive stimuli] even *without contextual factors* necessary for perceiving them as significant" (1996:181). The contextual features here are the rules of fictive engagement, and not the particular representational conventions through which fictive works might be portrayed, that is, they are fictive conventions rather than representational conventions. The representational features are not specific to fictive representation. Understanding the latter certainly seems necessary, but because they are not specific to fictive practice, this is no reason to doubt

---

[58] Steven Pinker provides a theory of art in which art is based on the exploitation of ancestral psychological mechanisms by modern technologies such as pictorial representation, music, and television drama (1997). Pinker's representative sample of art is likely to be too simplistic to convince any philosopher of the arts that he has a satisfactory explanation of the psychology of art. Pinker's *cheesecake* theory also plays down the role of *representational conventions* even in the model of art he provides. His treatment of fiction is prone to both of these complaints. I consider Pinker's theory later [7.2].

Feagin's contention. She merely claims that the affective systems need not be aware of the fictionality of the stimulus.

I have now opened the way for a substantive discussion of our naïve affective reactions to fictions. I said earlier that the objects of responses might be identified with the type of environmental information to which the affective systems are adaptively wired (such as the shadow in our peripheral vision that causes us to start, or human facial and vocal changes). This suggests a source for the disgust and fear elicited by *Brain Dead* and *Saving Private Ryan* respectively. It is also a possible source of our affective engagement with people, especially when those people, as they are in stage plays, are not even separated from us by anything like a plate glass window. Actors on stage appear exactly like people emoting concerning the dramatic situation around them (assuming they are good actors) because they literally *are* people emoting (externally) to the dramatic situation around them. There is no necessary link between an external emotional display and what is going on inside someone's head. People fake emotional displays all the time. Therefore, the fact the dramatic substance is fictional is no bar to actors externally emoting concerning the situation.

This is complicated by the fact that the facial displays of the emotions that Ekman discusses can be hard to fake because of the modular nature of affect programs at their basis. The automaticity that makes facial expressions good evidence of our mental states may also make it hard for actors to adopt *convincing* facial expressions. One manner in which actors may avoid this is by themselves eliciting the corresponding emotions via imagining incidents or via remembering real incidents with a congruous emotional character, that is, via *method acting*. Method acting may amount to an attempt to avoid the constraints on voluntary emotional expression imposed by the modularity of affect program responses.

Such perceptual features are given their emotional importance by having been established through the evolutionary development of the cognitive and affective phenotype, or the development of the individual, as worthy of eliciting an affective response. It is adaptive to find things like rotting food, anguished facial features, the breaking of social taboos[59], and sudden movements in the peripheral vision affectively engaging, and especially so for animals with an evolutionary history like ours, but which lack the ability to engage cognitively with their environment to the extent that we do.

---

[59] This seems to be an affective reaction elicited by a socially emergent elicitor, and as such is an effective example of how affective reactions can be *naïvely* elicited by what are socially constructed states. Joseph Le Doux notes an interesting study on how taboo words such as "bitch, fuck, Kotex, cancer" are processed in periods of time not consistent with their conscious appraisal of the meaning of such terms, showing that the emotional significance of such terms may be established unconsciously (1998:56).

It is also important to remember that the type of *information* meant here is not always the narrow form of cognitive manipulation of mental representations that comprise beliefs and desires. Le Doux makes the following suggestive remark:

> In fact, the amygdala of all creatures may be prepared to respond to species relevant cues. For example, faces are important emotional signals in the lives of primates, and neurones in the monkey amygdala respond briskly to the sight of monkey faces.
>
> (Le Doux 1998:254)

It is obvious why humans, their actions, and their perceptible features are objects of affective responses; a large part of what makes life good or bad is tied up with our relations to other people. At the very first we have an interest in orienting ourselves to others and this process is something we note in infants early on as they direct their attention toward the eyes of those around them. It seems so natural to us that other people are there that we tend to lose sight of the fact that the mind must have established systems very early on in life for realising this primitive object-directed attitude. There is evidence that it is via the various affective systems that we orient ourselves behaviourally to the world; affective programs provide us not only with salient physiological feedback (that is, feelings) they also *direct our attention toward the emotional stimulus* (Le Doux 1998:149, 289). Damasio thinks the emotions “enhance” the important objects in our environment, enabling cognition to function (1994:173; 1999:193). These functions certainly seem to be the germ of a *primitive object-directed attitude*. Should this be so, and noting the quick and dirty nature of the affect programs, it is easy to conceive how we might naïvely establish such attitudes toward fictions, especially when those fictions exhibit all the perceptual features of ordinary humans, as many fictions – most obviously live drama, but also as I have argued, films – do. Like Feagin, I think the content or information that is the subject of these responses can correctly be seen in many cases as *narrow* in that it is not dependent on the conventional form specific to the fiction; the amygdala just does not care. I must, therefore, include these narrow, perhaps non-doxastic, cognitive reactions within what it is to participate emotionally with a fiction.

In terms of the theory of fictive practice developed earlier, the *props* that help us generate the content of our pretenceful activities in fictive practice may also have perceptual features that tend to cue naïve affective responses. When the props have features that bear a close resemblance to the sorts of cues relevant to our simpler affective responses – most likely because they *duplicate* those cues – we will respond to them. This may not depend on our consciously registering that feature of the prop. The props that constitute works of fiction are replete with content that add to the richness, novelty, and objectivity of our games of pretence. Furthermore, some of this content may be derived by our perceptual and affective system without the mediation of pretence-bound cognitive participation. The props that guide our games of make-believe may therefore contain more than narrowly conceived propositional content. Fictive props may contain information other than propositional content.

Because the reactions do not acknowledge the representational conventions of the fiction, they are strongly isomorphic with cognitive states that engage with the non-fictive world. This is for the same reasons as for the isomorphisms that exist in pretence-bound participation. These cognitive isomorphisms, as displayed in our elaborative and non-elaborative inferential activities, exist in fictive and pretenceful practice because of a common structure at the basis of fictive and genuine processes [3.2&3.4]. In the case of the inferential isomorphism, it was because the representational tokens that subserved pretence inference were in the same code as those that subserved genuine inference, and the inferences were perhaps carried out on the same inferential mechanisms. In the case of the isomorphisms of naïve reactions to affectively primed stimuli, the two states again share in the same psychological mechanism. The structure that is activated in the brain in the perception of features of vocal or facial displays, or in the other situations I mentioned, is just the same in a *real-world* experience and an experience of *fictional worlds*. This is because the mechanisms involved are not able to tell the difference, in which case these perceptual experiences are strictly psychological *mistakes*, or just do not care. Thus the states are not only isomorphisms, but are *qualitatively identical* with equivalent *real-world* experiences. This, of course, is because they *are* real-world experiences.

This source of fictive emotionality, and the general idea that affect is mandatory and automatic, might help to explain a puzzling feature of our fictive emotionality. Our beliefs are *quarantined* from the cognitive states that make up our fictive involvement. Otherwise there would be the threat of irrationality as our actual beliefs were mixed up with what it is we make-believe and which it is *only fictional we believe*. Affective states sometimes seem less quarantined in fictive participation than cognitive content is, however. Emotional reactions to fictions seem to often spill over into our genuine emotional lives, even though we are aware of the fictionality of the states that elicited them [5.2]. I have met people who are unable to endure horror movies or movie violence; though they are quite aware it is only fictional there are horrible events or violence occurring. We really can also be made *very scared* of fictions; anyone who was allowed to watch horror movies as a child should be aware of this. Given that affective reactions are opaque and mandatory, these results might occur because we are unable to control the affective reactions we have. Though we are aware of the fictionality of the provoking stimuli, the naïve fear and disgust inspired by some fictions feel as bad as the real thing and are resistant to our cognitive appraisals of the eliciting situations. They are resistant for much the same reason phobias are resistant to our cognitive understandings.

***

We should survey some actual examples to see how this theory of naïve affective reaction accounts for various aspects of fictive practice. In a well-known device in horror, thriller, and splatter movies, directors use the sudden introduction of perceptual stimuli to induce startle in the audience. I noted this earlier, when I argued that an emotion thought to have been a highly cognitive appraisal of a situation might have been an affective reaction caused by a simple cue [4.5]. Jenefer Robinson argues that startle focuses our attention on the content within artworks (1995a). It is a domain specific mental module that orientates appreciators to the content or detail within fictions. Robinson also thinks the emotions involved in the appreciation of fictions, loosely characterised as to include startle, allow us to educate our sentiments (1995b). I am generally friendly to Robinson's views, especially her insistence that the narrow cognitivist conception of the emotion may be *too narrow* to cover all the emotional features of fictions we should have an interest in, especially the states with a greater physiological component.

One notable example of startle occurs in Stanley Kubrick's film adaptation of Stephen King's novel *The Shining*. A famous scene involves the main character, played by Jack Nicholson, apparently having gone insane, chasing his wife around the Overlook Hotel with an axe. The wife locks herself in a room and the audience, already on edge from the building tension, is startled as the axe bursts through the door. The sudden appearance of the axe caused me to jump when I saw it and I would not be surprised if this were a case of an involuntary startle reaction. If we assume this is so, we can see that the startle reaction would not have been caused just by the appearance of the axe through the door, but by the manner in which the axe appeared. The scene would need to be *perceptually salient enough to warrant a startle reaction*. When we look at the scene, this is what we find: the axe bursts though the door during a short lull in the action and does so with an extreme force and a significantly loud *thwack*. The startle reaction has been elicited by a sudden and extreme change in perceptual stimuli and it has caused the audience to react with the physiological and behavioural adjuncts of the startle reaction. The audience gives a little jump, some might shriek a little, their hearts race, and adrenaline is released. All of this makes the scene thrilling and terrifying. The appearance of Jack Nicholson's crazed visage through the hole made by the axe only serves to heighten the experience, given his prodigious ability to adopt crazed expressions.

Another excellent example of the startle effect occurs in the Tom Hanks movie *Castaway*. Tom's character is on board a doomed cargo flight somewhere over the South Pacific when it runs into turbulence that will ultimately lead it to crash into the sea. Before this the camera looks into the mirror in the aircraft toilet as the character checks a bandage on his finger. The scene is a *reflective* lull in the action, but is suddenly broken by a bright light and sharp sound as the mirror breaks, Tom is knocked to the ground, and the plane begins its thrilling *death plunge*. Another, similar air crash scene occurs in the wonderful movie *Fight Club*. Possibly these scenes also play on the established fear some of us have concerning flight, but the employment of this filmic tech seems to follow a similar profile: there is a lull in the action, everything grows quiet, it seems apparent the character is out of trouble, and then, as they say, *all hell suddenly breaks loose*. A form of Internet site replicates this dynamic profile. On being sent a seemingly innocuous link to some purportedly *hilarious* or *raunchy* content, you click on the link, and the site invites you to turn up the volume on your computer to better enjoy the experience. Typically there is some serene piece of content, or a task – I have seen examples involving pictures of cute kittens, a calming advertisement for ambient new-age music, and a *Where's Wally* style puzzle – that attempts to gain your close attention. It is when the site has your close and quiet contemplation that the *screaming face* appears, usually accompanied by an intensely loud screeching noise.

Another example of a naïve reaction to a fiction is elicited by the laugh track or canned laughter that is found accompanying many television sitcoms. Though many of us might find them tacky and in bad taste, laugh tracks do work, and they work because they exploit the fact that laughter itself is a perceptual cue that is prone to elicit more laughter. Furthermore, it is likely that a "neurological laugh detector" triggers it in much the same way affective emotions are elicited by "appraisal mechanisms" (Provine, 2000:149). I think this reaction can be characterised as *naïve* because the neurological laugh detector likely does not care that the laugh occurs as part of a work of fiction; it is merely the sound that counts. Laugh tracks *exploit* our cognitive architectures to engage our participation; laughing, after all, is fun. If you find yourself laughing at one of the many sitcoms you *know is just not funny*, blame your *neurological laugh detector!*

I can envisage another interesting case in which these naïve effects might be exploited. It should not matter to us who in particular acts a given role; it should be the fictionally characterised action of the fiction we are interested in. And yet, we all know that *movie stars* are a prevalent feature of modern drama. It would also be unconvincing to deny that our experience of an actor in one role has an effect on our perception of their other roles. Here is one among many possible reasons why. The phenomena of movie stars and character actors might exploit the fact that we become unconsciously affectively primed to actors during their roles - recall Dave, the short-term amnesiac who was able to exhibit preferences toward people even though he was unable to explicitly recall who they were. Subsequent experiences of the actors in other roles may draw on this original and ongoing affective priming. Perhaps we saw Tom Hanks in *Forrest Gump* and were primed to respond favourably to his endearing portrayal of Gump. Hanks' facial, vocal, and bodily expressive mannerisms might have been favourably established in our emotional memory by this experience. Subsequent experiences of Hanks, in *Apollo 13*, *Saving Private Ryan*, *The Green Mile*, might have traded upon this established emotional priming and might even further enhance it. It could even tempt us into seeing *Castaway*. This might also explain why Hanks plays a similar character in every movie he has ever made, why he may never make a good *Iago*, and why his role in *Road to Perdition* amounts to a departure from his oeuvre. A similar but more dramatic departure has occurred with Robin Williams, who, once typecast as the lovable and eccentric goofball has now taken to playing psychopaths in movies such as *Insomnia* and *One Hour Photo*.

***

There is an interesting issue here concerning the possible cognitive deferral to affective reactions, and the notion of *the suspension of disbelief*. Le Doux identifies one indirect source of cognitive involvement in affective responses that might be pursued concerning the issue of fictive emotional involvement. It arises through a subtlety in the way affect is wired up. Affective fear evolved primarily as a *quick and dirty* response to the environment and it retains that form in the human brain, the thalamus being still primarily responsible for *evoking* fearful responses (Le Doux 1998:161). With the rise of higher cognition in humans and our close relatives, the neo-cortex began to be used as a more efficient centre for guiding behaviour; quick and dirty responses tend to be over-sensitive, waste a lot of energy and time, and in some cases are less than effective in novel environments. The neo-cortex in the higher mammals has evolved to play a larger role in fear elicitation. This is shown by the increased anatomical connections that exist between the neo-cortex and the amygdala in such animals (Le Doux:161-168). But the evolution of the higher brain centres did not totally negate the effects of such *lower* systems and they are now wired parallel to higher cognition, with the result that affect fear can kick in before or despite our recognition of the environment. In fact, the quick and dirty responses still tend to do the affective work, with the lines of influence running from perception through them *and then* into higher cognition, where it is more closely studied. Le Doux notes, “The information from the thalamus is unfiltered and biased toward evoking responses. The cortex’s job is to prevent the inappropriate response rather than to produce the appropriate one” (Le Doux 1998:165). Higher cognition then *decides* to turn off the response (which it is sometimes less able to achieve than we would like as in the case of phobias [2.2]) or it takes the disturbance seriously and directs attention toward the source of the affect. Much of this happens in the computational unconscious and is captured by what Ekman describes as *coping methods* [2.4].

