EMOTIONAL RESPONSES TO INTERACTIVE FICTIONS

Andrzej Hagger

A Thesis Submitted for the Degree of MPhil
at the
University of St Andrews

2012

Full metadata for this item is available in Research@StAndrews:FullText
at:
http://research-repository.st-andrews.ac.uk/

Please use this identifier to cite or link to this item:
http://hdl.handle.net/10023/3595

This item is protected by original copyright
Emotional Responses to Interactive Fictions

Candidate: Andrzej Hagger
Degree: MPhil in Philosophy
Date of Submission: 31.05.2012
1. Candidate’s declarations:

I, Andrzej Hagger, hereby certify that this thesis, which is approximately 40,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September, 2010 and as a candidate for the degree of MPhil (Philosophy) in September, 2010; the higher study for which this is a record was carried out in the University of St Andrews between 2010 and 2012.

Date 22/09/2012 signature of candidate ...................

2. Supervisor’s declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of MPhil (Philosophy) in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 22/09/2012 signature of supervisor ………............

3. Permission for electronic publication:

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

Access to printed copy and electronic publication of thesis through the University of St Andrews.

Date 22/09/2012 signature of candidate .................... signature of supervisor ......................
Abstract

We commonly feel a variety of emotional responses to works of fiction. In this thesis I propose to examine what we understand by the terms fictional and narrative, and to describe what sorts of narrator might be required within a narrative work.

Of particular interest are interactive works of art, both narrative and non-narrative, and I provide a definition of what features a work should possess if it should properly be considered interactive. I discuss the notions of interactive narratives and examine how interactivity affects any possible narrator.

I examine the paradox of fiction - how it is that we can feel emotions towards characters we know not to exist, and suggest how the paradox can be dissolved. I further discuss how it can be rational to feel these emotional responses and note particular responses that it does not seem possible to feel rationally when engaging with non-interactive narratives.

I then examine what effect the introduction of interactivity to both non-narrative and narrative works has, and argue that it reduces the control the artist has to direct our emotions, but increases the range of emotions which we can feel. Finally I suggest that some of the emotional responses that would be irrational to feel when engaging with non-interactive narrative works can be rational when we are engaged with their interactive counterparts, but that at least one emotional response cannot genuinely be felt rationally even in interactive cases.
# Table of Contents

Introduction ........................................................................................................... 1  

Chapter 1  
  1.1 ............................................................................................................ 2  
  1.2 ............................................................................................................ 21  

Chapter 2  
  2.1 ............................................................................................................ 35  
  2.2 ............................................................................................................ 57  

Chapter 3  
  3.1 ............................................................................................................ 67  
  3.2 ............................................................................................................ 69  
  3.3 ............................................................................................................ 83  
  3.4 ............................................................................................................ 96  

Chapter 4  
  4.1 ............................................................................................................ 98  
  4.2 ............................................................................................................ 109  
  4.3 ............................................................................................................ 117  

Conclusion ........................................................................................................... 128  

Bibliography ......................................................................................................... 129
Introduction

We commonly experience emotional reactions to works of fiction, but there is much discussion over whether such responses are genuine emotions and, if they are, whether it can be rational to feel them. In this thesis I propose to examine what motivates these concerns and how they apply to interactive fictions.

In chapter one I will examine our understanding of what works should be considered fictions, how we can define and describe the quality of a narrative, and consider the role of the narrator within fictional works. In chapter two I will describe how we should understand and apply the term ‘interactive’ to works of art, and what would be required from interactive narratives and narrators. In chapter three I will lay out the problem of the paradox of fiction - whether we can feel emotions for fictional characters and, if we can, whether it is rational to do so, and argue that it is both possible and rational for most, but not all, emotional responses. Even when certain emotional responses can be shown to be irrational, there are good grounds for believing that it can at least be instrumentally rational to experience them. In the final chapter I will show how interactive works, both narrative and non-narrative, can elicit stronger emotional responses than non-narrative works, and explain why the artist loses some measure of her ability to control or direct just which emotions we feel. I will highlight the importance of the sense of identity players feel with their characters. I will also argue that the range of possible and rational emotional responses is greater for interactive works than it is for non-interactive works, but that some responses remain irrational.
Chapter One - Fiction, Narrative & Narration

In this chapter I will highlight some of the difficulties encountered when trying to find a single, all-encompassing definition of what it is for something to be a fiction, and I will present a framework which better captures the differing senses with which we apply the term. I will then go on to discuss what we mean when we talk of narratives and how we distinguish between narrative and non-narrative works. I will argue in favour of a broad and permissive definition of narrative and refine this through the notion of ‘narrativity’. I will then ask whether a narrative requires there to be a narrator within the fictional world of the work. I shall argue that there is no such requirement, that often it appears to be impossible or undesirable for such a narrator to exist, and that the argument that it is instrumentally useful to presuppose their existence is flawed.

1.1 Fiction

Our first step should be to consider just what a fiction is - what features distinguish fictions from non-fictions. At first glance it might be thought simply that non-fictions concern the real world - actual events and actual individuals - and fictions concern imaginary worlds - events and agents that are not real. This, however, is certainly insufficient: on the one hand there are scientific papers that purported to describe the real world but which later discoveries showed to be erroneous - discussions of a geocentric solar system, for example; whilst it could be argued that such works should
be counted as non-fictions because they sought to describe the real world and simply failed to do so, we might reply that, since they do not actually describe the real world, the only world we can fairly say that they do describe does not exist (i.e. they describe an imaginary, or fictional, world) and as such they cannot correctly be considered non-fictional - already we have run into difficulties. On the other hand there are ‘historical novels’ that give fictional accounts of real events featuring real individuals - Tolstoy’s War and Peace, for example, includes Napoleon and Alexander I and is set in Russia during the Napoleonic wars. Attempting to determine a suitable proportion of fact to fiction or fiction to fact is equally problematic, as it fails to account for grossly erroneous assertions of fact, about phlogiston, for example, and entirely or significantly, but accidentally, true works of fiction.

We might appeal to the intentions of the author, and claim that a work is fictional if the author intended us merely to imagine its content is true and non-fictional if she intended us to believe that its content is true, in which case papers on the properties of phlogiston are non-fictional but false, and accidentally accurate novels are fictional, but true. However we may have concerns about removing the truth or falsity of a work from consideration of its fictional status, and where the author’s intentions are unclear or cannot be known we appear to be unable to categorise her work either way. It therefore follows that what may be needed in determining whether a work should be considered to be fictional or non-fictional is an account, not of its properties or the intentions of its author, but of its function: is the work a suitable basis for the audience to imagine fictional worlds? Whether or not a work can be used as an imaginative tool tells us
nothing about whether it generally is used as such, so we might also wish to consider how works are used, or the way with which they are engaged, by their audiences; although the suggestion that a work is fictional or non-fictional iff its audience treats it as such would be curious¹, we should not overlook the role of the audience’s beliefs about a work. There are then (at least) three distinct senses in which works might be considered fictional or non-fictional: (1) whether they are accurate descriptions of the real world, (2) whether the author intended his audience to imagine the content of the work is true, or to believe it, and (3) how audiences in fact engage with the work. After outlining two possible positions that might be taken and some typical problems that they encounter, I will suggest a framework that captures these senses and provides scope for degrees of fiction within them and in doing so allows for a more holistic consideration of the nature of fictionality, and I will test this against the problem cases.

1.1.1 Authorial Intentions

The first position to examine concerns the intentions of the author of the fiction - the fictionality of a work depends upon how the author intended the work to function. In *The Nature of Fiction* Currie describes the author of a fictional work as intending the audience “to make believe that the story is true”², and notes that there are two key parts to the author’s intentions:

¹ It would follow from such a theory that audiences could not misunderstand a work; there can surely be no value in a theory that labels demonstrably accurate documentary accounts fictional and far-fetched tales of fantasy fact simply because their audiences believe them to be so.
The author intends that we make-believe the text … and he intends to get us to do this by means of our recognition of that very intention.³

In the case of a scientific paper about the properties of phlogiston, even though the substance is no more real than Sherlock Holmes, the author of the paper certainly did not intend to write a fiction; for Currie the work would only be fictional if the author had intended it to be. This view raises a number of questions: what if the author failed to achieve his intentions - he may have intended to write a fictional work but if it was received as non-fiction, or it accidentally described true facts about the real world, should we still call it fictional? If we had no information about an author’s intentions, how would we classify a work if it were discovered after the author’s death, and could be interpreted equally plausibly as fictional or non-fictional? What if the the author’s intention was to deceive his audience? In “What is Fiction?”⁴, Currie imagines that Defoe intended that Robinson Crusoe be received as a non-fictional account; would it therefore fail to qualify as fiction, and would it consequently have to be classified as merely a fraudulent non-fictional work? Currie’s solution is to suggest that there may be a category of “pseudofictions” into which Robinson Crusoe would fall as there is “a widespread practice of reading the work as if it were fiction”⁵. This is in line with (3) above. Currie defends this position by noting another of Defoe’s works, A True Relation of the Apparition of one Mrs. Veal, which was until relatively recently regarded as fiction:

⁴ Currie, “What is Fiction?”, p. 36.
⁵ Currie, “What is Fiction?”, p. 37.
Anyone who says, reasonably enough, “It was widely and mistakenly thought to be fiction,” must be making a distinction between being fiction and being regarded as fiction. So my answer to Crusoe-type counterexamples is this: Don’t confuse the (historically revisable) claim that these works are fiction with the (less revisable) claim that they are pseudofictions.

A strict application of Currie’s criteria (cf. (2) above) counts *Robinson Crusoe* as non-fiction, and his solution springs from his belief that “[t]o say that [it] is not fiction is not to say that readers ought to change their attitudes toward it” we can say that they ought not to do so because by (1) the work is fiction.