An ability of higher cognition to take seriously quick and dirty affect responses seems to give some sort of backing to that part of fictive practice that is often described as the *suspension of disbelief*.[60] It would be even more helpful if social cues could be found to contribute to those situations in which the neo-cortex defers judgement to the *lower* systems. This process of cognitive deferment might also go some way to explaining the *sympathetic* attitude necessary to our imaginative engagement with fiction [5.2]. This is admittedly speculative stuff, but it seems to suggest a method by which our enduring affective engagement with fictions might eventuate, despite our realisation of the fictionality of the material. What it shows is that the naïvely affective engagement we might have with various perceptual features can be subsequently located or used by higher cognition in contextualised practices, without basing the affective engagement *within* that contextualised practice. This fits well with Feagin's contention that we reflectively locate the affective reactions elicited by our theoretical practice of generating fictive understandings, and ponder their warrant. This point leads naturally to a discussion of how affect might act as a reflexive prop.

---

[60] Something the existence of which is denied by Walton (1990:241), Radford (1975), and indeed me [1.2].

5.4 Affect as a Reflexive Prop

Walton contends it is not just things such as stumps and the descriptions found within fictions that can act as props in games of make-believe (1990:210-215). In his story of Charles and the fictional slime, Charles' affective fear (his quasi-fear) of the slime is itself a prop in his game with the slime representations. Walton claims that the physiological disturbances Charles experiences make it true in his game that he is afraid of the slime; they provide the basis of his fictional attributions of emotions. I have argued that Walton has not explained much of what is at issue here – that is, has assumed and not given a substantial explanation of how fiction might cause affect, and his theory is vague in places we might want it to be more definite – but, these points taken, I think this idea is exactly right. Feagin expresses the idea in her similar contention that aesthetic appreciation involves theoretical, affective, and reflective components [3.4].

The case in which the fictional Charles describes himself as terrified amounts to Charles' reflexively incorporating his affective states as a prop into his game of make-believe. Charles notes the affective disturbances in himself caused by his imaginative and affective involvement in the fiction, links these disturbances to the fiction, and finds himself prone to describe himself as terrified, given the evidence of what he is fictionally viewing. Charles is going through a fairly typical case of a self-ascription of a mental state, though this disregards the dual facts that he is apparently afraid of a *blob of slime* – which is quite untypical of fear elicitors – and a *fictional blob* at that.

Remembering that affect not only gives rises to feelings and physiological changes but that it also directs the attention of the subject toward the object (Le Doux 1998:149; Damasio, 1999), we can see how affect can be a source both of the *fictional facts* we might want to express about our interactions with fictions, and of our participation with fictions. Affect serves in engaging us with our environment, as Damasio notes, *enhancing* the objects that comprise our environment. Affect is one of the mental processes that generate fictional truths within the game-world we play with works of fiction. The process goes something like this: I read in *Ulysses* the various descriptions of Bloom's attendance of Paddy Dignam's funeral and theoretically generate the truths to which I subsequently adopt pretence-bound participation. For the reasons set out in 5.2 and 5.3, various affective systems kick in and I really feel sad. I not only feel sad but the affect makes the senses of the sets of descriptions that I read *experientially salient*. This is part of what affect does and it explains a practical source of our becoming carried away by our participation in fictions. Next, I reflexively *attribute* these affective reactions to a response I am having toward Leopold Bloom and Paddy Dignam. This attribution is at the same time the sort of thing we commonly do in real life and is a pretence-bound participation that positions my affective reactions within a game of make-believe. I will argue in the next chapter that it plays a vital role in our cognitive or social emotional dealings with works of fiction, and our apparent dealings with their fictional worlds.

An actual example comes from my experience of the first thirty or so minutes of *Saving Private Ryan*. Because of a mix of my imaginative and my naïve involvement in the content and representational expressive features of the depiction of the D-Day landing in the beginning of the movie, I had various affective reactions. Given the character of the experience, I would be tempted to say I was sad, fearful, and disgusted with the landing battle. Empirical evidence would tell whether affective sadness, fear, and disgust were actual affective states elicited in me. Nevertheless, it seems clear that the presence of affective states did contribute to my discomfort and to my subsequent willingness to describe myself as sad, fearful, and disgusted of the situation. This again is a case of self-ascription of a mental state, though, because the cognitive appraisal of the affective states drew on fictive material in framing the emotional experience, it amounts to the reflexive incorporation of my mental states into the game-world of the fiction.

Thus the affective activity we have while experiencing fictions may often be a form of isomorphic participation. Affective reactions to thoughts of emotionally loaded situations will often be fictionally characterised. While reading one of the more graphic sections of *American Psycho*, it may happen that I will become nauseated by the descriptions of power tools and kitchen implements and what is done with them, because of an affective reaction to the imagined content. It is likely I become *fictionally nauseated* by the work-world of the fiction. Other affective reactions, such as a sudden surprise or startle elicited by the zombie bursting through the door, will make it fictional that the zombie bursting through the door fictionally surprises appreciators.

Feagin is not the first to think some sort of affective engagement is vital to aesthetic appreciation, as the idea has a very long history indeed. Without some sort of *affective* or *emotive* component – that is, with Feagin’s *theoretical* component alone – aesthetic appreciation would be rather dull. It is hard to see why many aesthetic works would merit our attention if they did not engage our emotions. *Affect draws us into fiction*. This is a popular component of many theories of the arts; the appeal of art is seen to lie, in part, in our finding it pleasing or displeasing. I think my approach is consistent with the interesting notion of mental props. I have, however, aimed to show there is much more going on than Walton realises, particularly in the form of facts available to the philosopher from the sciences of the mind, especially cognitive science, neurophysiology, and neuropsychology. Walton cannot be blamed for avoiding the wider issues of course; he is a philosopher. This may just again emphasise the importance of widening the scope of this study.

The manner in which we characterise our mental states might add an interesting aspect to fictive participation; it opens the way for an explanation of some of the more sophisticated, odd, or incongruous emotional reactions we have toward fictions. A mix in the emotional stimuli might lead us to attribute to ourselves sophisticated mixes of emotions such as *Schadenfreude* or horror. Aaron Meskin and Jonathon Weinberg (2000) argue that some of our complex emotional states directed toward fictions involve a mix of affective states. Also, in such cases we may not be sure what our feelings about the fiction are because the fictions are deeply ambiguous. In our dealings with the world we are often put in ambiguous situations in which we are not sure exactly what it is we are feeling. The point of the Schacter and Singer experiments was that sometimes this causes us to confabulate about our mental states [2.2]. In a fictive experience, a first reaction, perhaps affective fear or disgust, may subsequently be cognitively framed to take into account other aspects of the fiction, especially those obvious in stylistic or representational features, as well as our mental attitude toward those features. Some of the emotions we have in fictive contexts might combine in unexpected ways. Many fictions portray complex situations and types of content, with an implied attitude toward the situations that seems incongruous with their emotional nature. I am thinking particularly of extremely black comedies such as *American Psycho*, or deeply ambiguous and troubling works like Terry Gilliam's *Brazil* or David Lynch's *Blue Velvet*. This puts the appreciator of such works in a difficult position. When I originally read *American Psycho*, I was deeply unsure about my reactions to the fiction and how they should be characterised. Did laughing at Patrick Bateman – even though his actions shocked me – somehow implicate me in his crimes? Perhaps, though, it is a virtue of such a work that it touches such a raw and ambiguous nerve. Such complexities will open the way for affective reactions to generate more sophisticated fictive participation.

5.5 Affect, Fiction, and Behaviour

I have already discussed the place of action and behaviour in fiction appreciation, noting that, though it is often available, in characterising it we must make reference to its fictional nature. Acting in conjunction with fictions, in as much as it is apt to be reflexively incorporated into a game of make-believe, must be seen as *fictionally acting*. Shouting at Punch and Judy that the Alligator is behind them amounts to fictionally shouting so. But there are also some points to be made concerning the role of emotion in these actions.

There are two types of action that might arise in fictive participation. First are those actions that are licensed by the particular fictive game that is being played. Many games of make-believe, especially those in children and in role-playing and videogames, allow the participators overt physical actions. Role-players may thrust their tinfoil swords and halberds at each other; videogamers may aim their light-guns and wrestle with steering wheels. Of course, less overt actions are available to appreciators of cinematic and literary fictions in the form of the interpretative engagement with those fictions.

But there are also actions – less consciously deployed perhaps – that arise from our affective engagement with fictions. Among the physiological adjuncts of affective states are some that are distinctly behavioural [2.4]. The various affective reactions cause us to flinch, direct our attention, exhibit vocal, facial, and bodily expressive reactions, and further, to direct our higher cognitive abilities toward the stimuli and so engage our display rules and coping methods. Many of these behavioural aspects will become involved in our fictive activities, though it is likely that the more cognitive adjuncts of this behaviour, the coping methods, will be curtailed for the obvious reason that in appreciating fictions we are quite aware there is no situation with which we must cope (or if there is, it is quite different to the one only fictionally characterised). It is obvious from casual observation that some of these affective adjuncts are involved in our fictive experiences, however. Fictions are often the cause of obvious behavioural appendages of affect, such as laughter and crying, in fiction appreciators. Audiences at horror movies can be observed to flinch or shriek as the result of the startle elicited by a perceptual feature of the movie. Moviegoers may also squirm in their seats, smile or grimace, close their eyes, look away during a particularly nasty scene, or even grab at the arm of the person sitting next to them. These actions, when they occur, often follow simply from the engagement of affective systems that have been elicited by our imaginative or *naïve* involvement with the content or representational features of fictions.

When more complex actions follow, it is likely they are warranted by the rules of the type of fiction that is being experienced, though the affective states may have their influence there. That is, affective states and their behavioural adjuncts may show their effects in rule-based actions. In pantomime, though the activity of the children in shouting “the Alligator is behind you!” may in a large part be caused by the affective excitement the children feel, it is an available action because the children understand the rules that guide the engagement in fiction. In the other examples of fictive practice I imagined, for example where the audience is invited to wrestle Hamlet in a production of *Shakespearean Cage-fighting* [3.6], though the affect may be involved, it is likely the behaviour – the punches thrown, any *body-slams* or *pile-drivers* that may take place – arises because the participator understands this form of involvement is warranted in the game-worlds that appreciators construct with this particular type of fiction. The emotions likely to be involved in the game might add conviction to the role-players’ performances of course! Because they exist within a game of make-believe, however, any furious kicks aimed at Hamlet’s groin will only be fictional furious kicks aimed at Hamlet’s groin. This is the case even though the participator may be aiming a kick at the groin of the actor playing Hamlet, *with some considerable fury*. Incidentally, any pain Hamlet felt, should the kick connect, would be fictional pain, even if the actor playing Hamlet was writhing around on the ground and screaming. The actor’s pain would be a prop that made it suitable to say that Ophelia had levelled Hamlet with a well-aimed kick to the proverbial *family jewels*.

I would like to introduce one more example of fictive behaviour, one that I signalled in my introduction [1.2]. In my admittedly extensive experience of playing videogames, I have seen an *extremely strange behaviour*. It displays itself most obviously in car racing and driving games. The brilliant Playstation 2 game, *Gran Turismo 3*, which involves players driving various sports cars around challenging race tracks and attempting to amass enough prize money to buy better cars, players can often be observed to *lean into the corners, attempting to compensate for the movement of the car*. The same can be observed in motorcycle racing games and in adventure games when the character is jumping across a dangerous precipice; the player physically hops at the same time.

Colin Radford thought the case of tennis players doing a little hop in urging the ball over the net was irrational, but at least tennis players believe in the existence of the ball! The computer games players I have encountered, myself included, are clear that the Nissan Skyline GT-R super-car they happen to be driving is in fact a real-time computer animation, on which their physical movements (apart from those directed toward the handheld controller) can have no influence. But they can often be observed in this strange behaviour. An incident was recently related to me, involving a Playstation driving game that was being projected onto a large screen, in which the participant had trouble keeping his balance while standing and watching! What should I say of this puzzling behaviour?

Though I would gladly defer to the judgment of a cognitive scientist, I expect that, while it is explicitly realised the depictions of cars and racetracks are only fictional, the visual depiction of them exploits underlying features of our perceptual systems that give rise to the enduring perceptual orientation we have to our environment. It is likely the perceptual systems underlying our perceptions of visual scenes make *assumptions*, particularly about the orientation of visual scenes, allowing us to move around the world, rather than have the world move around us (Pinker, 1997:263). Further, this perceptual orientation is integrated with our physical orientation in our world, in short telling us which way is *up* and which is *down*, and to systems such as those that allow us to balance. Changing the orientation of the fictive viewer to the fictional scene, even though the viewer is aware of the fictionality of the stimulus, may naturally give rise to balancing behaviour because the perception of fictional visual scenes is subserved by the same mechanisms as those that subserve *genuine* representation of the world. Furthermore, it may be that balancing behaviour is automatic or mandatory. This explanation is parallel to my earlier explanations of our affective engagement in fiction, which involved fictive representations being *coded* identically to beliefs, and as such available for processing by the affective system that other mental representations are available to.

6. Fiction and the Cognitive and Social Emotions

6.1 Cognitive Emotions and Fiction

I argued that typical affective reactions directed toward fictions would be surprise and disgust elicited by a splatter movie, repulsion at graphic war representations, affective responses to images and to understandings of situations, or naïve responses to communicative facial features. Affective emotions are foremost short-term stereotypical states, and account for such reactions when they occur in fictive practice. Higher cognitive or social emotional responses to fictions would account for our *caring for* or *fearing for* characters, having *sympathy* for them, or *empathy* with them, *liking* or *hating* characters, or feeling *embarrassed for* their actions, or, if we think appreciators are being genuine when they report such attitudes, *falling in love with characters*. That is, they would be comprised of the more cognitive and social dealings we have with the emotional substance of fictions and of an emotional substance that involves an increased cognitive and cultural complexity.

The thesis that will be pressed in this chapter is that higher cognitive and social emotions, being states that frame our involvement in various social domains such as conflict, mating, and cooperation, allow us to participate in and understand the *social problem spaces* portrayed in many fictions. It is no surprise that the plots of most fictions conform to several basic types, being focused on loves, enmities, threats to familial and community groups, revenge, and discovery. These are among the domains we must deal with as a part of life. Our cognitive and social emotions guide our dealings within these domains.

Given that these emotional reactions are typically characterised as a relationship to a *person* or *situation* they seem to demand for their satisfaction the sense of game-bound cognitive pretence outlined in an earlier chapter [3.2-3.3]. They suppose a cognitively appraised relationship with a fictional individual *qua person*. Fictional cognitive or social emotions are engendered by the pretence attitudes of fictive practice because they are essentially fictional relationships. Walton explains how the fiction provides us with a prop with which we can generate our cognitive and social emotional relationships, though I contend again that the emotional work is done not solely with regard to conventional interpretative processes on a fictive prop, but in a substantive emotional way. I think also the emotional work will be done by *theoretical* make-believe states acting as *socially particular* elicitors of affective states or, as in the cases of self-report and emotional theorisation or simulation, by themselves forming the substance of the emotional attitudes [6.2].

Again, my observations await confirmation or disconfirmation via empirical research. You can get a little way on the observation methods available to the philosopher from the armchair, but empirical studies always have the final say. I think it likely that the observations and examples I introduce here will come close to getting it right, in that I have tried to take the results of other philosophical studies of fictional emotions and framed them in terms of the results of scientific studies of the emotions. If anything is novel here, it is the mixture of theoretical devices I am attempting to employ.

Some recapping is in order. Cognitive emotions have a greater social and cognitive dimension than purely affective emotion [2.4]. The cognitive states involved in the cognitive and social emotions are more easily characterised as folk-psychological entities, such as beliefs, intentions, and desires. They arise as we respond to our environment in a way that reflects long-term planned actions and that integrates with social structures and behaviours. It may be that some higher cognitive emotions are mental traits, perhaps modules, that contain *information* about a particular domain important to human life. As such they are “domain-specific” framing devices (Tooby & Cosmides, 1992:99).

The typical example, at least for Western eyes, is romantic love. Love has its canon of behaviours and occurs in certain circumstances. It has certain adaptive aims, though determining these is an empirical matter. To find out what love does and why it exists would involve a significant empirical study of the type Steven Pinker suggests (1997:363-520). Love has its own vocabulary. Self-report is important in romantic behaviour. Also, there are social institutions dedicated to the notion of love. Marriage is sometimes a matter of love, though in some cultures marriage often finds a different explanation. These aspects of the emotions seem to fit well with the cognitively based *coping methods* I earlier suggested formed the genesis of the cognitive or social emotions. These emotions seem to be comprised of the ability to track and cope with stereotypical tasks.

But again, some say love is a *feeling*, and a rather special, ephemeral one at that. There is also some aspect of love that resides in the brain (or some part of the body). One way to explain this would be to introduce an affective component to such higher cognitive or social emotions. There are certainly physiological systems in the brain and elsewhere in the body that give rise to lustful feelings. An empirical study would tell if these are the same systems that comprise affect joy or are specific to structures particular to mating drives. But it seems clear that some such thing is going to need to play a constitutive role in romantic love if this cognitive/social emotion is also occasionally to count as a feeling. This also allows that cognitive love need not always involve occurrent feelings; one need not always feel paroxysms of affective joy or desire to be in love. Such fits of affective joy accompany love only occasionally (their frequency perhaps corresponding inversely to the length of the loving relationship). And this is an important point generally: episodes of cognitive or social emotions need not involve occurrent affective reactions.

It is likely then, that affective states will often play constitutive roles in the higher cognitive emotions, both substantially and in their formation. That is, affect may be involved in cognitive emotions both in their *occurrence*, and in their *ontogeny*. Griffiths, though wanting to draw a clear line between affective and higher cognitive emotions, points out that instances of fear and sadness in early life *can* play a role in the construction of guilt displayed at a later age (1997:121-122). A guilt response might also include the types of affective states that go into the construction of the guilt. Romantic love is then a cognitive emotional state that sometimes might be accompanied by affective arousal. Damasio concludes that the social or “secondary” emotions combine affective reactions with the greater specificity of eliciting states made possible by the enhanced cognitive powers of our species (1994:134; 1999).

Cognitive emotions are different to affective responses in their outward and behavioural nature. Cognitive and social emotions are behaviourally complex in a manner affective emotions are not. Cognitive emotions also allow of a greater measure of cultural variation and social construction than bare affect programs, which often seem to owe their nature to ancestral traits. This means those who tout the socially constructed nature of the emotions can be satisfied in the theory Griffiths puts forth and that I follow here. It is often contended that romantic love is a rather modern notion and one that is not shared in all modern communities. If Plato’s *Symposium* is anything to go by, love was something rather different in ancient Greece than it is around these parts (though I guess that depends on where *these parts* happen to be). The Japanese have an emotion called *amae* that involves “a highly rewarding sense of being dependent on another person or organisation” (Griffiths 1997:101). Given the social realm within which cognitive emotions tend to exist, these cultural variations are understandable.

Cognitive or social emotions, because of their importance in human life, will form a significant percentage of the *facts* or *fictional propositions* and the emotional reactions creators of fictions generate via their works. How we encounter and interact with this content will form the basis for the cognitive emotional interactions we have with fictions. Reading a character-driven novel can often be an extended practice of reading off the cognitive emotional implications of the action and then of reacting to these. Such an interpretation would fit well with the theoretical and affective notions of fictive practice I have been arguing for here. I go on to suggest that the notion of our subsequent reflexive attitudes will also go a long way to explaining some puzzling cognitive emotional interactions we have with apparent fictional people.

6.2 Understanding Fictional Folk

It is a remarkable fact about humans that we are able to understand and predict the motives and actions of those around us [2.4]. Given a clear situation, it is likely each of us could work out what individuals, and at that, particular individuals, would both feel and do in that situation. We also have a natural interest in such an activity. These processes are also a component of our appreciative activities with fictions. Coming to understand any decently complex character-driven novel or film involves attempting to understand what the characters are doing, and why. Interpretation also involves our evaluating and judging these characters and their actions. If the actions and motivations of characters are opaque to us, we cannot usually be said to have interpreted a fiction successfully. Such an interpretative process adds a great deal to our enjoyment of fictions. Indeed, authors may exploit this real-world psychological activity to make their fictions interesting and engaging. Dennett (1987:252) thinks authors exploit what he calls “intentional systems theory” to set up cognitive problems for appreciators; we derive pleasure from solving these problems.

There are two sides to the debate concerning how people achieve this understanding: those who advocate *mental simulation*, and those who advocate tacit *theorisation*. The advocates of mental simulation propose that our understanding of others' mental lives and subsequent behaviours arises as we take their situations and plug them into our own mental architecture. We run *off-line simulations* that furnish us with the feelings, motives, and behaviours of a person in the situation we simulate. The idea of putting yourself in *someone else's shoes* is central here. In terms of the philosophy of art, Currie's notion of "secondary imaginings" is an application of simulation theory to fictive interpretation (1995b; 1997a:54-56). Feagin's view of fiction as centrally involving sympathetic and empathic appreciative stances is also a simulation theory (1996). Though not assenting to theory theory, Carroll argues that simulation is not as centrally involved in fictive practice as theorists such as Currie think (Carroll, 1998:342-356).

The proponents of *theory theory* think the more likely form of our understanding is the application of folk-theories that have been built up over the history of our dealings with others. The theories are tacit in that we do not always employ them in a conscious manner, often reasoning through the steps and eventually coming up with the results somewhere in the cognitive unconscious. They can be minimally defined as a *body of knowledge* employed in generating predictions and explanations of others. Nevertheless, the types of terms employed by such theories are familiar. They are comprised by the vernacular emotion vocabulary that has been a focus already in this essay and that should be seen as a discourse no less theoretical than scientific discourse. Among the central terms are "belief," "desire," "love," "hate," and the like.

There is even some suggestion that some of these interpretative activities might be automatic, mandatory, reaction-like, and even modular. This is suggested by an experiment carried out on subjects' reactions to animated line drawings. People are all too willing to see simple geometrical animations as having reasons and motivations, even emotional impulses, and, when those people are neurologically impaired, their ability is impaired (Damasio, 1999:343-344*n*11). Greg Currie and Kim Sterelny suggest that "modular mind reading" might be involved in the appreciation of fiction (Currie and Sterelny, 2000). Should this be the case, perhaps this is an aspect of our understanding of people that is more apt to be characterised as modular or even affective. If this process is involved in fictive appreciation, strictly I should include this interpretative attitude in the section of the last chapter that dealt with naïve appreciative attitudes.

Given the heterogeneity in the mental substance the debate between simulation theory and theory theory discusses, including also the heterogeneity Griffiths signals in his division between the affect and higher cognitive emotions, I think the disagreement might be only an apparent one. The notions of simulation and theorisation seem too vague and too ambitious to be theoretically useful. I share this opinion with Stich and Nichols (1997). I think it is likely we might both simulate and theorise *in some sense of those terms*, both aspects sometimes being mixed, and even that an aspect of our social understanding may be modular, as in the processes cited directly above.

Something like simulation occurs as a result of plugging others' situations into *our own inferential* or *affective* systems; which demands a lot of theorising if we are to have anything to plug into our own mechanisms. The fact that counterfactual thoughts and images are a proper object of inferential operations and affective responses has already played a large part in this thesis, so it is obvious how thoughts characterising others' situations can be inferentially relevant or affectively engaging for *us*. Such an activity engages *our* inferential and affective systems, which provided reliable predictions of the other's mental state, because luckily we have the same factory standard mental architecture they have.

Tacit theorising will occur in much the same way as it occurs for the Gururumba people in the case of being a wild pig, and more extensively as it applies to the broad range of *less exotic* higher cognitive emotions. In the case of the Gururumba, I noted the members of the society had a theory of a particular human behaviour, applicable whenever someone exhibited certain types of external behaviour. The theory involves what in our mind are the dubious theoretical posits of *spirits* and *ghost possession*. For this reason it should be obvious that such a description of the behaviour *does not involve a simulation*: the Gururumba could hardly plug the violent young men's situation into their own latent *wild pig modules*. We should not be too smug; there are those who think our whole theoretical schema, folk-psychology, is similarly flawed (Churchland, 1981). I think that we often employ such theories in our everyday understanding of others. Perhaps mental simulation also presupposes the existence of such a theory; what else could we plug into our own mental mechanisms? The importance of emotional talk and self-report in the cognitive emotions is likely to fit comfortably with the idea of mental theorisation.

I need to go back to several of the distinctions made in Chapter Three if I am to appreciate the place in fictive practice of this ability to understand and predict the actions of others. Whatever the final conclusion of the debate, it seems clear that an explanation of how we understand people is likely to provide much of the form of the explanation of how we understand fictional folk. Most likely, we simulate and/or theorise concerning the situations presented within representational artworks; the fictions are *props* for our doing so. The *theoretical* form of make-believe identified earlier [3.4], is an ideal way to understand the role of mental theorisation in fiction appreciation. It can be seen as a way we theoretically participate with fictions, generating the content that fills out their work-worlds. Understanding fictional folk involves the *theoretical or generative construction of the work-world of a fiction*, and, in the cases where we theorise about our relationship or attitude to that work-world, the *generative construction of a game-world*. Also, the participation here happens to be in a form we ordinarily attach to the real world; the pretenceful participation secured by Walton's notion of make-believe or by Nichols' and Stich's theory of pretence means we are able to attach these understandings *profitably* to fictional worlds. As such, the states are functionally parallel to real instances of mental theorisation. These activities will be isomorphs of the processes in real-world activities and will be easily characterised in a fictional idiom. This means the resulting *understandings* exist only within our game-worlds, and are *fictional understandings*.

Therefore, at the causal basis of these perceived isomorphisms in fictive practice are *common psychological mechanisms or structures*. When characterised as understandings and not merely as the activity of cognitive mechanisms, one has an actual object and it is only fictional the other has an object at all. In these cases, unlike the earlier case of the naïve perception of emotionally salient stimuli [5.3], the isomorphism here seems to demand that a pretence-bound attitude toward the fiction be adopted prior to the utilisation of the isomorphic theoretical or simulative method. Prediction and explanation require a complex fictional object – moreover character or situation – about which to predict. In order to be able to understand a fictional person we need to have formed the complex representational states that subserve this fictionally characterised object, and to have placed them in a possible world box. The interpretation of this object relies on our specifically emotional understanding, however.

So part of the process by which we generate the facts of the fictions we participate with relies on *domain specific cognitive traits* central to our interpretation of ordinary people. These are likely to be among the *knowledge structures* identified in the section on the cognitive and social emotions [2.4]. Our interpretative interaction with fictions draws on more than just an understanding of how to participate with fictions, as we also need to know how to participate with people. This understanding is achieved by the universal traits that ground our everyday ability to mind read. Our complexly cognitive social emotions play a role in allowing us to interpret fictive works. I might then like to add to the fictive generative machinery discussed by Walton [3.3] domain specific mechanisms that are not particular to fictive practice. By what principles do we generate the facts internal to fictional worlds? Not only do we apply the *real world*, *mutual belief*, *and existing folk lore or trans-fictive* principles, we also employ our natural mind reading abilities, be they based on simulation, theory, or a heterogeneous mixture of the two.

These predictive and explanative states are only *relatively isomorphic*, as pretence-bound states always are. The actual focus of pretence-bound states is representational tokens in a possible world box, or pretencefully held thoughts, and not people and situations. The isomorphism though, is in the structure of the generative processes and not in the foci of those processes, which *of course* is not isomorphic; we know fictional people do not exist. Subsequently, these processes are fictionally described as prediction, explanation of, and emotional reactions to, people and situations, even though what is really going on in the mind is the manipulation of complex counterfactual representational states.

At this stage, the *reflexive* attitude becomes important. That I pretencefully theorise from the descriptions in *Ulysses* that Leopold Bloom feels the need for a son to replace the one he lost years before, and one as brilliant and frail as Stephen Dedalus, and feel affective sadness elicited by these imagined cognitive emotional understandings, makes it fictional that I understand and feel sad for Leopold Bloom. This is allowed to happen because we adopt a cognitive reflexive stance to aspects of our own psychology, fictionalising them within our games of make-believe. The process is again one of pretence-bound participation. The cognitive emotional theorisation provides us with further representational tokens about works of fiction, to which we adopt pretence-bound participation.

Our affective reactions to fictions generate fictional truths in our game-worlds and, as such, affect plays a generative role in fictive practice [5.4]. The cognitive emotions give rise to similar features. In applying the idea of Walton's reflexive mental props to the higher cognitive emotions, because we form cognitive emotional understandings focused on fictive props, it will be fictional that we form emotional understandings about fictional people and situations. Depending on what conclusion comes out of the simulation theory/theory theory debate, it will be seen that we fictionally simulate or theorise about the actions of fictional characters. Again, I think the details will rely very much on the substantial discussions, but, via Walton's rules of make-believe engagement, we can see how we interact with fictional characters to which we lack the normal psychological attitudes. And because a great deal of what is commonly seen as great in art resides in the use of realistic characterisation, this ability to make-theorise or simulate seems to be one of the most important things we do when we appreciate novels and other fictions. It also follows that great authors and producers of other fictions are particularly good theorisers or simulators in that they are able to describe convincing emotional situations.

There are matters here concerning the epistemic access to others' minds that one is commonly granted in various fictive works, especially those of the Modernists. Theorisation exists in people to counter a practical form of the problem of other minds that notes we have uncertain access to other's particular mental states. Theorisation and simulation respond to this by using various aspects of our own psychology to bridge the gap in access. Modernist novels, however, often provide thoughts as part of the substance to be internally engaged in.[61] Molly's soliloquy at the end of *Ulysses* is one such case. Here there is no need to theorise the likely substance because Joyce gives us first hand access to Molly's thoughts. The technique has surprising consequences. In this case, it allows Joyce to populate Molly's mind with the kind of vocabulary one might find in the cheap romance novels Molly spends her days reading. The fact that Bloom had that afternoon purchased her a copy of *Sweets of Sin* – "*Her mouth glued on his in a luscious voluptuous kiss while his hands felt for the opulent curves inside her deshabilé"* – gives us reason to laugh. Given Molly's other favourite pastime, adultery, which presently involves a certain Blazes Boylan, this fact also makes understandable Bloom's wish to have Molly instructed in the literary arts by Stephen. In this case a novel narrative style gives us further access to the fictional content of a work-world, though it may be in a form that is only with great difficulty fictionally characterised as the content of a game-world given its demand that appreciators be seen as omniscient observers able to peer directly into the thoughts of others.