Regarding the inclusion of factual truths in a work, Currie is very clear that for a work to be fictional “if [it] is true, then it is at most accidentally true.” Currie is sensitive to the fact that works may include non-accidental truths alongside accidental ones, allowing that historical novels which are largely based on documented evidence but also include imagined scenarios can be considered fictions. The inclusion of a single imagined episode (even if accidentally true) in an otherwise documentary work would, for Currie, be sufficient to describe the work as a whole as fictional. This is problematic, as many works we would usually consider to be non-fictional posit imagined counter-factual

---

situations or ask their readers to engage in thought-experiments - Currie is at risk of counting the overwhelming majority of philosophical writings as fictional. Presumably he would fall back on his proposed solution for *A True Relation of the Apparition of one Mrs. Veal*: philosophical texts are generally read as if they are non-fictional works. Currie is forced to add yet another category to his classificatory system: alongside fictions and non-fictions we have pseudo-fictions, and now “pseudo-non-fictions”. I suggest that Currie’s appeal to our “widespread practices” is evidence for my claim that the term ‘fiction’ *simpliciter* is insufficient to capture just what we mean in a given context, and that more fine-grained definitions are required.

### 1.1.2 Function

The second position to examine concerns the way in which the fiction functions for its audience - the fictionality of a work depends upon the use its audience makes of it. The term ‘audience’ opens this position to its most obvious objection: what constitutes a work’s audience - a majority, or a certain proportion of the individuals who have engaged with it, or the consensus of ‘expert appreciators’? Alternatively, should each individual be considered a separate audience? If so, this opens up the possibility that a work is fictional to some audiences and non-fictional to others. Further, however constituted, how should the audience or audiences determine whether or not a work is fictional? As already noted, appeals to features of a work may too easily be insufficient. *Robinson Crusoe* again presents something of a problem to this account, as it would seem to require that it was initially a work of non-fiction until it was discovered not to be
a genuine account, at which point it became a work of fiction. This would also reinforce both the ‘which audience’ question and the dual-classification problem - either the change in status occurred when its inauthenticity was discovered by certain expert appreciators, or by a certain proportion of its audience, or it was simultaneously non-fictional for those unaware of Defoe’s deception and fictional for those who were not.

In *Mimesis as Make-Believe* Walton describes a child, Heather, who is pretending that tree stumps are bears\(^\text{10}\) . For Walton “to be fictional is, at bottom, to possess the function of serving as a prop in games of make-believe” \(^\text{11}\); in his example Heather is playing a game of make-believe in which she uses the tree stump as a prop, so (at least while she is playing the game) the tree stump itself is a fiction. In the same way, a novel is a prop that a reader uses to imagine the exploits of a remarkable ‘consulting detective’ or the tragic affair between two Russian aristocrats. Walton distinguishes between tree stumps and novels by suggesting that the former are “natural” - they are “ad-hoc props, pressed into service for a single game of make-believe on a single occasion” and the latter are “designed to be props” - they have been created with that function in mind\(^\text{12}\) . Although the idea that the way a designed prop becomes a fiction can *relate* to the intentions of its creator (cf. (2)), it is important to note that for Walton “it does not depend on them”\(^\text{13}\).

\(^{10}\) Walton, *Mimesis as Make-Believe*, p.21-2.
\(^{11}\) Walton, *Mimesis as Make-Believe*, p.102.
\(^{13}\) Walton, *Mimesis as Make-Believe*, p.88.
Walton suggests that “[f]ictions are cultural constructs” 14, i.e. that certain norms may govern the make-believe in which Heather indulges such that she imagines the tree stumps are bears but, being governed by different cultural norms, Jack might make-believe that they are demons. With regard to designed props, although our intuitions are that the creator’s intention is that they be used as a prop, or used as a prop in a particular way, Walton’s position seems to be that our cultural norms give rise to these intuitions and individuals with radically different norms need not heed the intentions of the prop’s creator.

Walton also discusses the notion of authorless works, and gives an example of cracks that appear in rocks on a beachfront that closely resemble words and intelligible sentences15. If someone were to read these sentences and use them as a prop for making-believe, it would appear that they should, by this account, be properly considered to be a fiction - quite a counter-intuitive result. We can extend this problem by imagining that, unbeknownst to the reader, the sentences in fact accurately describe a real-world scenario; the reader makes-believe in the same way he or she would were the prop an authored novel - the ‘work’ must surely remain fictional (this broadly parallels the problem of accidental accuracy in the intentionalist account).

It can also be questioned whether this account is more a way of describing whether a work is treated as fictional, rather than whether it is fictional. As we have seen, this

---

14 Walton, Mimesis as Make-Believe, p.88.

15 Walton, Mimesis as Make-Believe, p.87.
question echoes the Crusoe-type problem for Currie, insofar as such problems force him to fall back on widespread practices in place of his initial criterion, whether the author intends her audience to imagine or believe the content of the work. This is the heart of the issue, and thorough consideration of the implications of this objection may lead to a way to better understand the problem at hand; it would appear that trying to determine the status of a given work as fictional or non-fictional simpliciter blurs three distinct senses in which the concepts can be applied. However, although they are similar to our initial (1), (2) and (3) the difficulties we have seen with how Currie handles non-fictional works with imaginary episodes and Walton’s reliance on cultural norms show that they require some modification: (1A) how accurately the work represents the real world (not simply whether it is completely accurate); (2A) whether the author has attempted to describe the world accurately, to deceive the audience, or has invited the audience to make-believe in order to achieve some given end; (3A) how an audience has reacted to or treated a work at a given time.

At its core it may be said that my account will rest broadly on an intentionalist view of fiction, as Walton’s assertion that intentional acts are subordinate to the objects of appreciation (or the props of make-believe) strays too far from the general usage and understanding of ‘fiction’. Whilst cracks in rocks may be suitable for use as props for making-believe, it seems to me that they are not themselves works of fiction; without a possible appreciator it is difficult to see how the term ‘fiction’ has any useful application, as the notion of props seems to presuppose an agent able to make use of them. Here “possible appreciator” is a key concept - Walton’s example relies on the cracks in the
rocks resembling something that is in some way intelligible. If the cracks appear to form words, they are only suitable as props if the appreciator is competent in the language in which the words appear to be written. It could be argued that this is not a necessary condition, as another (non-present) individual competent in that language would be able to use them as a prop, however it easy to imagine possible worlds in which any formation of cracks in rocks would closely resemble an intelligible language to another agent. This would suggest that all cracks in all rocks, all shapes in the bark of all trees, all layouts of road or power-line networks are (or at least could be) fictions; the set of fictions (or possible fictions) thus extended includes not only all objects individually, but all configurations of groups of objects and makes 'fiction' an empty term. The notion of purpose does not help here: if the purpose of the layout of power lines is not to provide a prop, the same surely applies to cracks in rocks; furthermore, this notion of function arguably seems to have intentional acts built in to it, as if an object has the function of providing a prop for make-believe, an agent (in this case the audience, rather than the author) capable of the intentional act of making-believe is again presupposed. Walton’s assertion that cracks in rocks can be fictions denies the necessity of an author. Given the work the appreciator does in these cases, however, we should conclude that the fictions are authored by the audience, and it is less implausible to suggest that the works, such as they are, are instantiated in their imaginations than in the props they used in authoring them. Also, if we imagine cracks in rocks which happen to resemble sentences that accurately describe real-world scenarios, it is even harder to consider
them non-fictional in anything but the sense of their apparent truth-values, and the suggestion that the rock itself is a non-fictional work seems deeply counter-intuitive.  

1.1.3 Three senses of fictionality

How then should we understand ‘fictional’ and ‘non-fictional’? In order to capture the various senses of fictional previously described, I suggest the following: works can be considered epistemically non-fictional to the extent to which they accurately represent the real world and fictional to the extent to which they misrepresent it. Their status as authorially fictional or non-fictional depends upon the intentional state of the author while he or she is writing it, as opposed to the author’s intentions regarding how the work should be received, and they can be described as receptively fictional or non-fictional according to whether their audiences believe them to be epistemically fictional or non-fictional. I will examine each of these categories in turn.

Epistemic fictionality

Epistemic fictionality can be determined by examining the truth-values of the claims within the work. Applying this first to uncontroversially non-fictional works such as accurate scientific papers there are no surprises, nor in application to uncontroversially fictional works such as Tolkien’s *The Lord of the Rings*, and it should be equally acceptable to describe accidentally accurate works of artistic invention as epistemically

---

16 Walton does not claim that such rocks should be considered a non-fictional work, for him they are “no more than a curiosity” (*Mimesis as Make-Believe*, p. 87).
non-fictional. Where works combine uncontroversially true statements and uncontroversially false statements, such as Joyce’s *Portrait of the Artist as a Young Man*, the work as a whole might be described as epistemically fictional but it would be more informative to describe it as a mixed work, i.e. that it contains some statements that are epistemically fictional and some that are epistemically non-fictional, and this too seems to be acceptable. Turning to less intuitively obvious cases, it seems fair to call erroneous scientific papers epistemically fictional. We might feel more inclined to describe such works as non-fictional, but false, but this is a return to using a single all-encompassing notion of fictional/ non-fictional which leads us to the unintuitive position of describing a paper discussing phologiston as a non-fictional account of a fictional object. By separating epistemic fictionality from authorial fictionality we will see that such a paper is epistemically fictional and authorially non-fictional. For papers with less significant errors it would certainly seem odd simply to call them fictional in the broader, non-specific sense with which the term is generally applied, but I suggest merely that those errors can reasonably be described as epistemically fictional; as with the Joyce example, the papers-as-wholes could also be described as epistemically fictional, but it would be better to describe them as a mixed work - many of the statements might be epistemically non-fictional, with only a few (admittedly key) claims that are epistemically fictional. To take a slightly different example, consider a humorous poster highlighting the supposed dangers of “dihydrogen monoxide” (i.e. water) and calling for legislation to impose controls on its use. As a work we would consider it straightforwardly fictional even though, aside from the fictional claim that a pressure group is lobbying for its removal, all the remaining claims about dihydrogen
monoxide are factual; in terms of epistemic fictionality this too is a mixed work. If an individual with the relevant cultural norms to interpret cracks in a rock as intelligible language, and in so doing instantiates a work in her imagination, I suggest that there is no reason we cannot attribute epistemic fictionality to that work if the work she has authored is not an accurate depiction of the world; furthermore, if a second individual with cultural norms such that he interprets the cracks differently, and instantiates a different work that does provide an accurate description of the world, there is similarly no problem in describing his work as epistemically non-fictional.