---

[61] Carroll makes a similar point in an argument against seeing simulation as primarily constitutive of our interpretative participation with works of fiction (1998:350-51).

Here are some more actual examples of this process. During my undergraduate student days I had ample opportunities to observe flatmates watching soap operas. A typical conversation about a soap opera might proceed like this:

> Peter: Why is Chris doing that? [Motioning toward the television screen where Dr Chris is represented as changing the duty roster in the hospital in which he works.]
> Susan: Because he's in love with Maude.
> Peter: And?
> Susan: Maude's on a late night shift with Clint, and Chris suspects that Clint has *dishonourable intentions* concerning Maude. Actually, Clint *is* planning on getting Maude into the sack.
> Peter: What about Maude?
> Susan: What about her?
> Peter: Does she have feelings for Clint?
> Susan: No, she thinks he's slimy. I think she has secret feelings for Dr Chris.
> Peter: Oh. But he doesn't know about this?
> Susan: No.
> Gary: Doesn't this plotline strike anyone as trite and contrived?
> Peter & Susan: It's only a soap opera!

Susan and Peter are engaged in some pretty rudimentary elaboration of the plot of *Shortland Street* in making inferences about the explicit occurrences in the work-world of the fiction. What makes this case salient is that the inferences make reference to the postulates of folk-psychology, such as intentions. Furthermore, they make reference to emotional states such as the love and other feelings of the characters fictionally represented. I think this interaction is typical of appreciators of soap operas and, if it involves an inferential process particular to our prediction, explanation, and judgement of the mental states and actions of other people, then it will be the case that these states are involved in the appreciation of fictions such as *Shortland Street*.

This emotional manner of fictive interpretation also occurs in fantastical fictions. In *The Lord of the Rings* and countless other fantasy and science fiction works, though the characters may seem exotic, they are usually assumed to have mental states and motivations very much like our own. It is only in this way that we can not only understand Frodo's and Sam's mental journey, but also *empathically* identify with them. When we are asked to interpret the actions of characters with different kinds of mental states to our own – such as *Star Trek's* Vulcans and androids – our reasoning may be of a theoretical type that does not draw on our emotionally specific knowledge structures [3.3]. Given this, it may be no surprise that empathic identification with such characters is often precluded, and that such characters are often dramatic lame ducks. Consider also Kafka's *Metamorphosis*. Part of the special appeal of this story lies in the puzzling situation Gregor Samsa finds himself in. What does it feel like to have woken up as a beetle? This story may derive much of its force because it poses an *unexpected* question from which we struggle to draw out the implications. Its features may also track onto other emotional domains, such as alienation and rejection.

Our emotional theorisation will also often cause further emotions in us. One example is an emotion I often feel strongly while participating with fiction: *embarrassment*. I doubt I am alone in feeling embarrassment for some fictional characters. *Seinfeld's* George Costanza makes yet another social transgression – his mother discovers him *treating his body like and amusement park* – and I feel the shiver of embarrassment for his inability to conduct himself respectfully. The reaction is often displeasing enough to make me cringe and turn my attention from the particular fiction that caused the embarrassment. Though embarrassment is not usually classed among the affective emotions, it is similar. The state is elicited by highly complex social situations – and some not so complex ones, of course – and it gives rise to physiological and qualitative states. I often feel embarrassed for George Costanza and his social foibles, and I expect these states of apparent embarrassment, like comparable fictive states of socially particular shame, anger, or satisfaction, to find their explanation in the theory I am offering here. They amount to combination of pretence, emotional theorisation, and emotional reactions.

6.3 Who Loves Janet Weiss?

At this stage I am in a position to explain the actions of those who claim to have higher cognitive emotions toward fictional characters. What do we say of people who claim to fall in love with or to pity, hate, or care for fictional characters? Informal observation has shown me that people are sometimes willing to make such ascriptions. But can I fall in love with the character Janet Weiss from *The Rocky Horror Picture Show*? I have explained the presence of affective emotions in fictive practice, by seeing both imagination and naïve eliciting methods as important, but how can I reconcile the seeming existence of cognitive emotional *relationships* to folk we know do not exist?

There are several issues here. First is the proviso that such relationships, should they be possible, will be only *fictional relationships*, given their location within games of make-believe. Second is the question of why appreciators are tempted to describe themselves as having relationships with fictional characters. Third is the question, following Alex Neill's claim that our inability to play the requisite roles in fiction counts against our having certain types of emotions, such as jealousy (Neill, 1995:254-255), of whether we can have fictional emotional *relationships* with fictional characters *at all*.

The first issue is drawn from aspects of the theory of fiction and emotion I have presented here. Via our fictive theorising, our affective reactions, and our higher cognitive emotional understandings focused on works of fiction, we are supplied with a range of tokens representing the premises of fictions. Via make-believe, it becomes appropriate to fictionally ascribe the existence of fictional worlds and fictional states of affairs. Among the states of affairs are those that concern distinctly emotional matters and so we participate with distinctly emotional worlds. Of course these worlds do not really exist [3.5]. The ways in which we speak of them as real worlds merely reflects the rules of the discursive practice that comprises game-bound fictive practice. This, plainly, is my explanation of the ascriptions of attraction we might express for Janet Weiss, or those of pity we might express for Anna Karenina. The emotional states referred to are the apparent commitment of ascriptions arising from a fictive idiom, and as such are a matter of fictive practice.

This seems to explain the ontological, semantic, and logical difficulties in such emotional attributions, but why and how do these attributions of relationships arise? Why are we tempted fictionally to assert we have relationships with fictional characters? Indeed, how can they seem so real? This is one of the most puzzling aspects of fictive participation. There are various things we might take into account when attempting to explain this. Important is what Le Doux says about emotional attribution and what we have observed about the basis of various forms of fictive isomorphism. The clue to how the emotional attributions occur in people goes back to the Schacter and Singer experiments [2.2]. Le Doux thought the real lesson that should be learnt from these is that people are prone to explain inner disturbances by relating them to outer and obvious environmental phenomena. In previous sections I discussed how affective reactions might be wired up *naïvely* to external cues and also to *imaginatively held* thoughts and images. It is certainly adaptive to hold both forms of elicitors affectively significant. Furthermore, it must be remembered feelings are not the only conscious outcome of affect program reactions, as affective states also *direct our attention* toward the emotional stimulus. The affect programs are responsible for those physiological disturbances that focus not only on the natural environment but also on the cognitive emotional environment; that is, the world of human behaviour. The disturbances (feelings and object-directed attitudes) both draw us further into the fictional content and demand we explain what is going on inside of us, just as the Schacter and Singer subjects did.

Fictional *worlds* provide the obvious source of the content of our explanation of our emotion states, and so the explanations of our emotional *relationships* with fictional characters arise from language located within a fictive idiom. As far as our physiological reactions to fictions go – the states that might prompt us to ascribe such attitudes to ourselves – it might again be important to realise the type of information available for self-report is often not the type of information that is used in the processing of the affective emotional state. Le Doux (1998) makes this clear in his remarks about the type of brain systems that are involved in affective processes. Though we may in actuality be experiencing a simple affective arousal caused by imagination, the *opacity* of our affective reaction makes it all too easy to tell ourselves a complex story of a higher cognitive emotional relationship to a fictional individual. *The games we play with fictions may be abetted by the opacity of our emotional states.*

Fictive appreciation often involves characterising the nature of the game we are playing with fictions, a game that reflexively incorporates our mental states as props. The ascription of an emotional relationship to a fictional character amounts to a *reflexive construction of a game-world* in which our own mental states act as props. Abetting the formation of this game-world attitude is a combination of facts about the emotional mind. *Pace* Walton, people are often on a very uncertain footing with their emotions. This may explain why people might be tempted to describe themselves as in love with a character. The affect, caused in us as a result of our participating in the form of pretencefully held thoughts and obvious from the presence of feelings, seems to demand a self-report of the type suitable in real-world circumstances. This is an important way our psychological and, in particular, affective natures seems to draw us into fictions and into participating in the game-bound fictive practice.

It might be fictional that we pity Anna Karenina, fictional that we find George Costanza an embarrassment, or fictional that we fall in love with Janet Weiss. Perhaps at the basis of these apparent attitudes is an isomorphism with real-world states of pity, embarrassment, and attraction. Again, in this case, it is because they may share in a psychological process or mechanism, namely, a cognitive attribution and self-report of an emotional relationship with another individual, perhaps based on the casting of a perceived emotional disruption and the perceived environmental situation into a discrete mental episodic attitude. This isomorphism may also be abetted by the fact that the representational tokens that subserve genuine cognitive attitudes are in the same code as those that subserve our fictive attitudes, so that both have a similar causal profile, especially in respect to what emotional reactions they might cause.

Yet some might want to deny that these fictional relationships exist altogether. This brings me to the third issue. Neill argues that some emotions, such as jealousy, are unavailable to fictive participants because appreciators simply cannot play the requisite fictive role of seeing themselves as relating to a fictional character (1995:254-255). We cannot fictionally be in love with Janet Weiss because such an ascription would demand we could play a fictive role. I think that Walton shows that this is not necessarily true because in fictive practice we often fictionally characterise our interaction with characters in the construction of our game-world. We *do* play fictive roles because our mental states act as fictive props that make it appropriate to see ourselves as fictionally relating to a fictional character.

Neill's point might be given a stronger reading though. Earlier, I left open the possibility that love was conceptually and behaviourally a complex state [2.4]. Though affective reactions *can* be elicited by fictions, and we can be characterised as fictionally fearing or being surprised by fictional worlds, when we come to the socially and behaviourally complex cognitive emotions such as love, things may get more difficult. To have an emotional relationship with someone in these cases would demand our active *behavioural interaction* with them, or perhaps merely the preparedness to so act, in cases where emotion is unrequited. In these latter cases, such as where a movie fan might fall in love with a film star who they have little expectation of meeting, let alone of reciprocating their love, we would expect the individual to pursue lines of behaviour such as writing love poetry, dressing nicely, and hanging out around studios, on the small chance they did meet the individual. If someone who claimed to be in love with a *fictional movie character* did pursue behaviours preparing for such contingencies, however, we would suspect them of being irrational, moreover insane. To do more than to pay lip service to the fact that one would like to have a *relationship* with a fictional character – to dress nicely on the chance of meeting them – really would be irrational in Radford's sense [1.2]. On this view we could never *fictionally* fall in love with a character, because love may in part be constituted by a personal *interaction*, or the behavioural commitment to carry one out should the chance arise, and fictions just are not the type of thing we can physically interact with. If love is essentially a conceptual and behavioural phenomenon, then *fictional love* might be denied appreciators on *substantive* grounds, given that fictive props cannot respond to our behavioural involvement.

However, I think this complaint depends for its force on the contingent features of the fictions we are presently used to interacting with. It is true I could not expect or even hope for the reciprocation of my love for Janet Weiss of the movie *The Rocky Horror Picture Show* if I was to feel that way. Movies just do not respond to our input. But I argued that some fictions *do* involve our overt physical participation, and can respond to our physical interaction [3.6]. Videogames are the most important example of such an interactive fiction. It is possible in the future someone might create a fictive prop that could respond to us in a variety of physical modalities so it might be fictional that we had complex behavioural emotional *relationships*. At this point I will stop, before getting to sordid matters such as *cyber-loving*!

7. Fiction Naturalised

7.1 Fictions, Emotion, and Behaviour

Throughout the literature on fiction and the emotions, a persistent theme is the role fiction, and particularly the fictional emotions, plays in the wider world. Why do we emotionally participate in fictions? Why do we pursue emotional experiences that can often be saddening, repellent, or fearful? What is the moral significance of our emotional participation with fictions (if indeed there is one)? Where did the emotional participation come from and when did it begin? I will restrict my observations here to those that arise from a broadly naturalised vantage point. Among the things I hope to have achieved is the profitable application of naturalised ideas to areas where they have traditionally not had a great deal to say. Here I will finally give in to my desire to speculate on the evolutionary significance of fictive practice.

There are really several issues. The first issue concerns the origin of the cognition at the basis of the involvement of emotion in pretence and fictive practice. How might the behavioural aspects of pretence practice have played a role in the evolution of the cognitive traits at its basis? I make some preliminary observations concerning this question in this present section. A second question, dealt with in the next section, concerns how and why the types of fictions we have might have arisen. Is it possible to say anything from an evolutionary standpoint concerning not only the cognitive traits involved in pretence behaviour, but also the *artefacts* or *props* involved in the fictive practice? To do so would go a long way to providing a basis for a *naturalised philosophy of the arts*. I argue that there will be difficulties involved in naturalising aesthetics.

What might be the adaptive origin of pretence behaviour? It is obvious why it would be adaptive for an organism to count as emotionally significant thoughts about what might happen if it proceeded along certain paths of behaviour, and how such an adaptation as imagination – if imagination is an adaptation – might evolve. I have these persistent thoughts when I am using machinery with whirring bits and sharp edges: I am aware of what it would be like to catch parts of myself within their workings. I sometimes wince when having such thoughts. I generally avoid using machinery with whirring bits and sharp edges. It is likely we have an imaginative faculty for adaptive purposes concerning such things as counterfactual reasoning, long-term problem solving, and planning. It may be one of the central cognitive features of the human mind, freeing us from the present, and expanding our consciousness to include the past and future, allowing us to *track* our environment more fruitfully. The adaptive potential of such a faculty is obvious. It allows me to realise that walking out onto a busy street without *looking left, looking right, looking left again* might put me in a difficult situation involving a bus.

Yet without an ability to find such imaginings emotionally pertinent – that is, without the conjunction of affect and imagination – the past or future could hardly be reasons for motivation in the present. Imagination, and the representational states that subserve it, would be causally and behaviourally inert. It is clear, then, why our emotional systems should be wired into the imagination. It is going to be adaptive for me in the long run to take heed of the unpleasant emotions I have when thinking about nasty chainsaw accidents. The emotions *orientate* us to the content we pretencefully engage with. I have thoughts of nasty chainsaw accidents, and subsequently I avoid chainsaws whenever I can. In his defence of popular art Carroll offers a similar example of how the imagination is tied into affect, except that he enjoins the reader to imagine sticking the knife in their eye (1998:273). Feagin's favourite example of a thought that might engage our affective natures is the *vomit sandwich* (1996). I do not think this propensity to become emotionally aroused by thoughts of nasty accidents with sharp implements or of ingesting disgusting foodstuffs is peculiar to those with an interest in the philosophy of the arts. The conjunction of such counterfactual reasoning and affective or emotional reactions will give people a reason and method by which they can formulate long term plans or coping methods. If it evolved for the reason that it played this role in the lives of organisms ancestral to us, then imagination, and its conjunction with affect, is an adaptation.

Several of the theorists I have been discussing have signalled this issue. Walton thinks "evolutionary pressures may be responsible for our being organisms of a kind susceptible to quasi emotions in situations in which they might enrich our psychological participation in games of make-believe" (1990:245fn). Nichols and Stich suggest an evolutionary function from which our pretence abilities might have evolved, namely "to facilitate reasoning about hypothetical situations" (Nichols and Stich, 2000). Currie has also made such observations (1995b), as has Ruth Millikan in her discussion of the adaptive nature of strategy testing and conditional planning (1984).

Damasio goes a step further than claiming pretence cognition and affect are necessarily conjoined if they are to be behaviourally efficacious, and thus adaptations. His view is the stronger one that without emotions, *cognition generally*, and especially what he calls "practical rationality," will not properly function (Damasio, 1994, 1999). Affective emotions are necessary to the successful functioning of higher cognition and also important to consciousness. Our emotions, as sensed within our bodies, help make us conscious of the environment by enhancing our awareness of the objects in the fictional and external environment with which we must deal. The emotions for Damasio are fundamentally related to homeostasis and to orientating ourselves toward the objects in our environment necessary to securing this. Emotions map us onto, or allow us to track, the significant aspects of the external and internal environments. Remember that Le Doux also thinks part of the role of the affective emotions, such as fear, is to make objects of our perception salient [2.2 & 2.4]. Damasio thinks consciousness involves a *second order* mapping (of these first order maps) that shows how the individual relates to or *knows* the objects in its environment (1999:176-194). Consciousness involves mapping the organism in relation to the objects in its external and fictional environment so that the complex tools of higher cognition, including counter-factual reasoning, can be brought to bear on the problems that result from the demands of homeostasis. Because emotion is essentially involved in homeostasis, and because consciousness is a more complex way of representing to the organism the needs of homeostasis, the objects of consciousness should be drawn from those representations emotion has already gone a long way toward establishing as salient. Given this, imagination would necessarily involve the subtle or not so subtle influence of the emotions in enhancing the content of imagination. It is not merely adaptive to have the imagination conjoined with our affective natures, as I have been claiming; it is necessary that imagination be conjoined to affect if it is to function. Damasio's emotionally driven theory of consciousness provides a theory of pretence with an ideal opportunity to derive the connection between the imagination as employed in make-believe and our affective emotions.

This ties into issues of interpretation of fiction. If affect enhances the objects of the imagination, thereby framing our understanding and participation with imagination, then the emotions are likely to play a large role in our interpretations of fictions. In fact, emotion plays a central role in our interpretative abilities. It does so because emotion plays a large role in our interpretation of the non-fictional world. The cognitive and affective systems that allow us to interpret and act in the actual world subserve our participation in fictional worlds. Robinson claims affective states such as startle help to focus our attention of fictions (1995a) and Currie holds that "secondary imaginings" involve us employing our own affective systems to interpret what characters are thinking and feeling about the situations they are in (1995b; 1997a:54-56). Emotions are knowledge structures that allow us to interpret the environment. The conjunction of these knowledge structures with imagination allows us to interpret fictions.

The connection between affect and imagination, especially its utility in helping us to deal with the *real world*, raises problems for how we are to characterise the affective reactions. Because imaginatively elicited affective reactions are central to the development of emotional coping methods and longer-term adaptive behaviours, they are very much *reactions within this world*. Imagining what it would be like to catch various parts of yourself within the workings of farm machinery, stabbing yourself in the eye with a knife, eating a vomit sandwich, or being run over by a bus, and finding such imaginings emotionally unpleasant, is likely to make you a little more careful in what you do around farm machinery, sharp cutlery, vomit sandwiches, and roads. The emotions are real ones directed at real content. This is the source of Lamarque's doubts that the emotions inspired by fictions are fictional emotions. Affective reactions to pretencefully held thoughts are responses to, and of utility in, the *real world* rather than fictional worlds.

Emotions elicited by fictions are likely to be useful in this world because they might give us an understanding of situations and forms of life with which we have not yet been involved. Such understandings could be useful in the future, both personally and more widely in terms of social groups, in engendering sympathy with others. Feagin thinks this sympathetic and empathic engagement one of the most important aspects of fiction in general (1996). Such games may allow us to rehearse plans of action in the safety of our own heads. The may also allow us to understand others so better to interact with them.

Because they have pretence representations, images, or associations as their proper cause, these affective reactions *are not precisely isomorphic* with the similar common or garden emotions that would arise should the situations represented in the thoughts become a present reality. I used the notion of *isomorphism* to signal a similarity in dynamic structure that engenders fictive practice, in that the aspects were often fictionally characterised. The affective reactions we have to pretencefully held thoughts *strictly* are non-isomorphic in both their afferent and efferent dynamic, yet are similar enough that they may be fictionally characterised. Pretencefully held thoughts obviously contribute in a different manner as inputs to affective reactions. On the efferent side of the equation, affective reactions are different in that their physiological and behavioural manifestations are *very much in this world*, affecting what we do and feel in this world. When I go home from the cinema after seeing *Saving Private Ryan*, and finding the first thirty minutes affectively gruelling, I might be tempted to have more sympathy with war veterans; ANZAC day might come to have more importance for me. Our affective interaction with fictions occurs in the real world, and the reasons for its existence – and there might be evolutionary reasons for its existence – must exist in the real world. This is not surprising, given that really there are no fictional worlds!

7.2 The Ubiquity of Fiction

Why do we participate with fictions, why are they so prevalent, and why do they have the features they do? Fictions are evident in all cultures in the world, and in Western culture, they have developed to a considerable sophistication. Peter Jackson's movies of *The Lord of the Rings* are enormously complex objects, utilising the most sophisticated technology, and guided by highly conventional fictive rules. Yet the reason for their incredible success is likely to be related to the manner in which they emotionally involve us. The creature Gollum in *The Two Towers* is a masterpiece of technology. He is the integration of traditions of dramatic writing and characterisation, acting, filmmaking, but also of the very modern technology of digital animation. Gollum is a complex technological fictive prop. But when I saw Gollum initially recalling his former name Smeagol in *The Two Towers*, this did not prevent me from feeling an enormous sense of pity for the creature. The existence of fictions such as *The Two Towers* is tied to their ability to so affect us. There are a variety of observations to be made on the cultural ubiquity of fictive practice in relation to emotion. I discuss three theories that come from a broadly evolutionary standpoint, and then make my own suggestions.

Steven Pinker has formulated a theory of the arts, including fiction, built on the principles of evolutionary psychology (1997). Though I see some merit in Pinker's theories, they are not free of serious problems. Foremost, Pinker's model of art is unlikely to satisfy many philosophers of the arts to whom it will seem simple-minded and reductionist. Pinker's theory might be called the *cheesecake* or *cultural pornography* theory of art, since art is seen as calculated toward pushing our affective buttons. Pinker says of music that it is a "pure pleasure technology, a cocktail of recreational drugs that we ingest through the ear to stimulate a mass of pleasure circuits at once"

(1997:528). Music, and the other arts, including fiction, are not adaptive, or adaptations, but are *serendipities* based on the lack of fit between our cognitive architectures, which were established in the Pleistocene era, and the modern world, including its technologies:

> Our minds are adapted to the small foraging bands in which our family spent ninety-nine percent of its existence, not to the topsy-turvy contingencies we have created since the agricultural and industrial revolutions. Before there was photography, it was adaptive to receive visual images of attractive members of the opposite sex, because those images arose only from light reflecting off fertile bodies. Before opiates came in syringes, they were synthesised in the brain as natural analgesics. Before there were movies, it was adaptive to witness people's emotional struggles, because the only struggles you could witness were among people you had to psych out every day.
>
> (Pinker, 1997:207)

Many of the aspects of the emotional participation with fictions I have noted, including our affective engagement and our interest in the motivations and emotions of fictional people, are based on ancestrally established cognitive and emotional traits that in the past would have been engaged only in our dealings in the world. More often than not these cognitive traits are seen as *modules* that arose as adaptations to conditions in the Pleistocene. Fiction is a technology that allows us to throw the affective switches and to observe the trials and tribulations of others without the danger or hassle of being embroiled in real situations warranting emotional responses. Pinker also notes that fiction and drama explore universal human themes: "There are a small number of plots in the world's fiction and drama… More than eighty percent are defined by adversaries (often murderous), by tragedies of kinship and love, or both. In the real world, our life stories are largely stories of conflict: the hurts, guilts, and rivalries inflicted by parents, siblings, children, spouses, lovers, friends, and competitors" (1997:427). There have been similar claims concerning the aesthetics of natural environments (Orians and Heerwagen, 1992).

Though there must be something to these explanations, and they fit with much of what I have said here, they also leave a great deal out. A real problem with the possible scope of Pinker's thinking and of the thinking of evolutionary psychology in cases like this is that the type of culture we are discussing when we discuss fiction seems *unprecedented* in the conditions in the Pleistocene. There simply does not seem to be a precedent in hunter-gatherer lifestyles for the type of extremely culturally contingent design features we find in some fictions, especially some of the *art fictions* I have discussed here. Pinker's explanation seems far better as an account of pretence, that is, of the cognitive abilities underlying fictive practice. This ability certainly has its precedents in the Pleistocene. But this does not amount to an explanation of fictions, let alone art fictions.

Furthermore, Pinker does not seem to acknowledge how *conventional* the features of these cultural artefacts are. He contends that a feature of visual perception – namely, the set of mental *assumptions* necessary to interpret a set of stimuli on the retina as *the world* and not just as a pattern of “smeared pigment” – is exploited in depth perspective in pictorial representation, and in illusions (1997:215-218). Indeed, the artworks Pinker intends to account for – photographs, television, film, and realistic painting – are characterised by him as illusions. Pinker presents what is essentially an illusionist theory of art that is unlikely to sit well with the philosophers of the arts whose theories I have been discussing here.[62]

Geoffrey Miller’s theory of sexual selection and cultural development provides another avenue for explaining the cultural ubiquity of fictional practice (2000). Sexual selection is an evolutionary process whereby traits, not in themselves adding to the fitness of the organism, are selected because they advertise the fitness of an organism. The famous example is the peacock’s tail. It is thought the peacock’s tail evolved because it successfully signals to females of the species that the male sporting it is a strong specimen, as only strong specimens are able to *waste* enough resources to grow a large tail. Also, allowing the mechanism to function is an imbalance in the cost of mating between the sexes. Peahens make a significantly larger investment than the males in producing progeny. Therefore, they exist in a *buyers’ market* and can afford to be choosey. Males must compete to secure mating opportunities and, in peacocks, the males compete by *display*. In the past, many similar claims have been made about the growth of the hominid brain that led to *Homo sapiens’* prodigious cognitive powers. Geoffrey Miller extends these claims, and goes so far as to maintain that forms of art arose because they are expensive signals of prodigious cognitive powers. To put it crudely[63], artworks are ornaments aimed at advertising the potency or fitness of the individuals that produced them, by stimulating the senses of prospective mates. Furthermore, “Modern human culture is a vast, collaborative attempt to chart out this space of all possible stimulation, to discover how to tweak our brains in pleasurable ways” (2000:155).

While there is something to this, I am not sure it will do as an explanation of why art fictions have arisen. The time periods in the development of fictive practice are much too small to have been the basis for the development of an adaptation. In consequence, much of the forms particular to fictions will have to do with forces arising from culture. Miller’s theory seems to provide a source of the cognitive complexity underlying our fictive abilities, that is, the prerequisites of aesthetic fictive behaviour. But does this say anything about aesthetics itself? Given that Miller’s story is equally applicable to the psychology of sports cars, we might doubt it is going to be very successful as a theory of art. Miller’s theory is a theory of the motivation of aesthetic behaviours, and not necessarily of the props that play a role in that behaviour.

---

[62] There are various arguments against this type of illusionist theory (Currie, 1995a), and others for it (Anderson, 1996).

[63] Miller is not afraid of putting it crudely: he compares Led Zeppelin’s music to a baboon’s rump, a comparison that, though I can see the logic of it, I am not altogether comfortable with (Miller, 2000:70).

It also might be argued that Miller merely explains the motivational psychology behind the production of fictions, and this does not show why people want to experience them. He does not say much about the *experiential* psychology of aesthetic or fictional practice. This may be an unfair criticism, though. Miller goes to great pains to explain that consumers as well as advertiser of fitness must understand complex sexually selected traits. If art is to be a sexually selected trait, it must not only be worthwhile to produce; it must be worthwhile to select. Sexual selection also provides a story for why it would be fun to experience artworks, because individuals must have a reason for seeking out mating with high fitness partners, and that reason is that they are fun to be around.

Perhaps fiction is as ubiquitous as it is, and fictions have developed to take the form they have, because the fiction *meme* is a potent replicator? Meme theory, introduced to the world by Richard Dawkins (1976/1989), has recently gained several new adherents (Blackmore, 1998; Dennett, 1995; Lynch, 1996). Memes are putative units of cultural inheritance sharing a functional identify with genes. Dawkins originally made the claim that any system that displayed the features of replicating and mutating individuals, an environment of limited resources in which the individuals replicate, and a method whereby those individuals better suited to their environment would come to replicate with more ease and to dominate the population of individuals, evolution by natural selection would take place. He also wondered whether, other than life, this evolutionary process was exemplified in any other system presently in evidence: “But do we have to go to distant worlds to find other kinds of replicator and other, consequent, types of evolution? I think that a new kind of replicator has recently emerged on this very planet. It is staring us in the face. It is still in its infancy, still drifting clumsily about in its primeval soup, but already it is achieving evolutionary change at a rate that leaves the old gene panting far behind” (1976/1989:192). The evolving system is culture and the replicator at the centre of the evolution is the meme. Memes are shadowy things – Susan Blackmore (1998) comes closest to spelling out exactly what they might look like – but a meme would deserve the name if it exemplified the features of evolutionary replicators. As far as fiction goes, meme theorists might claim the fiction meme, particularly those putatively embodied in the various forms fiction take – novels, movies, short stories, and soap operas – are particularly good at grabbing our attention and, thereby, at replicating. In this way they spread through society as fiction has done as a cultural form.

Is fiction culturally ubiquitous because it embodies a meme that is a potent replicator? I am suspicious of such claims. The works I have read that attempt to develop the cultural evolution analogy into a strong claim about a real evolutionary process, are often extremely speculative, and leave large questions unasked. For example, just what are the units of selection that persist over generations and evolve, that is, what are the *identities* that are selected through the evolutionary periods? Ideas seem to mutate, combine, split apart, recombine, with a speed and ubiquity that seems disanalogous to the relative constancy of the replicators, or even vehicles, in genetic theory. The question remains: have adherents of meme theory done enough to establish there is a *substantive functional identity* between the processes of biological and cultural evolution, or does the comparison remain one metaphor among many for characterising cultural development? Indeed, there are other metaphors that seem better suited to explain what happens when an idea *spreads* through society; the *epidemiological* metaphor seems well-suited. It is ironic that one book on memes, Richard Brodie's Virus of the Mind: The New Science of the Meme (1996), makes reference to the competing analogy of epidemiology in its title!