**Authorial fictionality**

A work should be considered authorially fictional if, while writing it, the author’s intentions were to create a work that was not an accurate representation of the real world. All (seriously written) scientific papers are therefore authorially non-fictional. This is an important departure from the notion of works being fictional in virtue of inviting the audience to make-believe, as it accurately captures the status of scientific works concerning, for example, quantum physics; in such papers the theories posited are not directly empirically testable - the reader is invited to make-believe that the described scenario is the case, and examine how well it measures up to or conforms with facts about the world that are observable or determinable\(^\text{17}\). Equally this accurately captures works such as *Robinson Crusoe* in which the author is attempting to deceive the

\(^{17}\text{Although such papers satisfy Currie’s requirement that the author intends the audience to engage in make believe, the author is attempting to describe truths about the real world so if the paper is accurate it is almost certainly not accidentally so, and would therefore not be fictional according to Currie.}\)
audience into accepting an epistemically fictional work as non-fictional. Style is also accounted for - an author may write an epistemically non-fictional work in a narrative style that may unintentionally be mistaken by an audience for fiction; it is nevertheless authorially non-fictional. It also provides an intuitively acceptable counterpart to the epistemically non-fictional status of accidentally accurate works by describing them as authorially fictional. It might be objected that works of realist literature are attempts to describe accurately the real world, but are nonetheless fictions. There are two key points to note however: firstly, although such works attempt to provide faithful representations in general, George Eliot’s Middlemarch, a work widely accepted as being within the canon of realist literature, is set in a fictitious town and as such it cannot be claimed that Eliot intended to provide an accurate portrayal of Middlemarch or of the lives of the majority of the characters within the novel as they did not, in fact exist; secondly, we have no reason to believe that Eliot intended Middlemarch to be received as a non-fictional account of real places and people. We can therefore be confident in asserting that the work as a whole is authorially fictional. What this example demonstrates is that authorial fictionality provides scope for granularity of the status of a work - a work as a whole may be authorially fictional while at the same time containing elements that are authorially non-fictional, such as Middlemarch and War and Peace. The extent to which the authors intended their works to portray accurately real people and places determines their levels of authorial fictionality.

Receptive fictionality
This is straightforwardly understandable, but adds a final dimension necessary for accurately describing works which have been assigned an incorrect epistemic fictionality status by their audience. Not only does it also allow for granularity of status assignment, it comfortably allows for works to change their status, and for different audiences to assign different statuses to works, or parts of works.

The previous attempts at assigning a single fictional/ non-fictional status to the Crusoe problem ran into difficulties; this framework allows us to describe it as epistemically fictional, authorially fictional and, whilst for a time it was receptively non-fictional it is now receptively fictional. Returning the same status for all three tests would perhaps allow us to describe it simply as fictional, in accord with our intuitions.

Summary table

<table>
<thead>
<tr>
<th>Epistemically Fictional Works</th>
<th>Authorially Fictional</th>
<th>Authorially Non-Fictional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receptively Fictional</td>
<td>Robinson Crusoe (Latterly)</td>
<td>Accounts of phlogiston (Latterly)</td>
</tr>
<tr>
<td>Receptively Non-Fictional</td>
<td>Robinson Crusoe (Initially)</td>
<td>Accounts of phlogiston (Initially)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Epistemically Non-Fictional Works</th>
<th>Authorially Fictional</th>
<th>Authorially Non-Fictional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receptively Fictional</td>
<td>Accidentally accurate fictions received as fictions</td>
<td>Accurate but disbelieved/ misunderstood scientific papers</td>
</tr>
</tbody>
</table>
Potential objections

One objection that may be raised in reference to the notion of epistemic fictionality is how it deals with works which contain both factually correct statements and factually incorrect statements. It is clear that a work entirely devoid of false statements should be described as epistemically non-fictional, and one entirely devoid of true statements (eg a fantasy or science fiction novel) should be described as epistemically fictional. Whilst individual statements can be assigned truth-values relatively straightforwardly, consideration of the entire work is perhaps a little less clear. Although it might be argued that there is some merit in allowing for a small proportion of false claims in an otherwise epistemically non-fictional work in order not to describe almost-perfect scientific papers and the like as fictional, this approach quickly runs into difficulties: what proportion can be deemed acceptable? A work accurately describing real historical events, but making unsubstantiated claims as to the motivations of the individuals should probably not be described as non-fictional (here I assume the claims are presented as definitive facts, rather than “may have wanted” etc.). If pushed to make a definitive pronouncement on the epistemic fictionality of a work as a whole (i.e. if ‘mixed work’ is, for some purpose, an unacceptable determination), we might elect to base our decision either on the number of correct and incorrect claims, but this would make the dihydrogen oxide poster non-fictional. It would be better, then, to decide based on the
correctness of the most important or significant claims - for the poster this would be the
existence of the pressure group and the ‘need’ for legislation. Clearly, just how to
determine the importance of a claim will vary from case to case - a scientific paper that
reaches an accurate conclusion through an erroneous methodology should still be
considered epistemically fictional as, although the conclusion may be important, for
scientific works it is the rigour and accuracy of the work done to reach that conclusion
that is paramount. The difficulties deciding the importance of individual claims within a
work are less a shortcoming of the notion of epistemic fictionality than a spotlight on the
undesirability of avoiding the ‘mixed-work’ conclusion. The most logical solution, if we
simply must judge the work as a whole, is to hold that even a single false claim is
sufficient to call the work epistemically fictional. Although this labels almost-perfect
scientific papers as fictional, it is important to recognise that this is simply a description
of the truth-value of the paper as a whole, not a claim about the intention of the author.
So understood, this criterion should be uncontroversial. It does, however, highlight
another potential criticism - that the notion of epistemic fictionality is no more than a re-
labelling of a work’s truth-value. This point stands, however the category captures an
important sense of the term ‘fiction’ as generally applied, and is a necessary component
to give a complete description of a work’s status.

Objections may be raised regarding authorial fictionality in terms of unknown intentions -
the term seeks to capture information about the intentional state of the author, which
may not be known. Whilst this may be a fair criticism, authorial fictionality does not
suffer any more from this objection than other intentional accounts of fiction. We can
appeal to other known facts about the author and the context of the work to help decide, for example whether the work was published in a scientific journal, whether there is independent documentation that corroborates the claims made, whether the author accepted prizes or awards for fiction etc. Arguably we cannot know that certain ancient Greek plays were not intended to provide documentary accounts of real-world people and events, but we are generally confident in believing this not to be the case. Beyond this, it should be no more difficult, indeed perhaps easier, to determine whether the author intended to provide a factually correct account of the world than whether she intended an audience to make-believe the scenarios in the work; we can be confident that Defoe did not intend to describe the real world, but we might debate whether he intended his audience to make-believe or to believe the story. Returning to the example of the quantum physicist, she certainly invites her readers to make-believe that her description is the case, but she also intends to describe the real world; my account of authorial fictionality better captures this case.

It may be suggested that my definition of receptive fictionality fails to capture how the audience believes the artist intended the work to be received in favour of the truth-values they assign to it, and indeed this is true. Authorial fictionality is intended to describe a truth about the author’s real-world intentions, rather than what an audience believes about his intentions. As described above, there are cases where the truth or falsity of that claim may be impossible to determine and here we are left with no option but to make an educated guess, and it seems acceptable for me to admit of the
possibility that an audience, however constituted, might assign a value to a work’s authorial fictionality that is determinably mistaken.

If consideration of a work returns the same value for all three senses of fictionality, there should be no controversy in describing it simply as fictional or non-fictional. If a mix of values are returned this is more problematic and it might be argued that this framework fails to provide a straightforward label in these cases. However, this objection only applies if we are trying to give a single label to a term we have already seen has three distinct senses, and that is the very thing which this framework attempts to avoid.

1.1.4 Conclusions

I hope to have shown that our notion of just what it is for something to be fictional is unclear, as the term can be applied in a number of ways. Reliance on author’s intentions is insufficient, as he may fail to achieve his intentions, or the intentions may not be known.

 Appeals to widespread practices of audiences of a work are also insufficient, as audiences may simply be wrong in how they judge the work, and different audiences may, through variations in cultural norms or available evidence, treat the same text in different ways. I have provided a tripartite description of what we mean when we talk about a work’s status as fictional or non-fictional, and argued that the result of ‘mixed work’ in any of these categories is both desirable and useful, and in keeping with our common practices.
If it is necessary or desirable for a specific purpose to judge a work as a whole this framework provides suitable criteria with which to do so. Objections to the results of applying these criteria dissolve when their distinct senses are properly understood and separated out. Where the intentional state of the author is not known we can base our determinations on other information, and these problems are no greater for this account than for other intentional accounts of fiction. The key benefit of this account is that it highlights and deals individually and independently with the three senses of fiction that are usually conflated.