Furthermore, I agree with Sterelny and Griffiths in that it is doubtful there is anything that needs explaining by meme theory; memes might be an explanation in search of a phenomenon (Sterelny and Griffiths, 1999:333-334). In the case of life, it is pretty clear there is something that needs explaining, that is, an organism that seems *designed* to fit its environment. For many years people postulated a *literal* designer to account for the inherent design. Evolution, however, removes the need for a literal designer and we end up with a metaphorical one: evolution. Evolution is a *metaphorical* designer, because, unlike real designers, it is literally unintelligent. In the case of cultural evolution, we already have somewhere to look to find the source of the design inherent in cultural artefacts: *creative people*. Calling Beethoven's Seventh Symphony – or even the famous theme in the slow movement (Dennett, 1995:344) – a meme is a slight on Beethoven, who drove himself to extreme frustration in designing his symphonies; likewise with *Ulysses*, *The Lord of the Rings*, and other complex fictions. Indeed, meme theory, with its claims that creative people are those who are particularly disposed to *catch and spread* memes, significantly undermines our notions of creativity.

Memetics is a different type of theory to the Pinker or Miller theory. Here it is not the cognitive trait that is being explained, but a cultural trait. And this is the problem; has it been shown there are cultural traits in anything but a metaphorical sense? Meme theory seems to be merely a metaphorical framework for discussing cultural development, which is a feature of the universe that presently seems perplexing and not yet able to be described by a single overarching theory. Not enough attention has been paid to the question of when a structural analogy between differing states or processes becomes a substantive theory of qualitative identity.

There is some explanatory virtue in all of these theories, yet each seems to underplay the sophistication of art by looking for one feature to place above the others. Perhaps the accurate characterisation will see the ubiquity of fiction as emerging from a heterogeneous mixture of developmental structures that are only presently hinted at in the various literatures that seek to explain cultural change. Nevertheless, I think our emotions are central to theories of cultural development. I believe a naturalised epidemiology of cultural change allows us to characterise the features in fictive practice and to explain its ubiquity. I mentioned in the Introduction that evolutionary psychologists Tooby and Cosmides think common features of human life, such as language and, perhaps, fictive practice, "may be explained as the expression of our universal psychological and physiological architectures in interaction with the recurrent structure of the social or non-social world" (1992:117-118). The traditional studies of fiction have more often than not been interested in features of fictions seen as *culturally particular*, owing their existence to their place in a realm of cultural or social change autonomous of biology, physiology, and perhaps even psychology.

Yet evolutionary psychologists make extremely sensible points concerning the inherently *biological* nature of culture in their revision of the nature/nurture debate. They argue that the existence of the particularities of cultures demands a high degree of pancultural uniformity in the cultural and biological interactions that ground psychological development (Tooby and Cosmides 1992:77-93). They also claim that the concept of culture has been insufficiently examined in the past. Tooby and Cosmides make distinctions between "metaculture," "evoked" culture, and "epidemiological" culture (1992:121). The first amounts to the type of pan-human cultural features that are the precondition for humans to socialise within a culture at all. These are generated by the pan-human cognitive modules evolutionary psychologists see as the basis of human cognition. Tooby and Cosmides argue that anthropologists have too often been blind to the existence of the metacultural regularity that establishes their ability to carry out ethnographic study at all. *Evoked culture* accounts for the within-group similarities and between-group differences that lead ethnographers to make the claims of substantial between-culture differences that they so often do. As the name implies, these cultural differences are *evoked* by the particular environmental and social conditions in which a culture finds itself, yet they are evoked *from* a pan-human cognitive architecture. *Epidemiological* culture concerns cultural particularities, such as the representations that are transmitted between individuals; that is, the types of informational structures that seem to *spread* through a society. Though epidemiological representations are contingent upon historical precedents – the type of representations present in a society will depend on the particular representations that were established earlier – "domain-specific mechanisms [will] influence which representations will spread through a population easily, and which do not" (Tooby and Cosmides, 1992:121).

It is the transmitted or *epidemiological* portions of fictive practice that have been most studied. The lexicon of *genres*, *schools*, and *movements* – the language spoken in the humanities – is a lexicon of transmitted cultures. Yet metaculture, and the physiological and psychological facts that underlie it, do play an important role in generating the transmitted culture. In the case of fiction, it may be a dual set of psychological facts that grounds our fictive abilities and gives rise to the ubiquity of fictive practice. Firstly, we have the ability to envisage possible or imagined situations. This allows us to track our environment more profitably and to allow for future states of affairs. The cognitive mechanisms underlying this ability are beginning to be described and are certain to be central to fictive practice. I have spent a great deal of time in this thesis drawing together the theories of these states that exist in the philosophical literature. Of foremost importance is the Nichols/Stich cognitive theory of pretence. Nichols and Stich think, "that the mind uses the PWB for a variety of tasks including mind reading, strategy testing, and empathy," and they suspect "that the original evolutionary function of the PWB was rather to facilitate reasoning about hypothetical situations" (Nichols and Stich, 2000:122). This pan-cultural ability of pretence will make possible the evoked cultural practice of fiction where the extra-machinery of fiction – the spoken and written word and, increasingly, other forms of electronic media – is available. Walton's theory of fictions as *props* that play a role in our games of pretence, making them richer and more objective, is a theory that shows how transmitted culture trades in a currency of pan-cultural cognitive traits.

Second, our affective and emotional systems are involved in the cognitive ability of pretence, probably for adaptive reasons [7.1]. I think it will be the case that our affective and emotional natures play a central role in determining what types of transmitted fictive form will be prevalent. The props that have developed in fictive practice – the novels, television sitcoms, videogames, and movies – exist because they are entirely apt to engage our emotional natures. In this I agree with Steven Pinker's observations that fictions will reflect the content of actual human lives (1997:427). The psychological and physiological features of fictive practice that have been my subject – the ability of pretence, affective reactions, feelings, emotional evaluations, and emotional *relationships* – may generate our involvement in fictions and may determine the character of those fictions that spread more easily through culture. Studying the genres, schools, and movements of epidemiological culture may not be enough. We must pay attention to our underlying biology and to how it generates the possibility of culture.

In the case of fiction and the emotions, our biological core provides the means to explore a *rich set of problem spaces* in fictive works that epidemiological culture has elaborated on and transmitted to us. The constancies of fictive practice are likely to be based on the underlying psychological mechanisms, and the cultural particularities – such as Modernist novels, Japanese Noh drama, pantomime, videogames, soap operas, modern digitally enhanced movies such as *The Lord of the Rings* – owe their existence to the features of other dynamic cultural pressures, some of which are discussed by the social sciences and some of which are yet to be discovered.

Does fiction have a function? Is it an adaptation? I think this unlikely. If anything has a function or is an adaptation in fictive practice it is the *cognitive and affective systems that underlie it*. The fictions that engage us so forcefully may be a serendipitous side effect of established cognitive traits, but one with immense complexities due to the forces that shape culture.

8. Conclusions: Emotional Participation in Fictions

8.1 Representational Participation

Before I summarise and discuss how we emotionally participate in fictions, I think it necessary to discuss our non-emotional participation. As previously argued, this participation falls roughly into two kinds, though any fictive appreciation inevitably involves measures of both of these forms of participation because they are invariably mixed. First is the participation that does make an *apparent* acknowledgement of the fictive nature of the fiction, and second is the participation that does not make an *apparent* acknowledgement of its fictive nature. The first type is *representational* participation and the second is *isomorphic* participation. The inclusion of the qualifier “apparent” is necessary in our characterisation of our appreciative activities. Though there are exceptions based on interesting features of our psychology [5.3], our appreciative attitudes toward fictions must make an *actual* acknowledgement to the fictive status of fictions. That isomorphic participation does not make an *apparent acknowledgement* signals the fact that such participation does not make an immediate reference to the fictionality of the appreciated material. It seems to be of a kind with our relationships with aspects of the non-fictional world. In this, it can be fictionally described as relationships and mental attitudes toward fictional people and situations.

This *apparent ignorance* of fictive status by our mental and linguistic participation with fiction is the aspect of our fictive appreciation that originally seemed paradoxical. If our appreciative attitudes toward fictions did not make an actual acknowledgement of fictive status, *we would be mistaking fiction for fact*. Indeed, if my characterisation of the various *naïve* manners of emotional interaction is accurate, then these states do involve the affective systems mistaking fiction for fact, a mistake that is understandable given that they are *quick and dirty* [2.4]. This paradoxical confrontation with an *actual* acknowledgment and an *apparent* non-acknowledgement led Radford to conclude that appreciators of fictions are irrational and inconsistent [1.2]. The situation is not paradoxical and appreciators are not irrational, because the basis of the apparent non-acknowledgement, that is, *isomorphic participation*, is based within games of cognitive pretence (though again, there are exceptions [4.3]).

Representational participation, then, does make an *apparent* acknowledgement of the fictive status and conventions of fictions. In this it is not apt to be fictionally characterised (or characterised as being within a game-world). There are various examples of non-emotional representational participation I can introduce to make this clear. In the case of written fictions, the most important act of representational participation is the understanding that appreciators are to adopt a pretence-bound attitude toward the descriptions found in the works of fiction, and *not* to the biographical and publication information found in the pages before the text proper. It is the *central* act of representational participation to understand that we are to adopt an attitude of what Walton would call make-believe to the descriptions and other representational features of fictions. This is not likely to be fictionally characterised; it is not usually fictional of the world of the fiction that it is meant to be appreciated by readers via an act of make-believe or that it is described in a fictive work. Unless, of course, it is a *strange* fictional world, rather like the *second order* fictional world that seems to intrude knowingly into the *first order* fiction of Woody Allen's *Deconstructing Harry* or the marginally coherent *Adaptation*.

Likewise, it is an act of representational participation when we note the features of the author's language in describing the fictional material or world that is to be seen through the pretence-bound participation central to fictive practice. Joyce's *Ulysses* contains some spectacular representational stylings. But it is not the inane content of some of the passages that makes them remarkable; it is the representational form of the narrative. In the *Ithaca* section of *Ulysses*, Joyce's odd and pedantic dialectical style seems the most unlikely of manners in which an author might relay content that is to be pretencefully held or make-believed. Indeed, much of the chapter – which is over 20,000 words long – seems to frustrate the reader's ability to participate successfully with the fictional world. Though the example of *Ulysses* is an extreme case, coming from the forefront of avant-garde Modernism where it became increasingly customary to draw attention to the qualities of the representation, much fiction demands we respond to it in ways that cannot be characterised in fictional idioms. When we admire and take pleasure in cinematography and editing techniques, we are responding to features of fictions that engender representational participation. It is not usually fictionally true of the world of the fictions that any editing, cinematography, or other representational conventions are taking place. The exceptions to this are the cases of *nested* fictions of which the famous example is *Hamlet*, and fictions about people creating fictions, such as the recent film *State and Main*.

8.2 Isomorphic Participation: Theoretical and Reflexive Attitudes to Fictions

Yet when we observe the characters in a movie such as *Magnolia* quickly going about the business of tormenting themselves, the swift editing does contribute to the fictional state of affairs of the film. The excellent editing allows the events to be both hectic and integrated; as we watch, the states of affairs quickly compound and conspire. It is fictional that the events are quickly moving and compounding in terms of their effects, and also fictional that we see so. Observing the speed and implications of these events is an act of isomorphic participation. Though depending on editing to achieve the presentation (and isomorphic participation is generally dependent on representational participation), the engagement that allows us to see the speed and implications of the fictional events makes reference to facts fictional to the work-world. We see how Earl Partridge's attitudes toward death and to his deceased wife at the time of her death have affected his son T. J. Mackie in his attitudes toward women and toward his father; we see how, in all his hurry, nurse Phil Parma has just dropped a bottle of painkillers and we observe that one of those dogs is just bound to end up dead, something confirmed in the dénouement.

Isomorphic participation comes in various forms. We decide Bloom has a bushy moustache; that Dedalus looks like a youthful Joyce; that parts of Dublin, especially the parts frequented in the *Circe* episode, are grubby. Sometimes, as in movies and television, the images are ready-made, and we are naturally described as gazing on in amazement at the Roman Colosseum. We are interested in all the sorts of non-emotional truths presented in works of fiction. We sometimes appraise the *moral* situations presented within fictions. Often we desperately try to *make sense* of exactly what is happening in fictional worlds; in Terry Gilliam's *Brazil*, we wonder what exactly is real and what all those odd bureaucrats are doing, and by the end of the movie we may simply be confused, perhaps making it fictional that we are confused!

Drawing on distinctions made by Walton and Feagin, we see there are two main types of non-emotional isomorphic participation: theoretical, or generative participation, and reflexive participation. Theoretical modes of isomorphic participation involve, not surprisingly, our *theorising* about the content of the fiction, attempting to derive truths about or *elaborate* on what Walton calls the work-world of a fiction. We note that Gonzo the Muppet has just learnt he is not from this earth and we decide he will want to find his alien origins. We suspect the handsome Ricardo will walk in on his wife, the buxom Sabrina, embraced in the arms of the dastardly Dirk. We conclude that the butler all along was the only one with access to the gun cupboard around the time that Colonel Reginald Upthwaite became the *late* Colonel Reginald Upthwaite. Sometimes, given the triteness of such plots, the theoretical leap imposes no strain on our interpretative capacities, and we judge that *we could see that coming from a mile*. Conversely, sometimes the facts of a fictional world elude the appreciator, and it is only in the dénouement of the fiction that the appreciator is able to make the appropriate elaborations and understandings. The film *The Usual Suspects* is an excellent example of a work where the inferences made in interpreting a fiction are turned on their head in the final scenes. Furthermore, as Nichols and Stich remind us, theoretical and generative elaboration of this work-world comes in two further flavours: inferential elaboration and non-inferential elaboration, the former directly inferred from the premises of the pretence, the latter elaborated by the appreciator under their own steam. All of this theorisation or generation arises because, within a game of make-believe we *pretend* to adopt attitudes and to theorise about the people and situations that we pretend exist. Fictional people and situations are fictionally assumed to have the same sort of properties as normal people and to be subject to the same type of reasoning that is applicable to the real world. The reasoning is isomorphic, which means that our theoretical activities will have the surface appearance and structure of normal real-world theoretical activities. There is some functional identity.

Reflexive participation also involves theory in this sense, but is different in that the theory is focused on our place in regard to fictional worlds, since such reflexive participation regards the generation of game-worlds. The theoretical observations make reference to an observer as well as an observation. *I note* that Gonzo the Muppet has just learnt he is not from this earth and *I suspect* that he will want to find his alien origins. Fictive practice demands that we ascribe mental states to appreciators. This is on par with mental attitudes as they occur in the real world; theoretical observations are not disembodied phenomena but are the psychological actions of subjects. Such reflexive theoretical attitudes must also be fictional, however, because it cannot be the case that such ascriptions are true simpliciter of appreciators of fiction. They are true only within a game-world appreciators play with fictions; that is, they are fictionally true. These reflexive theoretical observations, making apparent reference to appreciator's attitudes, prove of much importance when we come to characterise the nature of the fictional emotions.