1.2 Narrative and Narration

Having discussed what it is for something to be fictional the next area we should examine is that of narrative. It should be clear that fiction and narrative describe different, if often related things – whilst perhaps we mostly commonly associate narratives with fictions, neither are all narratives fictional nor all fictions narrative. In addition we must note that although narrative is perhaps most often associated with literary works it is in no way restricted to literature – films, dance, music and pictures can all include narrative content. After providing a rough definition for each term I will examine Currie’s notion of ‘narrativity’ and whether it can be usefully applied. I will also look in detail at whether some form of narrator is presupposed by a narrative. We begin, then, by attempting to describe just what narrative and narration are.
1.2.1 Outline Definitions

The simplest descriptions of narrative and narration, albeit question-begging in terms of definitions, are that narration is the relating of the details of an occurrence and a narrative is a story. For ‘details of an occurrence’ to be related, we might conclude that someone must be doing the relating; that narratives require an author or authors is uncontroversial (for Waltonian ‘cracks in rocks’ it is the make-believer who authors the narrative through her imaginings) but does a narration require a narrator? We might be tempted to think so, but as we shall see there are persuasive arguments against this view. If narratives are stories then we can say that they differ from shopping lists in that they describe situations and events, they differ from scientific papers in that they describe specific individual events rather than attempting to provide rules with which to describe general states of affairs and, by requiring a narrative to describe some sort of change in the situation, we can differentiate a narrative from a mere description of a scene. However, if ‘a description of a change in a situation’ were all that were required for something to be considered a narrative even the single sentence “the cat sat on the mat” could qualify (given the assumption that “sat” was describing an action that the cat was taking, i.e. the sense of “sat down”, rather than “was sitting”). Whilst this might be acceptable, perhaps we would not wish for such a simple statement to be considered a narrative, especially in such cases where the sense of verb is ambiguous when devoid of further context. One possibility might be to require more than one event and, in order to avoid similar problems of ambiguity with statements such as “the cat sat on the mat and the man looked at the picture”, for those events to stand in some sort of relation to one another. Livingston lists a number of options for what sort of relation we might require:

i. a temporal ordering

ii. a causal sequence
iii. a more general 'x occasions y' relation

iv. at least one event manifesting some agent's problem-solving activity, that is, some (intentional) actional or goal-directed activity

v. two or more actions involving purposeful activity, with the additional requirement of a form of 'closure', meaning that there must be events that function as a coherent sequence of purposeful activity, with a beginning and resolution (provided, for example, when some agent realizes or abandons some overarching goal)\(^\text{18}\)

Livingston also suggests that some sort of explicit or implicit narrator may also be required for something to be a narrative and notes that (iv) or (v) could be expanded to require a 'point' or “some other expressive or communicative goal" that is achieved by the author through the structure of the story, and that we might additionally require that something about the agent's goal or efforts to achieve the goal must be in some way interesting to the audience.

It would appear that statements such as “the cat sat on the mat, and then the man looked at the picture”, “the cat relaxed, and sat on the mat”, “night fell, and the cat sat down on the mat”, “the cat chose the mat to sit upon” and “the cat was cold, so sat on the mat by the fireplace and curled up” would satisfy each of the initial criteria respectively; however, we still might find ourselves reluctant to call any of them narratives, and appeals to the expanded versions of (iv) or (v) seem to be fraught with difficulties – how are we to determine whether the author has succeeded in communicating the point of the tale (perhaps it is subtle, or relies on the audience being aware of certain facts about the real world not explicitly stated in the story\(^\text{19}\)), or whether


\(^{19}\) They might even be reliant upon certain beliefs about the real world that were commonplace at the time of writing but that have since been disproved and now largely unheard-of, i.e. the author's success in
the efforts the agent makes are suitably interesting? Indeed, given it is not unusual for narratives to be criticised for being uninteresting or for having a muddled or contradictory 'point', such criteria might seem to be unsuitable. I suggest, however, that the reluctance to call extremely simple sentences narratives is misplaced. It is motivated by appeals to works (I use this term here very loosely, and include single sentences as works in this context) of greater complexity and seeks, by implication, to conjoin complexity of narrative with aesthetic value. Whilst there is undoubtedly some truth in the idea that simpler works are less satisfying, it should be clear that length and complexity are by no means the only criteria on which a work is judged. What we are looking for here is simply a way to capture a highly minimal sense of narrative, and our initial attempt seems capable of doing this: a narrative is a description of specific individual situations and events, and must include in its description some sort of change in the situation. It might be objected that this is considerably over-inclusive, that “the cat sat (down) on the mat” cannot reasonably be held to be an example of narrative. Narratives, it might further be argued, must contain one or other elements from Livingston’s list. However, what this objection amounts to is merely that, were the sentence to contain such elements, or be expanded to multiple sentences in order to include a number of those elements, it would be a better narrative. In terms of a minimal definition it suffices, and I find no problem in endorsing “the cat sat (down) on the mat” as an example of an extremely minimal narrative.

1.2.2 Narrative and Narrativity

Gregory Currie attempts to move the discussion from what does or does not constitute a narrative to the notion of narrativity\(^\text{20}\). This is the suggestion that narrative is better thought of as communicating the point of the story may change over time.

as a spectrum on which texts etc. can be situated than a category into which a given text might or might not be placed; when something's status as a narrative is apparently disputed this can more usefully be understood as measuring its level of narrativity against different points on the scale – depending on the level of narrativity one has in mind a given work may or may not qualify as 'a narrative'. Shopping lists and scientific papers establishing general rules with which to describe the world would be at the lowest end of this spectrum, the previous statements about cats and mats would be a little higher on this spectrum as they display at least some hints of narrativity, and works such as War and Peace and The Brothers Karamazov would be at the high end of the scale as they have many temporally and causally related events, agents engaged in purposeful problem-solving with efforts and resolutions that are interesting to the audience and so on (Currie calls such items “exemplary narratives”). This spectrum should remind us of Livingston's increasingly demanding criteria for what constitutes a narrative - as we move through the list from less to more demanding, so we should move up Currie's spectrum of narrativity. Currie provides two additional criteria for judging a work's narrativity: the inter-relatedness of the characters in the story and the work's thematic unity. In the film Paris, Je T'Aime various short scenarios are played out against the backdrop of Paris with a large number of characters; whilst several broad themes such as love and chance encounter reoccur, and all the stories are linked in virtue of taking place in Paris, the characters are principally isolated in their own stories\(^{21}\) so the film's thematic unity is merely general and does not contribute significantly to the narrativity of the film as a whole. In Run Lola Run, however, we see the same characters three times and each time the story begins the same way, but the outcome of each iteration varies radically due to differences in the decisions taken by the characters and in seemingly insignificant events that occur each time Lola's journey unfolds.

\(^{21}\) Occasionally one story may lead into another, and so characters from one may appear briefly in the preceding or following tale. However, there is little, if any, causal connectedness between each set of characters, so the narrative of the film as a whole cannot be said to be meaningfully unified.
Here the thematic unity of the film is highly particular as the stories concern the same characters and the same situation, and we see not only how the main characters are inter-related but also how the main and minor characters are inter-related\textsuperscript{22}.

It should be noted that there are some difficulties with this account. Firstly, it appears to place considerable weight on the length of the work when determining its narrativity – longer works seem more likely (although clearly it is not necessarily the case) to include a greater number of events that could be temporally and/or causally related, more opportunities to provide the agents of the story with interesting problems to solve or difficulties to overcome, more characters to connect and so on. Currie says little about how one might determine either the level of narrativity of a single work or the comparative levels of different works, suggesting instead that the high degree of subjectivity in judgements about interestingness and success in communicating one's point, and the fact that the wider contexts in which works occur must be taken into account, make such determinations difficult to specify. We must be aware therefore of the distinction between the level of narrativity of a work and the aesthetic quality of that narrativity; a work ten times the length of \textit{War and Peace} may well define a hitherto unimagined high-point on the narrativity spectrum but, even if we allow that there are a manageable cast of characters and that the individual episodes of their lives are interesting, if the work is so long that no reader can reasonably be expected to retain much beyond a relatively small proportion of the work as a whole its aesthetic value in comparison to a shorter work might be in question. In addition, even if the episodes of the characters' lives are interesting individually, the work as a whole might well be too long to maintain the reader's interest, or the overarching story arcs

\textsuperscript{22} This inter-relatedness is particularly pronounced in that we see how differently the lives of the minor characters turn out as well as the main protagonists – for example, due to the differences in their brief interactions with Lola, one character either loses her children and steals someone else's baby or wins the lottery, and another is either injured in a car crash and later commits suicide or begins a relationship with a colleague.
might become overly complex and convoluted to be satisfying if considered in its entirety. Alternatively another, much shorter, work might explore the same themes and deliver resolutions just as (or even more) satisfying without the, albeit interesting, side stories (as we shall see in the next chapter, these issues extend to interactive media also, and perhaps to a greater extent).

There is something of an internal tension within narrativity too: on the one hand narrativity can be determined by quite objective features – changes in states, relations of one sort or another between agents – but on the other hand (relative) interestingness, context and plausibility (both in terms of agent actions and causal relations) must also play a part, and Currie does not integrate these features into his framework. It is therefore fair to say that Currie’s account is at best incomplete.

1.2.3 Narration and the Narrator

Our next task is to establish who it is that relates the details of the story to us. The author of the work has written the narrative, but is there a reason to believe that another agent is recounting the events in question? Currie’s proposal is straightforward:

for virtually all cases of narrative we are likely to come across, there is no distinction that should or can be made between authors and narrators, for there is no distinction to be made between narrative-making and narrative-telling.\(^{23}\)

\(^{23}\) Currie, p. 65.
One argument in favour of narration requiring a fictional narrator is the notion that a reader is justified in questioning from where the details of the events she is reading come, and that since questions of this ilk relate wholly to the fictional world we cannot appeal to the actual author of the text; to satisfy this relation we would require an agent within the fictional world of the work, a ‘fictional narrator’. Where the narrative is told in the first person it is reasonable to assert that the fictional narrator is the character relating the story, and there would be grounds for suggesting that the opinions expressed by an explicit fictional narrator are not necessarily those of the author, and that this narrator might well be unreliable as within the fiction he would not have epistemic access to the inner workings of other characters’ minds or their feelings or motivations. Livingston points out a clear example of where this approach might run into difficulties by asking “how a manuscript could have been written by an illiterate storyteller”\(^{24}\). If in reply we were to posit an unknown transcriber, we would be forced to allow that not only might the storyteller be unreliable, so too might the transcriber - he might have misheard the storyteller, or decided to exclude or modify what was told to him, or even to add details never mentioned by the storyteller. Stories which end without survivors or where the characters are forever stranded do not easily admit of a narrator who both has epistemic access to the incidents described and is able to relate those occurrences to us. Anna Sewell’s *Black Beauty* is another difficult case, as the story is related in first-person by a horse. Are we to imagine the horse wrote the work, or somehow psychically communicated the story to a (fictional, and potentially unreliable) human writer? In these cases we can say that character from whose perspective the story is told is the ‘internal narrator’ - the author of the work according to the fiction - but he is distinct from the ‘external narrator’ as it is not actually the case that he wrote the work in question. Currie suggests that “[i]f Doyle had represented himself as an agent with the [Sherlock Holmes] stories … and had thus told a fictional story partly about himself, he

\(^{24}\) Livingston, p. 280.
would thereby be an internal and an external author/narrator. However, this may not be quite accurate, as it is reasonable to suggest that the Doyle within the story is a fictional character, Doyle$_F$, distinct from the actual Doyle writing the manuscript, Doyle$_A$. Doyle$_A$ does not know a famous consulting detective named Holmes and has never visited 221B Baker Street; these statements can only be true of Doyle$_F$. There may be statements that are true of both Doyle$_A$ and Doyle$_F$, that they are Scottish, for example, but this is not enough to establish identity. Even in cases where it is fictionally the case that the work is narrated by a character within the story, however, there may be reasons to insist that there is another narrator ‘behind the scenes’, chief amongst which are cases where the character relates information beyond his epistemic field. Depending on the extent to which such knowledge is displayed we might be warranted in believing that an additional narrator is required, but where this knowledge is slight and serves no apparent artistic purpose it is probable that we are straying into the area Currie describes as ‘silly questions’; Currie’s reasonable request is that “we do not postulate internal narrators unless there is positive reason to do so.” Where a narrative, with or without an explicit internal narrator, includes much that no ordinary, non-omniscient character could be expected to know, it is difficult to make the case either that an implicit fictional narrator exists or that positing one serves any useful practical or aesthetic purpose. In *Mimesis and Make-Believe* Walton puts forward one possible case for doing so:

…recognizing narrators in a wide range of cases, while allowing that some may have scarcely any significance for criticism or appreciation, encourages sensitivity to the subtlest hints of narratorial intrusion and facilitates recognizing

---

25 Currie, p. 67.
26 Currie, p. 68.
infinitely fine variations in the degrees (and kinds) of prominence narrators in different works have...\(^{27}\)

Whilst there is certainly something to be said for sensitivity, there is no reason to presume that an assumption in favour of implicit fictional narrators unless there is good evidence to the contrary is a better way of achieving it than the reverse assumption - in fact, if we approach works with the presumption that there is no narrator any subtle clues that this presumption is incorrect should be all the more conspicuous. Walton concedes that there are cases where positing a narrator has no positive effect on our experience of the work, and seems to suggest that we should attempt to do so in order simply to become better-practiced at spotting potential cases where there is even the slightest possibility a narrator might be called for, an argument bordering on being circular. We can surely develop the same level of awareness and sensitivity by starting with a presumption against implicit fictional narrators and then examining the work in detail for evidence that this assumption is incorrect for this particular work. The advantage of taking this approach is that we will discard those cases where such a narrator offers “scarcely any significance for criticism or appreciation” and form judgements about other cases according to the evidence. I suggest that we will also be better placed to properly identify those occasions where what seems like evidence for a narrator is merely an authorial conceit - at worst we should be no worse off. Given that the positive outcomes are likely to be largely the same regardless of our starting assumption, if we are to choose between approaches we must consider the potential negative outcomes. Where we find the need for an all-knowing, all-seeing narrator we seem to have little choice but to imbue them with god-like powers - at best this does not alter our aesthetic appreciation of the work, at worst we might be led to interpret the work as told from a deity’s perspective and in so doing radically - and unwarrantedly - altering our

\(^{27}\) Walton, p. 367.
understanding and appreciation of it, perhaps even to the point where we could be accused of having misunderstood the work altogether. Walton seems to suggest that an approach such as Currie’s runs the risk of failing to recognise an implicit narrator when the author intended us to find one, with potentially the same resultant misunderstanding, and such unfortunate cases would arise either when we have failed to ‘take the hint’ or where the author has failed to make the hints suitably recognisable. However he does not show why this is necessarily the case, and as I have suggested there is no reason why evidence contrary to our assumptions should be any less conspicuous than evidence supporting them. It is fair to say that most books are not written with aestheticians as the specific intended audience, it is therefore also fair to say that if it takes a trained philosopher to spot the clues that an implicit internal narrator should be assumed in order to gain a proper aesthetic appreciation of the work the author has perhaps been too subtle for her audience. Since the presumption against implicit internal narrators dispenses with the need for deity-like powers and places the reader in no worse, and arguably a better, position to distinguish between genuine cases and those of authorial conceit we must conclude that Currie has the better approach.

The problems discussed earlier become even more significant when we consider films and television programmes - should we posit some sort of camera operator within the fiction? As many, including Currie and Gaut have pointed out, if so he would presumably have to be invisible to the other characters, able to float in mid-air, switch between two locations instantaneously in order to capture both sides of a phone call, miniaturise himself and all his equipment so that he can travel into the body of an ailing character to record the virus attacking her cells and achieve all manner of other difficult-to-countenance feats. Whilst there are examples of films where, within the fiction, there is a camera and camera operator (e.g. “found-footage” films such as The Blair Witch Project) in general there seems little to be gained from
imagining such an entity with the sort of super-powers previously noted; even if we were to take
the view that what we see on the screen is a view from “behind the eyes”, so to speak, of the
narrator in order to dispense with the need to explain the existence of film equipment in ancient
Greece we would still be left wondering why nobody noticed the narrator and how he can
breathe under water, teleport and so on. Our last recourse might be to suggest that we do not
see through the narrator’s eyes, rather we see his memories played out in his mind. Indeed,
they need not even be his own memories: they could be his imaginings based on testimony
from the characters.

This approach would allow us to explain how the narrator could “show” us events to which he
was not a party without the need for invisibility, miniaturisation or gills. There are various films
for which this approach might well be justified: *The Shawshank Redemption*, for example, has
both moments of explicit internal narration through the audience-directed voice-over of Morgan
Freeman’s character Red and scenes where he was definitively not present. There are many
more films which have no explicit internal narrator but for which the existence of a narrator could
be argued, but often this would have to be an unusually well-informed narrator - one who we
would have to imagine had perhaps interviewed all the characters after the events had unfolded
and who had pieced together from their testimonies the ‘whole story’. But what are we to make
of the presentation of a character’s final private moments before committing suicide? The best
that this line of thought can offer is to fall back on the idea that the narrator simply imagines how
that scene *might* have played out, and we might - indeed, should - have questions as to how
reliable those imaginings are. This becomes more significant when the character in question’s
final moments are a departure from her behaviour up to that point - if we see her full of remorse
for her earlier actions, on what possible evidence is this imagined repentance based? Towards
the end of *The Shawshank Redemption* Warden Norton realises there is no escape from the
consequences of what he has done and, rather than face them he kills himself. Without invoking a narrator we are warranted in believing that he is motivated by fear and cowardice, but if this is merely what Red or some other narrator has imagined or been told took place, we must be open to the possibility that those events may have played out differently from how they have been presented to us, and that it is merely the narrator’s (potentially biased or misinformed) opinion that he did not regret more than the simple fact that he had been caught. We can even go so far as to suggest that there exists at least the possibility that he did not commit suicide at all but was killed accidentally, perhaps while cleaning his gun.

Unless the film maker specifically wishes to create this doubt for some artistic purpose there can be little reason for us to go to such lengths to introduce it - it does not add to our aesthetic appreciation of the work but detracts from it, as we can only infer something about the narrator or his sources and not the character him- or herself. There are also films for which even appealing to an unusually well-informed narrator is unsuccessful: take the end of 2001: A Space Odyssey for example - even if the story up to that point could have been pieced together from testimony, recordings, logs, transmissions etc. Dave Bowman’s final experiences could not. Once again, stories which end without survivors or with ‘no way home’ do not readily admit of a narrator that could both know what had occurred and relate it to us. Even though there are circumstances in which a narrator could be reasonably postulated without the artist making this explicit, often there seems to be little to be gained from this enterprise that could not be achieved through careful and considered appreciation of the work and the potential for much to be lost; as Livingston concludes, “[t]he burden of proof is placed on the shoulders of those who claim that we ought to engage in such imaginings”28.

28 Livingston, p. 280.
1.2.4 Conclusions

I hope to have demonstrated that, whilst it is unable to make what might be termed quantitative measures, the usefulness in Currie’s move from narrative as a category to narrativity as a spectrum lies in providing a way to consider the level of contribution a work’s narrative makes to its overall aesthetic richness, and a way to maintain narrative as a part of the aesthetic richness of a work of lower narrativity when comparing it to a work of higher narrativity instead of tempting us to deny the former work’s narrative outright. It should be noted that Currie’s conception of this spectrum includes such things as shopping lists at the very low end; I however, have argued that there are minimum criteria for something to be considered a narrative and shopping lists and mere descriptions of a scene are not narratives and should not therefore find any place on the narrativity spectrum.

Although it might be argued that Currie’s approach would be improved by including such things as interestingness, context and plausibility, these are separate considerations from narrativity itself; a work may be more or less interesting or plausible regardless of its level of narrativity. A given work may suffer from too much or too little narrativity, and the skill of the author lies in finding the right balance between length, narrativity, complexity and the multitude of other criteria by which aesthetic judgements are made.

As I will show in the next chapter, with regard to interactive media, allowing different parts of the story to be higher or lower in narrativity permits us not to have to equate everything outside of cut-scenes with entirely non-narrative works, and in so doing we can be more fine-grained in our critical appreciation of its overall aesthetic quality.
Chapter Two - Interactivity

‘Interactivity’ is a term applied in a myriad of contexts and to an ever-increasing range of items, so in this section I will examine the notion of interactivity and how it relates to works of art in general, and to works of fiction more specifically. I will suggest just what is required for a given work to be considered interactive.

I will also discuss how narratives and their narrators can be interactive, and argue that implicit internal narrators are difficult to countenance in interactive works but that explicit narrators can and have been achieved without significant difficulty. Finally I will argue that a given work is likely to have episodes of greater and lesser narrativity, and that increasing the interactiveness of narratives runs the risk of reducing the quality of the narrative.