Invariably, isomorphic and representational modes of participation will be mixed, since any successful fictive appreciation will make necessary reference to the fictionality of the material in coming to both an interpretation and an evaluation of the work; otherwise we might as well be mistaking fiction for fact. A successful appreciation of most fictions involves *integrating* our representational participation with poetic effects, editing, cinematography, and other artistic contrivances, with our isomorphic participation with people and situations and plights. If we are prone to finding *messages* in works of fictions, it is only representational participation that will allow us to find these messages; the fiction must be seen as a communicative act between an author and an audience. We often engage with fictional worlds to appraise how *life-like* or *truthful* they are; doing so requires direct reference to the fiction *qua fiction*. We become interested in interpretative issues, asking, for example, what Hermann Hesse's eventual attitude in *The Glass Bead Game* was to the utopian intellectualism he portrays; why does the central character Joseph Knecht leave the cloistered schools of Castalia and then promptly die? In making such judgements, we are likely *both* to adopt a pretenceful stance toward the contents of the novel – reasoning about Knecht's motives, for example – and also to make reference to the factual features of the author's artifice – for example, asking whether Hesse changed his attitude toward the possibility of a utopian intellectual state during his writing of the work. We may even quietly commit what some call the *intentional fallacy* (Wimsatt and Beardsley, 1978), by pondering on the author's domestic situation and on how this might have contributed to his production of the novel. And surely such participation is often warranted; a reading of Richard Ellmann's masterful biography of Joyce is bound to rightly influence our interpretations of that author's works.

8.3 The Paradox Resolved

The paradox of fiction grows out of a tension between *apparent* and *actual* reference to fictional entities by intentional states and semantic representations. These intentional and semantic states are most apparent from the surfaces of verbal and written sentences that seem to make observations about fictional worlds. The paradox is historically evident in the philosophy of language; especially in the thread that deals with the apparent reference fictive proper names hold to fictional people. Our intentional attitudes also seem paradoxical for just the same reason:

ascriptions of our propositional attitudes make an apparent reference to people and entities where it is plain that people and entities do not exist. The resolution to the tension between apparent and actual commitments is the realisation that the apparent commitments are something like *pretended commitments*. They are commitments, obvious from our verbal and written observation sentences concerning fictional worlds and from our mental ascriptions concerning our own attitudes, made within pretence-bound games. The emotions share in this paradox in as much as they are seen as intentional states, or even causal states, directed at or caused by *people and situations*. The *paradox* of the fictional emotions is a symptom of a more fundamental problem and is resolved in a similar manner. The emotions we have for the people and situations within fictional worlds are obvious from our observation sentences and mental ascriptions concerning our fictive participation. Appreciators do not have emotions *for* fictional characters; it is fictional they have such emotions for fictional characters. It will be true simpliciter that they have genuine emotions cased by fictive props, however, if we loosen our theory from the grips of a very naïve cognitivism, as I have argued we should.

8.4 Emotional Participation

There is much that is of interest in our appreciative activities that has a specifically emotional focus and there *are* problems of the emotions and fiction. For example, what is the cause of the affect and feeling in us that motivates us to pretend to attribute intentional emotional relationships or states to ourselves? How does our reflexive attribution of emotions to ourselves relate to how we attribute emotions to ourselves in real-world occurrences of emotional episodes? What is the relation of the imagination to emotional states and how is this incorporated into fictive practice? Are there distinctly emotional facts we theorise about? These questions are dealt with when we begin to detail the nature of our emotional participation in fictions.

Though our emotional participation in fiction shares many of the forms of the non-emotional participation discussed directly above, coming in *representational* and *isomorphic* varieties, it also arises in its own distinct types. Indeed, I think that because it is likely that emotions do not form a natural kind but are diverse in kind, there is bound to be heterogeneity in our emotional participation in fictions.

8.5 Representational Participation and Emotion

Though my focus has been on the isomorphic generation of fictive emotion, I also think that representational participation provides a significant class of our emotional participation in fictions. Much of the emotional appreciation that is aroused by movies, for example, is directed to the representational form of those movies, and only some of these emotional attitudes – those that are relatively isomorphic with real world processes – might be later characterised fictionally. Soundtracks can elicit strong emotions, often highlighting aspects of the film with which we must isomorphically interact. A well-placed swell in the strings alerts us to the fact that the scene involves much passion and drama; a sudden piercing and angular phrase primes us to react with horror as yet another innocent college student is butchered by the man in the hockey mask. Editing and cinematography can achieve similar emotional effects. Often, representational conventions, in the form of narrative techniques, press our emotional involvement with fictive works. The use of black and white film stock might make us nostalgic if we happen to pine for the golden age of Hollywood. In Spielberg's *Shindler's List*, a representational technique adds to the emotional effect of our isomorphic relationship to an aspect of the content. Though the film is in black and white, several sequences introduce colour to heighten the effect of the material. A small Jewish girl stands out in a sequence where the Nazis are clearing out a Polish ghetto because her jacket is red, while the rest of the scene is the black and white of the film stock. While we are not meant to make-believe the girl is the only coloured thing in the work-world, the technique adds to the emotional effect when later we see her body, still dressed in the red coat, on the back of a dray, because it allows Spielberg to direct our attention. Soap operas also have a wide range of all too familiar representational *tricks of the trade* that help to make them more emotionally evocative. Leaving the action on a cliff-hanger at the end of each show is the most recognisable of these, though there are many others. Much more could be said about such fictive participation.

8.6 Isomorphic Participation and Emotion: Theoretical, Affective, and Reflexive Emotional Attitudes

I think our isomorphic emotional participation in fictions comes in three varieties, corresponding roughly to Feagin's theoretical, affective, and reflective modes of participation. Our *theoretical* isomorphic emotional participation arises when we theorise about the emotional significance of the content within fictions. In such cases it is fictional that I am so simulating or theorising. The participation is then of an isomorphic kind as it fits with what we might do were fictional situations real. There is a problem here though. The debate between theory theory and simulation theory seems far from settled [2.4, 6.2] and, if simulation theory should turn out to be at least partly correct, I think our simulative participation in fiction would better be classed as an *isomorphic affective* attitude, given that it relies on running our own affective system offline. Anyhow, I do not think the disagreement between simulation and theory theorists affects my argument here; either of these activities is easily characterised as an isomorphic form of participation where it is focused on fictional people and situations. Also, whichever form is involved in our interpretative abilities, our generative participation with fiction will thereby involve *domain-specific* knowledge structures particular to our ordinary emotional lives.

Our *affective* isomorphic participation arises when the content within fictions, whether pretencefully held or reacted to in a naïve manner, elicits affective reactions in us. The reactions are short-term stereotyped physiological and neurophysiological states. Often such reactions will arise from our theoretical activities concerning fictions, as when we realise the hunky Steve has yet again cheated on the demure Mavis with the brazen Maggie, and our blood boils. We pretencefully hold such thoughts before our attention and affective anger (or quasi-anger as Walton might call it) arises in us.

Sometimes the affective reactions will be activated in what I have been calling a naïve manner. Though many have reservations about calling startle an affective reaction, I think we should include it in the class of naïve reactions to fictional stimuli that do not make an actual acknowledgment of the fictional nature of fictions and yet seem to draw us into fictions. Other such *affective* reactions are the laughter caused by canned laughter, the anxiety caused by editing and cinematographic techniques, reactions to expressive features of human faces and vocalisations, sexual arousal elicited by sexually explicit representations, disgust responses to disgusting stimuli, and naïve responses to the objects of our fears and desires. Without too much difficulty, we could also imagine an avant-garde artist producing a film composed of people communicating entirely by graphically represented yawns, and thereby causing most of the audience to respond with yawns of their own.[64] This would be a case of a naïve isomorphic affective reaction to fictive stimuli. By the way, are you yawning yet?

Our reflexive isomorphic participation in fictions arises when we emotionally frame others and ourselves in respect to fictional characters and situations. This will often happen when we note in ourselves an affective disturbance and, on that basis, attribute to ourselves an emotional state directed at a fictional character or situation. Film spectator Charles, noting in himself the presence of affective fear, attributes to himself an attitude of fearing the slime. This self-attribution and attitude may follow the production of affect in either a naïve or imaginatively elicited manner, or the observation of such in others, and it is isomorphic because the process of attribution shares a structure with a feature of our real-world emotional life [2.4]. Therefore, such an ascription might be initially confused, develop over time, or be quickly revised as further evidence comes to hand.

8.7 Conclusion: The Heterogeneity of Emotional Participation in Fiction

Are the emotions we have concerning fictional characters real or fictional, then? There is room for them to be both real and fictional, though they are also a variety of different things [2.4] and are characterised as *real* or *fictional* in response to distinct problems [3.8]. Strictly, of course, to say *the emotions* are both real and fictional is to commit an equivocation across idiomatic loci. If our focus is directed toward the seeming intentional attitudes we have concerning fictional characters – that is, to people – then the emotions can be seen only as fictional. Real psychological attitudes to such people do not exist, because real such people do not exist. To explain this I need to introduce the variety of theories canvassed here. Considered *substantively* as responses to *external* characters, that is, considered as the senses of sets of descriptions, or as responses to images, sounds, and representational aspects of fictions, the emotions, more precisely the affective emotions, are real.

---

[64] If I did not suspect this was *bound* to already have been done, I would be tempted to make my directorial debut!

Such things are the proper cause of emotional elicitation. The explicative principle here is an attention to various substantive facts about humans: imagination, pretencefully held thoughts, images, perceptual cues, *tokens of representational states in a possible world box*, all provide a set of eliciting conditions for affective emotions and other affective reactions.

Furthermore, the emotions we have in connection with fictions – as part of the process of fictive appreciation that draws us into fictions – are a diverse bunch. Affective and physiological reactions, feelings, judgments of the emotions of characters, and reflexive cognitive appraisals of our relationships with fictional folk all seem to involve us emotionally in fictions.

I think there are various observations that follow from this diversity in kind. We might suspect the term “make-believe” is not a truly accurate one when describing the nature of our interactions with fiction, especially concerning our emotional interactions with fictions. The use of the term implies too much to be useful. Certainly, there is an element of this interaction it captures perfectly and that any theory of fictive participation will have to explain (even, as I found, a *thought theory*). This aspect of fictive practice is the *pretence* or *make-believe* attitude we hold toward the content within fictions, and subsequent fictional reasoning and self-report, identified by Walton and Currie and further explained by Nichols and Stich. But there is a great deal of other *stuff* that goes on between us and fictions and much of that is not usefully described as involving beliefs or narrowly characterised contentful mental representations. To explain our interactions with fictions, the term “make-believe” may have outlived its usefulness because it is too specific and implies too much. Walton’s theory is an important characterisation of fictive practice, but I hope to have shown there are a variety of states and attitudes involved in fictive practice that Walton does not detail. We should switch to a more inclusive term better suited to the job of characterising a heterogeneous set of attitudes and other mental phenomena that relate us to fictions. My choice is the purposely deflationary term “fictive practice.” We can understand it to denote not only the rule-governed and pretence-bound theoretical practices involved in fiction appreciation, but also the sort of *emotional* and *affective* responses that have been the subject of this essay. Fictive practice, especially in terms of emotion, is composed of a heterogeneous set of mental states and attitudes, just as our participation with the real world is:

- ☐ Cognitive and affective reactions to the representational features of fictions.
- ☐ Cognitive pretence relationships.
- ☐ Pretencefully (imaginatively) elicited affective reactions.
- ☐ Naïve affective reactions to emotionally loaded stimuli.
- ☐ Affective reactions to socially particular stimuli.
- ☐ Higher cognitive and social emotional theorisation/simulation.
- ☐ And, interestingly, reflexive cognitive appraisals of affect program and higher cognitive emotion effects as *relationships* to fictional worlds.

Fictive practice has long been characterised as involving states of make-believe. I think the most important of my general conclusions here is that we should be willing to dispense with the lexicon of “make-believe,” “make-beliefs,” and even of “pretence,” if we have been prone to see these terms as telling the exclusive story of fictive practice. Though these states are involved in fictive practice, and perhaps centrally involved, I think they are also overly specific and exclusive, as well as misleading concerning the types of physiological states, psychological states, and extra-psychological processes that go into our participation with fiction. I think fictive practice involves the employment of a prodigious array of mental states. Some of these states are in no easy way reconciled with a simple belief-desire psychology of the type held by many cognitivists. This reflects my general view of a fractured, modular, and often-opaque human psychology, a view I draw from a variety of sources and that is common within the naturalised thread of the study of the mind. The fictive stance, like the psychology at its basis, involves a range of stances.

Bibliography

Anderson, Joseph D. 1996. The Reality of Illusion: An Ecological Approach to Cognitive Film Theory. Carbondale and Edwardsville: Southern Illinois University Press.

Averill, J.R. 1980. “Emotion and Anxiety: Sociological, Biological, and Psychological Determinants,” in Explaining Emotions, A.O. Rorty, (Ed.) 1980, Los Angeles: University of California Press.

Barkow, J.H., L. Cosmides, and J. Tooby, Eds. 1992. The Adapted Mind: Evolutionary Psychology and the Generation of Culture. New York: Oxford University Press.

Beardsley, M.C. 1978. “The Aesthetic Point of View,” in Philosophy Looks at the Arts: Contemporary Readings in Aesthetics, Joseph Margolis, (Ed.) Philadelphia: Temple University Press.

Blackmore, S. 1999. The Meme Machine. Oxford: Oxford University Press.

Carroll, N. 1990. The Philosophy of Horror or Paradoxes of the Heart. New York: Routledge.

Carroll, N. 1995. “Enjoying Horror Fictions: A Reply to Gaut,” *British Journal of Aesthetics* 35(1):67-72.

Carroll, N. 1998. A Philosophy of Mass Art. Oxford: Clarendon Press.

Carroll, N. 1999. Philosophy of Art; A Contemporary Introduction. Routledge Contemporary Introductions to Philosophy, (Ed.) P.K. Moser. London: Routledge.

Churchland, P.M. 1981. “Eliminative Materialism and Propositional Attitudes,” in Mind and Cognition, W.G. Lycan, (Ed.) 1990, Oxford: Blackwell.

Churchland, P.S. and Terrence J. Sejnowski. 1990. “Connectionism and the Philosophy of Mind: An Overview,” in Mind and Cognition, W.G. Lycan, (Ed.) 1990, Oxford: Blackwell.

Cosmides, L., J. Tooby, and J.H. Barkow. 1992. “Evolutionary Psychology and Conceptual Integration,” in The Adapted Mind: Evolutionary Psychology and the Generation of Culture, J.H. Barkow, L. Cosmides, and J. Tooby, Eds. 1992, Oxford: Oxford University Press:3-15.