2.1 Interactivity

We should begin with the loosest understanding of interactivity: something to which we can do something, or which can do something to us. We can interact with apples, chairs and flowers in that we can eat, move or pick them, with buttons, levers and dials by pressing, throwing or turning them, and so on; both we and the object with which we interact are affected. Although this two-way relationship is what is implicit in a strict definition of interactivity, there is a sense in which one-way relationships can also be thought of as interactive - we get soaked through when we interact with rain and
warmed when we interact with sunshine. In these cases we are merely acted upon, so it is not interactivity in the strict sense, but we can take a more active or passive role in this action and it is in our choices about just how we are acted upon that we can see this looser sense of interaction - passively we may be warmed as an unintended consequence of our walk to the library, but to get an even tan we actively change position while lying in the sun to ensure its interactions with our skin are uniform. If we allow this understanding of one-way interactivity we might even suggest that we interact with clouds and stars by being moved or inspired by them, and the acceptable uses of ‘interactive’ can all-too-quickly expand to be so all-encompassing that they include practically everything, and consequently tell us almost nothing.

When dealing with the notion of interactivity in art, then, it is clear that all works of art can be thought of as interactive in this one-way sense, inasmuch as we hear music, see paintings and so on - we must even turn the pages of a book and move around a sculpture in order to see the works in their entirety - but this sense is too broad, and as a result, uninformative for our purposes. We can take a small step in refining this extreme sense of interaction by suggesting that a work requires some sort of engagement on the part of the audience before they can be said to appreciate the work: we must look at, and not merely see, a painting, perhaps taking note of the figures represented, examining the brush strokes of the painter and inferring something about the emotional content of the work or the emotional state of the artist, we must read and understand the words of a book and not just leaf through the pages glassy-eyed, and we might draw conclusions as to an underlying message to take from it or evaluate the
quality of the plot or the expressiveness of the language, and we must listen to, and not merely hear, musical compositions, and gaining an understanding of their structure, rhythm or melody may help us better appreciate them. So we might say that to interact with a work requires something of an active role of its audience. This understanding of interactivity is still too loose and includes too much for our needs, however, so we must find a way of narrowing its scope.

Although we now require an audience to be active appreciators, it can hardly be the level of activity required that determines a work’s interactivity - larger sculptures or longer symphonies are not more interactive because it takes more effort to circumnavigate them or more time to hear them. In “What is Interactivity?”, Aaron Smuts suggests that interactivity “must be a relational, not an intrinsic, property”. Smuts is attempting to provide a definition of interactivity in general, not specifically related to works of art, in order to “differentiate interactivity from related but incompatible concepts with which it is often confused”, principally the notions of control and responsiveness. In order to deny that an item is interactive either simply if the agent engaging with that item can control the way it presents information (e.g. viewing or reading chapters of a work, whether a DVD or a novel, out of sequence), or if the item responds in some completely determinable way to the actions of the agent (e.g. pressing pause on a DVD player or changing the channel on a television) Smuts suggests the following two definitions:

*X and Y interact* with each other if and only if (1) they are mutually responsive, and (2) neither X nor Y completely control the other, and (3) neither X nor Y responds in a completely random fashion.

*Something is interactive* if and only if it (1) is responsive, (2) does not completely control, (3) is not completely controlled, and (4) does not respond in a completely random fashion.\(^{31}\)

Smuts has been led to these definitions by taking conversation as the ultimate archetype of interaction. He describes two examples of conversation that he believes do not qualify as interactions; in the first our interlocutor responds to whatever we say with random, entirely unpredictable replies, in the second her responses are confined to repeating exactly what we have said to her, or she “barks once for each syllable we utter” and he asserts that in both cases “we would not say that we were interacting with that person … rather, it must be that a certain kind of [non-random] responsiveness absent of control and predictability is necessary for there to be interaction”.\(^{32}\) This requirement that responses must fall into a middle ground between randomness and predictability if the conversation is to be considered an interaction has consequences which are so counterintuitive they suggest it is too unclear to be useful, at the very least for our purposes, and that the above definitions may simply be incorrect.

\(^{31}\) Smuts, “What is Interactivity?”, p. 65.

\(^{32}\) Smuts, “What is Interactivity?”, p. 63.
Smuts is happy to allow that an item may initially be something with which we can interact, but once we gain sufficient mastery of its use that we can reliably predict the output a given input would occasion it ceases to be interactive for us, citing the video game *Pac Man* and word-processing software as examples.\(^{33}\) Presumably the same applies to driving a car - when we are learning to drive we are interacting with the car as we are not able, or at the very least not confident in our ability, to predict with much degree of certainty the effect a given action we might take would have; as our skill improves, we interact less and less with the car. However, if a competent driver is not interacting with her car, what is she doing? Smuts might assert that she is simply using the car, but it is unclear how the situation has changed beyond the transition of her actions and responses to the feedback she receives from the car from the conscious to the subconscious. Whether an item is being interacted with therefore appears for Smuts to be contingent on the level of concentration of the agent, and if this is the case Smuts would be forced to say an inattentive driver is not interacting with his car; this is clearly suspect - if we were to characterise the actions of our bad driver in terms of interactivity it is far more intuitive to say that he is interacting poorly, or interacting without due care. With regard to video games Smuts inadvertently highlights the nub of the problem with his picture of interactivity - he claims that “[n]ot only does an easy game cease to be interactive for you, it ceases to be interesting”;\(^{34}\) he has conflated novelty and interactiveness.

\(^{33}\) Smuts, “What is Interactivity?”, p. 66. There are certain established patterns of movement that, once learned and perfected, prevent the player from being captured by the ghostly adversaries in the game.\(^{34}\) Smuts, “What is Interactivity?”, p. 67.
There are further problems with the application of Smuts’s criteria for interactivity stemming from the vagueness of the not-random-but-not-predictable requirement. Using his example of *Pac Man*, even if I have not mastered the patterns which allow me to defeat the game, I can predict with certainty that if I move the joystick to the left, Pac Man will move to the left. Smuts’s position on this is not stated but it is reasonable to believe that for him Pac Man’s change in direction is not evidence of an interaction but of a response, and it is only with respect to the wider goal of defeating the game that I am interacting with it. Smuts is motivated to take this approach in order to deny that DVDs and televisions are interactive, but in the sense that the actions an agent performs on such items have an effect both on the item (the disc starts spinning, the channel changes) and on the agent (she sees the film, or a different programme) it is clear that they are interactive, even if only in a very weak sense. It is only when we wish to ask whether they constitute an interactive artwork (or, as previously noted, a novel experience) that we need a more precise definition, and as I shall show Smuts’s suggestion has a serious failure in this respect.

A successful definition of an interactive artwork ought to describe a Ming vase as non-interactive. However, if I push the vase over there will be a non-random, non-predictable outcome: the vase will not float or transmute into a potato, it will certainly shatter - but just *how* it will shatter we cannot know with any certainty. Since I can interact with the vase, Smuts’s framework would describe the vase as an interactive artwork (at least for me) and this is clearly unacceptable.
Returning finally to Smuts’s paradigm cases of conversational interaction, he again inadvertently provides us with a clue as to how we might better describe the situations when he states that “we would not say that a successful interaction takes place”. It is not that we have failed to interact with the person, rather our interaction has not resulted in the outcome we desired. Another way to describe the situation would be to draw a distinction (admittedly not always observed in general discourse) between talking to someone and talking with someone - we can talk to a brick wall, a vocal monkey, or an individual who shares no common language with us or who is suffering from some peculiar condition, but we would be unlikely to describe the exchange as talking with them. Equally we can interact with a monkey, at least in the sense that he is affected by our question (he hears our voice, and might be interested or frightened by it) and we are similarly affected by his response but he is unlikely to tell us the time. This brings us to a further problem with Smuts’s characterisation of an interaction in the form of a conversation: if I were to ask my watch-wearing friend the time, I can predict with almost absolute certainty that her response will be to tell me the time and not to recount a recipe for apple strudel, to shriek like a banshee or to attack me with her shoe. It is likely that I can even go so far as to predict, within reason, what time she will tell me it is - certainly if the sun is high in the sky she is unlikely to tell me that it is midnight (assuming, of course, that I am not in some far-north location) but if I know what time I left my office, and I have completed the walk to the park which I know usually takes me about ten minutes and I have only just arrived, I can reasonably predict to within two or

three minutes the time she will tell me. Despite the fact that we are both affected by the exchange, and I achieve my goal of learning the time, Smuts seems to be forced to say that this cannot be considered an interaction, or at best that it is an extremely weak form of interaction. This too is an unacceptable conclusion to draw.

Ultimately Smuts's definition of interactiveness is insufficient for a discussion of interactive works of art because it attempts to reduce too many senses of “interaction” into a single definition; it is more useful to describe the various senses of interaction as resulting from a paucity of language and to concentrate on the sense of interactive that is relevant to the particular discussion at hand, and to be clear about the sense we have in mind.

In *A Philosophy of Computer Art* Dominic Lopes provides a better way of capturing the sense of interactivity we require: 36 ‘users’ here refers to the audience member(s) interacting with the work (as opposed to others watching but not participating in the interaction) and for Lopes the ‘display’ need not be something visual - it refers to an instantiation of the work as appreciated by its audience regardless of medium:

> a work of art is interactive just in case it prescribes that the actions of its users help generate its display … [and it] is interactive to the degree that the actions of its users help generate its display (in prescribed ways). 37

36 Although Lopes is primarily concerned with its role in digital art, interactivity clearly applies to other media and we will examine some such examples in this section.

37 Lopes, pp. 36-7.
Users must have some degree of choice over the actions that they take, and it is fundamental to this conception of interactivity that, had a user made different choices the resulting display would have been different. If no member of the passive audience becomes a user, no choices are made and, depending on how a particular work functions, either the display remains unchanged or there is no display whatsoever (I have in mind a work which has no visible display until it is interacted with, but of course the lack of display could itself be described as the display of the work in its ‘passive’ state).

Lopes’s caveat that the actions of a work's users must be prescribed is important as it discounts such actions as adding graffiti to portraits, beheading sculptures or leaping on stage to protect Mercutio from turning hitherto non-interactive works into interactive ones; Tracey Emin’s *My Bed*, for example, was not made interactive by Yuan Chai and Jian Jun Xi when they jumped on the bed and had a pillow fight - and indeed Chai and Xi consider their actions to be a separate work, a piece of performance art they called *Two Naked Men Jump Into Tracey’s Bed*.

Lopes notes that there are some non-interactive works whose displays change, citing Alexander Calder’s mobiles as an example, and goes on to claim that the reason such works should not be identified as interactive is that they do not “go the way they do because of what we do in the process of appreciating them”.

38 Lopes, p. 42.
*Dark Matter* is a work with certain similar characteristics to Calder’s mobiles but it is displayed in a setting where “what we do in the process of appreciating” it makes it “go the way” it does to a much greater extent; if a work becomes or ceases to be interactive simply because of where it happens to be displayed Lopes’s criteria may be insufficient.

Parker gathered together various unwanted items from the garden sheds of friends, placed them in a shed and had it blown up by the British Army. The remains were then hung much like a mobile, recreating the moment just after the explosion, with a lightbulb suspended in the centre providing the primary illumination and casting shadows on the walls. Hung in a room deep within the gallery, the work was installed in a relatively sealed environment (insofar as how much it was subject to the wind from outside the building), but the traffic of appreciators around the work displaced enough air to cause very slow movement of the individual suspended pieces. Lopes denies that Calder’s mobiles are interactive because the actions their audience members take when appreciating them do not alter their displays, but what should we make of *Cold Dark Matter*? Movement around the piece was a prescribed action, the audience members had choices with respect to their movement insofar as the speed, direction and any pauses they wished to make were concerned, and their choices (what they did in the process of appreciating it) served to alter the display of the work (made it go the way it did) - the orientations of the pieces were changed and consequently the shadows they cast; this work therefore satisfies Lopes’s conditions for interactivity. The key difference we can observe is the extent to which movement around the piece affects its display.

Calder’s mobiles fall into two camps: those originally intended to be indoors (which, given their reliance on air currents generated by passing audiences would, it would
appear, leave Lopes no choice but to call them interactive works) and those intended to be outdoors and moved by the wind. Calder’s outdoor works included mechanisms that required a greater movement of air than could be generated by the movement of audience members around them, but we can imagine a piece more like a traditional mobile that would not; on a calm day the (prescribed) movement of audience members around the work necessary for its proper appreciation would cause its display to change. The changes in its display would not go beyond what was prescribed, so the work would be interactive to the degree to which the movement of the audience affected it. Since this would imply that the work was only interactive on days calm enough for the audience’s effect to be significant, presumably Lopes would argue that we can only assess the degree of a work’s interactivity after we have agreed that it is interactive in the first place - if it “prescribes that the actions of its users help generate its display”, and if this is not so prescribed any apparent interactivity such as we have imagined is purely accidental. Since part of the intention of the Cold Dark Matter was to capture the instant of explosion in a non-static manner, Lopes could argue that the movement of the pieces caused by the audience’s movement around the work is an integral aspect of the piece and Cold Dark Matter should therefore be considered interactive as there is no straightforward way of denying the audience’s prescribed participatory role, whereas in our example of the Calder mobile on a calm day the movement of the piece was not prescribed to be brought about by the circulation of its audience and any apparent interactivity should not be considered interactivity proper. This appears to be inconsistent with other practices in art, however - if a painter were to produce a piece that embodied, through pure chance, much of the styles of cubism we would
nonetheless want to describe the painting as cubist, or at the very least as sharing properties with other cubist works, even if the artist herself were totally unaware of the style or the piece had been completed prior to the accepted beginning of the movement. Certainly some might view this as a contentious claim - my choice to use cubism in this example is not accidental, there is much debate over when the movement began and even what features are work is required to have in order to be included as cubist. However, consider the following two examples: firstly, a movement in art, let's call it schmubism, with clearly defined artistic goals and required features - let's imagine a written manifesto - and a universally-accepted first example. An earlier work is then uncovered which embodies all the required features of schmubism, and a note written by the artist is pinned to the frame, describing a set of principles and artistic goals identical with the schmubist manifesto, even suggesting the name schmubism for this style of art. If it could be demonstrated that the artist of the hitherto-accepted primary example of schmubism had seen both this earlier work and the note pinned to it, had passed off the manifesto as his own and had created his first schmubist painting in response, we would have no choice but to admit the earlier work into the schmubist canon and revise our beliefs about the first example of the movement and who should be credited for its inception. Secondly, consider the practice of deliberate mummification of corpses. Various cultures around the world are known to have carried out this process. Even though the process was discovered and refined totally independently we give all such examples of the practice the same label. Modifying our first example, if it could be demonstrated that the two artists were unaware of each other and that the two manifestos were coincidentally identical, we have a situation
analogous to the mummification example, and we would be similarly justified in calling
the newly-discovered work schmubist and revising our beliefs about the first example of
the movement. We are therefore justified in describing the becalmed mobile as
interactive, even if accidentally so.

A possible counter-objection Lopes might raise to this argument is that the work’s
apparent interactivity arises because of a change in state - the lack of wind - and if we
were to bring an outdoor work inside so that any movement or change in its display can
only be caused by actions of the audience the work would not suddenly become
interactive; like Chai and Xi we would have changed the original work into a new one.
However, this is too bold a claim. If we were to add a mechanism in front of the Mona
Lisa which, when audiences walked over certain pressure-plates in the floor nearby,
caused flecks of paint to be sprayed on it we would certainly not have made the Mona
Lisa interactive, we would indeed have created a new work, just as if we placed a
mobile in a sealed box with a fan which could be operated by audience members via a
switch. In the becalmed mobile example, however, it seems that we better describe the
work as having features which can be interactive, albeit inadvertently; we have not
created a new work simply by viewing it on a calm day.

There is an additional problem for Lopes’s claim that an interactive work “goes” the way
it does because of what we do in appreciating it: it is inconsistent with a claim he makes
earlier in A Philosophy of Computer Art. Lopes describes a work by Hisako Yamakawa,
Kodama - Mischievous Echoes (2005): a room containing moving images of a forest which is silent while you are in it but from which voices can be heard when you approach or leave it; the voices are recordings of audience members made while they are in the room. The problem for Lopes is this: what you do in appreciating the work is to hear the voices as you approach, to discover the room empty (but for the forest images) and silent, and to hear the voices again as you leave. Should you happen to speak whilst in the room, you might hear your own voice as you walk away from the it, but speaking whilst in the room is neither a required nor a prescribed activity necessary to appreciate the work - it goes the way it does (particularly as you approach the room, before your voice can be recorded) because of what others have done. It is not enough for Lopes to respond that the work goes the way it does for you because of what others have done in appreciating it: the work cannot distinguish between what is said in the room in the act of appreciating the work and, for example, gallery staff or visitors disinterested in the work who have walked through the room talking without appreciating the work at all, or visitors who are interested in the work but whose conversation strays into areas not related to the appreciation of the work (i.e. they are initially appreciating the work, but at the moment at which the work records what we might call the “irrelevant” part of their conversation they are not, although they may subsequently return to appreciating it) - the work goes the way it does because of things that were done outside of the act of appreciating it.

Lopes, however, claims that Kodama is an

39 Lopes, pp. 25-6.
40 It is even arguable that hearing recordings of conversations not related to the appreciation of the work adds greater variety to what we hear when approaching and leaving the room, and that this variety adds to our experience of the work; rather than hearing only comments about the piece we hear snippets of conversations about all manner of subjects, or perhaps just one side of a conversation conducted on the phone, and this might make the piece more interesting for us.
interactive work. Clearly there are many other works which would be open to the same objection. Lopes introduced this requirement in order to exclude certain works with changing displays from being mis-labelled, in his view, as interactive. However, other works he would describe as interactive do not necessarily satisfy this requirement. I suggest that this claim should be dropped and we instead allow firstly that inputs may come from users who are not engaged in appreciating the work, and secondly that some works might be very weakly interactive - so weakly interactive that in their usual settings the interactivity is so weak we would be justified in disregarding it, or not appreciating the work as an interactive piece. In the right settings however we have no choice but to recognise certain latent interactive features, even if in usual circumstances there is no norm that users’ actions alter the work’s display. Lopes asserts that a work is only interactive when it “prescribes that the actions of its users help generate its display”; we should instead assert that works may be interactive if there is no prescription against users’ actions changing its display, as this allows us to include works with the sorts of extremely weak interactive features just described but still excludes acts of vandalism such as spraying flecks of paint over the Mona Lisa or acts that create new works in the manner of Chai and Xi’s Two Naked Men Jump Into Tracey’s Bed.

We should note that the display itself need not be created directly by the artist; for Lopes the artist’s role is to have a certain vision of how user inputs can be turned into display outputs. This distinction can be made clearer by considering the work of art as

a whole and the individual instantiations of the work; just as it is intuitive to credit the composer for a musical work and the musicians for a particular instantiation of it, so the artist is credited for the work-as-whole. By extension it would seem logical - and indeed necessary - to credit the users of an interactive work for an instantiation of it. At first this might seem counter-intuitive - for it is reasonable to assume that most users would lack any skill at ‘performing’ a piece of interactive art in the way we would expect musicians at a concert to be skilled at performing a piece of music. Indeed, ad hoc collections of amateur musicians could not be expected to deliver the accuracy and subtlety that audiences would require to gain a proper appreciation of a musical work. However, this is the point at which interactive art is able to diverge from the analogy; although certain examples of interactive art may, with their users, generate instantiations which themselves are of aesthetic interest, unlike musical compositions interactive art can easily accommodate examples whose instantiations are of limited or no aesthetic interest altogether. Instead, each instantiation helps to elucidate the vision of the artist, and it is the work-as-whole that is of aesthetic interest.