Crittenden, C. 1991. Unreality: The Metaphysics of Fictional Objects. New York: Cornell University Press.

Currie, G. 1990. The Nature of Fiction. Cambridge: Cambridge University Press.

Currie, G. 1995a. Image and Mind. Cambridge: Cambridge University Press.

Currie, G. 1995b. “Imagination and Simulation: Aesthetics Meets Cognitive Science,” in Mental Simulation: Evaluations and Applications, A. Stone and M. Davies, eds. 1995, Oxford: Basil Blackwell.

Currie, G. 1997a. “The Moral Psychology of Fiction,” in Art and its Messages, S. Davies, (Ed.) 1997, Pennsylvania: Penn State Press.

Currie, G. 1997b. “The Paradox of Caring; Fiction and the Philosophy of Mind,” in Emotion and the Arts, Hort, M. and S. Laver, Eds. 1997, Oxford: Oxford University Press.

Currie, G. & Kim Sterelny 2000. “How to Think About the Modularity of Mind-Reading,” *Philosophical Quarterly*, 50:145-160.

Damasio, A. 1994. Descartes’ Error: Emotion, Reason, and the Human Brain. New York: Avon Books.

Damasio, A. 1999. The Feeling of what Happens: Body, Emotion, and the Making of Consciousness. London: Vintage Books.

Darwin, C. 1872/1999. The Expression of the Emotions in Man and Animals. Three (Ed.) London: Harper Collins.

Davies, M. and Stone, T. Eds. 1995a. Folk Psychology: The Theory of Mind Debate. Oxford: Blackwell.

Davies, M. and Stone, T. Eds. 1995b. Mental Simulation: Evaluations and Applications. Oxford: Blackwell.

Davies, S. 1994. Musical Meaning and Expression. New York: Cornell University Press.

Davies, S., (Ed.) 1997. Art and its Messages. Pennsylvania: Penn State Press.

Dawkins, R. 1986. The Blind Watchmaker. London: Penguin.

Dawkins, R. 1989. The Selfish Gene. 2nd (Ed.) Oxford: Oxford University Press.

Dawkins, R. 1998. Unweaving the Rainbow. London: Penguin Books.

Dennett, D.C. 1987. The Intentional Stance. Cambridge, MA: MIT Press/Bradford Books.

Dennett, D.C. 1988. “Quining Qualia,” in Mind and Cognition, W.G. Lycan, (Ed.) 1990, Oxford: Blackwell.

Dennett, D.C. 1991. Consciousness Explained. London: Penguin Books.

Dennett, D.C. 1995. Darwin’s Dangerous Idea: Evolution and the Meanings of Life. London: Penguin Books.

Devitt, M. 1984/1991. Realism and Truth. 2nd Edition. Oxford: Basil Blackwell.

Dilworth, J. 2002. “The Fictionality of Plays,” *The Journal of Aesthetics and Art Criticism*, 60(3):263-273.

Eco, U. 2000. Kant and the Platypus. London: Vintage.

Ekman, P. 1972. Emotion in the human Face. New York: Pergamon Press.

Ekman, P. 1980. “Biological and Cultural Contributions to Body and Facial Movements in the Expression of the Emotions,” in Explaining Emotions, A.O. Rorty, (Ed.) 1980, Los Angeles: University of California Press.

Ekman, P., and W. V. Friesen 1975. Unmasking the Face: A Guide to Recognising Emotions from Facial Expressions. Englewood Cliffs, N. J.: Prentice Hall.

Engell, J. and W. Jackson Bate Eds. 1983. Biographia Literaria: The Collected Works of Samuel Taylor Coleridge. New Jersey: Princeton University Press.

Feagin, S.L. 1983. “The Pleasure of Tragedy,” *American Philosophical Quarterly*, 20(1):95-104.

Feagin, S. 1996. Reading with Feeling. New York: Cornell University Press.

Feagin, S. 1997. “Imagining Emotions and Appreciating Fiction,” in Emotion and the Arts, Hort M., (Ed.) 1997, Oxford: Oxford University Press.

Fodor, J. 1975. The Language of Thought. New York: Crowell.

Fodor, J. 1983. Modularity of Mind. Cambridge MA: MIT Press/Bradford Books.

Fodor, J. 1987. Psychosemantics. Cambridge MA: MIT Press/Bradford Books.

Gaut, B. 1993. “The Paradox of Horror,” *British Journal of Aesthetics*, 33(4):333-45.

Gaut, B. 1993b. “The Enjoyment Theory of Horror: A Response to Carroll,” *British Journal of Aesthetics*, 35(3):284-89.

Gazzaniga, M.S. and J. Le Doux. 1978. The Integrated Mind. New York: Plenum.

Gould, S.J. and R. Lewontin 1979. “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist programme,” *Proceedings of the Royal Society of London*, 205:581-98.

Greenspan, P. 1988. Emotions and Reasons. New York: Routledge.

Grice, H.P. 1989. Studies in the Way of Words. Cambridge, MA: Harvard University Press.

Griffiths, P. (Ed.) 1992. Trees of Life: Essays in the Philosophy of Biology. Dordrecht: Kluwer.

Griffiths, P. 1997. What Emotions Really Are. Chicago: University of Chicago Press.

Hacking, I. 1995. Rewriting the Soul: Multiple Personality and the Sciences of Memory. New Jersey: Princeton University Press.

Harré, R. (Ed.) 1986. The Social Construction of the Emotions. Oxford: Blackwell.

Haviland, J.M. and M. Lewis, Eds. 1993. Handbook of Emotions. New York: The Guilford Press.

Herz, J.C. 1997. Joystick Nation: How Computers Ate Our Quarters, Won Our Hearts, and Rewired Our Minds. Little Brown.

Hort, M. and S. Laver, Eds. 1997. Emotion and the Arts. Oxford: Oxford University Press.

Jackson, F. 1982. “Epiphenomenal Qualia,” *Philosophical Quarterly*, 32:127-36.

James, W. 1884. “What is an Emotion?” *Mind*, 9:188-205.

Joyce, J. 1922/1993. Ulysses: The 1922 Text. Edited with an Introduction by Jeri Johnson. Oxford: Oxford University Press.

Kenny, A. 1963. Action, Emotion, and Will. London: Routledge and Kegan Paul.

Kivy, P. 1990. Music Alone: Philosophical Reflections on the Purely Musical Experience. Ithaca: Cornell University Press.

Lamarque, P. 1983. “Fiction and Reality,” in Philosophy and Fiction: essays in Literary Aesthetics, Peter Lamarque, (Ed.) Aberdeen: Aberdeen University Press.

Lamarque, P. 1995. “Tragedy and Moral Value,” in Art and its Messages, S. Davies, (Ed.) 1995, Pennsylvania: Penn State Press.

Lamarque, P. 1996. Fictional Points of View. New York: Cornell University Press.

Lamarque, P. and Stein Haugom Olsen. 1994. Truth, Fiction, and Literature. Oxford: Clarendon Press.

Le Doux, J. 1993a. “Emotional Networks in the Brain,” in Handbook of Emotions, J.M. Haviland and M. Lewis, Eds. 1993, New York: The Guilford Press.

Le Doux, J. 1993b. “Cognition versus Emotion Again – This Time in the Brain,” *Cognition and Emotion*, 7(1):61-4.

Le Doux, J. 1998. The Emotional Brain. London: Weidenfeld & Nicolson.

Levinson, J. 1997. “Emotion in Response to Art; A Survey of the Terrain,” in Emotion and the Arts, M. Hort, and Sue Laver, (Ed.s) 1997, Oxford: Oxford University Press.

Levinson, Jerrold (Ed.) 1998. Aesthetics and Ethics: Essays at the Intersection. Cambridge: Cambridge University Press.

Lewis, D. 1983. “Truth in Fiction,” in Philosophical Papers, vol.1. New York: Oxford University Press.

Lycan, W.G. (Ed.) 1990. Mind and Cognition: A Reader. Oxford: Blackwell.

Lynch, A. 1996. Thought Contagion: How Belief Spreads Through Society. New York: Basic Books.

MacLean, P.D. 1980. “Sensory and Perceptive Factors in the Emotional Functions in the Triune Brain,” in Explaining Emotions, A.O. Rorty, (Ed.) 1980, Los Angeles: University of California Press.

MacLean, P.D. 1990. The Triune Brain in Evolution: Role in Paleocerebral Functions. New York: Plenum.

Maynard Smith, J. 1964. “Group Selection and Kin Selection,” *Nature* 201:1145-47.

Maynard Smith, J. 1976. “Group Selection,” *Quarterly Review of Biology* 51:277-83.

Maynard Smith, J. 1982. Evolution and the Theory of Games. Cambridge: Cambridge University Press.

Meskin, A. and Jonathon Weinberg. (Unpublished). *Emotions, Fiction, and Cognitive Architecture*.

Miller, G. 2001. The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature. London: Vintage.

Millikan, R. 1984. Language, Thought, and Other Biological Categories. Cambridge, MA: MIT Press.

Neill, A. 1995. “Fiction and the Emotions,” in Arguing About Art: Contemporary Philosophical Debates, A. Neill, and Aaron Ridley, Eds. 1995, New York: McGraw-Hill.

Neill, A., and Aaron Ridley, Eds. 1995. Arguing About Art: Contemporary Philosophical Debates. New York: McGraw-Hill.

Nichols, Shaun and Stich, S. 2000. “A Cognitive Theory of Pretense,” *Cognition*, 74:115-147.

Nisbett, R.E., and T.D. Wilson. 1977. “Telling More than we can Know: Verbal Reports on Mental Processes,” *Psychological Review*, 84:231-59.

Nussbaum, M. 1986. The Fragility of Goodness: Luck and Tragedy in Greek Tragedy and Philosophy. Cambridge: Cambridge University Press.

Nussbaum, M. 1990. Love’s Knowledge: Essays on Philosophy and Literature. New York: Oxford University Press.

Orians, G.H. and J.H. Heerwagen. 1992. “Evolved Responses to Landscapes,” in The Adapted Mind: Evolutionary Psychology and the Generation of Culture, J. Barkow, L. Cosmides, and J. Tooby, Eds. 1992, Oxford: Oxford University Press:555-579.

Parrot, W. and J. Schulkin. 1993a. “Neuropsychology and the Cognitive Nature of the Emotions,” *Cognition and Emotion*, 7(1):43-60.

Parrot, W. and J. Schulkin. 1993b. “What Sort of System Could an Affective System Be? A Reply to Le Doux,” *Cognition and Emotion*, 7(1):65-70.

Pinker, S. 1994. The Language Instinct. New York: Harper Collins.

Pinker, S. 1997. How the Mind Works. London: Penguin Books.

Plotkin, H. 1997. Evolution in Mind; An Introduction to Evolutionary Psychology. London: Penguin Books.

Poole, Steven 2000. Trigger Happy: The Inner Life of Videogames. London: Fourth Estate.

Provine, R. 2000. Laughter: A Scientific Study. London: Faber and Faber.

Radford, C. 1975. “How Can We Be Moved by the Fate of Anna Karenina?” *Proceedings of the Aristotelian Society*, Supp. Vol.49:67-80.

Robinson, J. 1995a. “Startle” *Journal of Philosophy, 92,* 53-74.

Robinson, J. 1995b. “L’Education Sentimentale” Australasian Journal of Philosophy 73: 212-226.

Rorty, A.O. (Ed.) 1980. Explaining Emotions. Los Angeles: University of California.

Sartre, J.P. 1939. Sketch for a Theory of the Emotions. London: Methuen.

Schacter, S. and J.E. Singer. 1962. "Cognitive, Social, and Physiological Determinants of Emotional State," *Psychological Review*, 69:379-99.

Scherer, K.R. 1993. "Neuroscience Projections to Current Debates in Emotion Psychology," *Cognition and Emotion*, 7(1):1-41.

Scruton, R. 1974. Art and Imagination. London: Methuen and Co.

Searle J.R. 1969. Speech Acts: An Essay in the Philosophy of Language. Cambridge: Cambridge University Press.

Searle J.R. 1974. "The Logical Status of Fictional Discourse," *New Literary History*, 6:319-332.

Solomon, R.C. 1986. The Passions. New York: Doubleday.

Solomon, R.C. 2000. "Nature of Emotions," in The Concise Routledge Encyclopaedia of Philosophy. London: Routledge.

Sterelny, K. 2001. The Evolution of Agency and Other Essays. Cambridge: Cambridge University Press.

Sterelny, K. and P.E. Griffiths. 1999. Sex and Death: An Introduction to Philosophy of Biology. Chicago: The University of Chicago Press.

Stich, S. 1983. From Folk Psychology to Cognitive Science. Cambridge, MA: MIT Press/Bradford Books.

Stich, S. 1996. Deconstructing the Mind. New York: Oxford University Press.

Stich, S. and Shaun Nichols. 1997. "Cognitive Penetrability, Rationality and Restricted Simulation," *Mind & Language*, 12:297-326

Stolnitz, J. 1969. "The Aesthetic Attitude," in Introductory Readings in Aesthetics, John Hospers, (Ed.) Macmillan.

Tooby, J. and L. Cosmides. 1992. “The Psychological Foundations of Culture,” in The Adapted Mind: Evolutionary Psychology and the Generation of Culture, J.H. Barkow, L. Cosmides, and J. Tooby, Eds. 1992, Oxford: Oxford University Press:19-136.

Turner, M. 1996. The Literary Mind: The Origins of Thought and Language. Oxford: Oxford University Press.

Walton, K. 1976. “Categories of Art,” *Philosophical Review*, 77:334-67.

Walton, K. 1978. “Fearing Fictions,” *The Journal of Philosophy*, 75:5-27.

Walton, K. 1990. Mimesis as Make-Believe. Cambridge MA: Harvard University Press.

Walton, K. 1997. “Spelunking, Simulation, and Slime; On Being Moved by Fiction,” in Emotion and the Arts, Hort, M. and S. Laver, Eds. 1997 Oxford: Oxford University Press.

Warnock, G.J. 1980. J.L. Austin. London: Routledge.

Williams, C. 1998. “Is Tragedy Paradoxical?” *British Journal of Aesthetics*, 38:47-62.

Williams, G.C. 1966. Adaptation and Natural Selection. Princeton: Princeton University Press.

Wilson, E.O. 1975. Sociobiology: The New Synthesis. Cambridge, MA: Harvard University Press.

Wimsatt, Jr., W. K. and Monroe C. Beardsley. 1978. “The Intentional Fallacy,” in Philosophy Looks at the Arts: Contemporary Readings in Aesthetics, Joseph Margolis, (Ed.) Philadelphia: Temple University Press.

Yanal, R.J. 1994. “The Paradox of Emotion and Fiction,” *Pacific Philosophical Quarterly*, 75:54-75.

Yanal, R.J. 1999. Paradoxes of Emotion and Fiction. Pennsylvania: Penn State Press.

Zajonce, R.B. 1980. “Feeling and Thinking: Preferences need no Inferences,” *American Psychologist*, 35:151-75.