2.1.1 Audience and User

One interesting feature of interactive art is how it opens up a second category of audience - that of the user. A user is most often, although not always, also a member of the audience of a work (those who perceive the display of the piece), but a user interacts with the work in prescribed ways to alter its display. It could be claimed that this category already existed in the role of musicians performing a composer’s work, at
least in terms of the individual expression a skilled musician adds when performing a piece. Interactive art makes much lower demands in terms of its users' skills and this category is thus far more inclusive and accessible. More importantly, few musical works are composed with the musicians' aesthetic enjoyment of their role in the performance in mind, whereas interactive works are often created such that the change in display is immediately (and, hopefully, engagingly) apparent to the user(s) for their appreciation. This new, or at the very least greatly expanded, category of audience might engage more directly with a work and in such works any emotional responses the work elicits can reasonably be expected to be stronger and more immediate for them than those of the passive, non-user audience, as not only can the users become at least part, and potentially the primary, focus of the work, often they experience actually being part of the work themselves. One example of this is Panda Eyes by the Jason Bruge Studio, in which one hundred plastic panda-shaped collection-boxes from the World Wide Fund for Nature are arranged ten-by-ten and fitted with motion sensors and motorised, rotating bases. When a single user approaches, all of the pandas spin round and face him or her; when multiple users approach, the pandas split their attentions between them, and when a user moves around the piece the bears rotate to follow them. Whether the work is viewed from the perspective of a user or passive audience-member there is a sense that the bears are standing in silent judgement at people’s failure to prevent the destruction of habitats and endangerment or extinction of species.

42 Equally, a work could be designed such that users are distanced from the output and the audience given a more privileged relation to it; in this case those audience members who perceive the isolation of the users from the output are likely to be afforded a greater opportunity for engagement than those who do not.
However, for a user looking at row upon row of pandas, whose mute, glassy-eyed and condemnatory stares follow them inescapably wherever they move, this feeling is much more powerful, their apparent accusations seem more personally directed and the user feels a stronger sense of unavoidable accountability.

*Panda Eyes* also serves as a good illustration of how interactive works, and the different experiences audiences and users have of them, can be appreciated in different ways; for audience members it seems reasonable to assume that their appreciation of the work-as-whole will be more satisfying than the instantiation of it they witnessed (“wherever anybody went they could not avoid the gaze of so many small and silent protesting pandas” vs “when the chap walked around the case some of the models turned in his direction”), but it is harder to conclude that the same holds for users - they will have the same appreciation of the work-as-whole, but their experiences of (or, indeed, as part of) an instantiation of it will be quite different - “when I tried to walk around them they kept turning and following me, and even when other people came close enough to attract the attentions of some of them the pandas closest to me were always staring right at me”. The directness of this relationship between work and user is a key difference between interactive and non-interactive works that can affect the strength of emotional responses a work elicits\(^{43}\) and, since the actions of the user play a role in determining the display of the work, this opens up the possibility that some works

---

\(^{43}\) One can further imagine a work which gives rise to different reactions from its users and its audience, and if the experiences of user and audience-member were each not apparent to the other aesthetic appreciation of it might even be three-fold: appreciation from the perspectives of user and audience-member and, after experiencing both perspectives, appreciation of the work as a whole.
may be able generate emotional responses that are unique to interactive media. One area that immediately lends itself to a discussion of the active role a user plays when engaging with an interactive work is that of videogames, where users (in this context, players) interact with a fictional world, fictionally take actions within that world and make certain things fictionally the case, and - importantly - often have the ability to go back and try things again a different way.

### 2.1.12 Reversibility

Reversibility is an important feature in interactive works; we should note, however, that although reversibility implies some amount of interactivity in a work, not all interactive works are similarly reversible. In non-narrative works there may be no end-point, they are reversible in the sense that the displays are transient and can be changed again and again depending on the inputs they receive from their users. In other non-narrative works there may come a point where the display reaches some sort of end (after a certain amount of time, or once the display is ‘filled’ in some sense) after which it resets and allows the users to start again. In narrative interactive works, and video games in particular, the sense of reversibility is very strong - the death of the player’s character will require either the entire game or, more often, the current section or ‘level’ of the game to begin again, or the player can revert back to a previous ‘save point’ or elect to begin the game anew whenever she chooses. We should recall that there are some films such as *Run, Lola, Run* and *Blind Chance* in which multiple scenarios are played out, but it is not the viewer who determines which scenes are played out or who
chooses when to stop or how to alter a given scenario so they are not ‘reversible’ in the same sense. The reversibility in video games or interactive narratives more generally might make interactive media an extremely effective tool for ethical learning, as it allows players the freedom to make mistakes without real-world consequences, and to go back and try different strategies to test the results that each course of action would bring about.

In “Videogames and Interactive Fiction”, Grant Tavinor uses Walton’s terms “work-worlds” and “game-worlds” to differentiate between facts that are fictionally true arising from the author of the fiction and facts that are fictionally true of a player's interactions with a fiction. What makes interactive media so different from traditional fictions is that much of the work-world is contingent upon the actions taken in the game-world, and this blurs the line between the two worlds as “the game-world effectively projects into the work-world because of the fictional interaction”. Because players have a certain degree of freedom to explore the fictional worlds with which they are presented, both in terms of exploring areas of territory and interacting with non-player characters that might divulge useful or informative facts about the fictional world, the player has a certain responsibility for his knowledge about his and other characters and their places within the world that does not occur with non-interactive media.

Note that Tavinor’s use of ‘game-world’ differs from Walton’s in that it does not include games of make-believe independant of the video game, i.e. using a video game as a prop to make-believe something totally different from the content of the world of the video game as designed by its creators.

In both interactive and non-interactive narratives there might be references to real places or historical figures and knowledge of some fact about them not explicitly stated in the work might be necessary to understand or appreciate the work fully (indeed, for sequels and series such as Terry Pratchett’s *Discworld* novels background facts about fictional characters and places might be required), audience members might be considered to be responsible for learning this information, for being a relevantly-informed audience - they would also likely be held responsible for following the story as it unfolded by paying due attention. What is unique to interactive fictions is the ability for users to choose to alter their game-world such that additional explicit information about the work-world can be uncovered that goes beyond implicit details we might learn by making-believe what a particular glance between characters in a film might plausibly imply or what the personal character of the subject of a portrait might be by calling to mind individuals with certain relevant traits. Participants can concentrate on achieving the core tasks quickly and experience a game-world that is something roughly analogous to a novella or short story, or they can press non-player characters for additional information, take on and complete additional “side” tasks and generally explore the work world with which they are engaged to experience a game-world more closely analogous to an epic novel. Gaining access to these details about the fictional world and the histories and motivations of the characters within it is, to the extent that is his choice, the responsibility of the user; films, books and plays present information about the fictional world only when the author chooses to, so audience members have no such responsibility. As we shall see, in some modern video games players are also given a certain degree of responsibility for the path which the narrative of the game
takes, not just the amount of narrative revealed. In these games, should a player decide or be required to replay part of the game, a new game-world is instantiated, and this can cause a very different work-world to be created in which different decisions can be made from those in a previous instantiation, either because they were mistakes that the player does not wish to repeat or because there is interest in seeing how the effects of taking an alternate decision play out.

Although Tavinor suggests that complex narratives would require so vast an amount of programming as to be impossible, in recent games like *Heavy Rain* the work-world the player experiences can be radically different depending on seemingly insignificant choices – not all the characters need to survive in order to complete the game, and there are a variety of possible final outcomes. I will examine this in greater detail in the following section of this chapter.

### 2.1.3 Conclusions

I hope to have demonstrated a reasonable way of understanding what we mean when we talk of interactive works, and how their interactivity goes significantly beyond that of menus on DVDs and on-demand films. This definition excludes vandalism and other unprescribed actions, but is otherwise quite inclusive, and allows for unintentionally interactive works and works which take inputs from users not actively appreciating the work. I have also argued that this new category of audience, the user, allows for potentially three separate perspectives, emotional responses to and judgements about
the work: those of the user, the non-user audience, and (where there is some separation between how users and the non-user audience relate to the work) those arising from people who have been both audience and user. Since the actions of the users in part determine the display of the work, there is room for a greater range of emotional responses than can be elicited by interactive works. I have also argued that users have a certain responsibility for investigating their game-world in order to discover more about the work-world that does not apply to non-interactive works, and that reversibility is a key factor in what makes interactive works interesting, particularly where new instantiations of game-worlds allow for radically different work-worlds to be created. I have also suggested that this reversibility has the potential to make interactive works effective tools for ethical learning.

2.2 Interactive Narrative and Interactive Narration

We turn now to the possibility and the aesthetic quality of introducing interactivity to narration and narratives. I will show that interactive narration is certainly possible - at least where an explicit internal narrator is present - and that, although possible to some degree, interactive narratives pose a significant challenge and that they may require us to approach our aesthetic appreciation of them differently from non-interactive works.

2.2.1 Interactive Narration
In the previous chapter I argued that explicit internal narrators were, by and large, unproblematic, and that we should only posit an implicit internal narrator when we had good grounds for doing so, with the result that many non-interactive works had only an external narrator, namely the author. Although there are examples of video games that include an explicit interactive internal narrator, I will argue that implicit internal narrators are difficult to make sense of in an interactive work, in no small part as a consequence of the authorial role the player takes on when engaging with such works. This discussion will be focused on video games as the main exemplars, as they are by some measure the most accessible forms of interactive works which incorporate a narrative aspect.

We shall begin with a simple sketch of what we would require to describe something as an interactive narration. A non-interactive narration is the relating of occurrences within the story and the key feature of an interactive work is the changing of a work’s display as a result of the act of appreciation by the work’s user (so long as there is no prescription against an appreciator changing the work’s display), so a good starting point for a conception of an interactive narration would be the relating of occurrences within the story that have been brought about by the player engaging with the work. Clearly the level of interactivity within the game will determine the maximum potential interactive narration as the fewer meaningful choices a player has throughout the course of the game the fewer opportunities there will be for satisfying our definition of interactive narration. Where a game is broadly linear (i.e. few meaningful choices), *Crysis 2*, for example, the narration within the work cannot hope to be more than
minimally interactive. This minimal narrational interactivity is likely to be nothing more than the triggering of a cut-scene or dialogue (usually from non-player characters) once a given task has been completed or a certain checkpoint reached. By far the most common example of this sort of interactive narration will follow a simple structure: the player achieves the goal in question, the game then provides the player with new information - usually a brief confirmation that the goal has indeed been reached and a longer description of the events in the story that lead the player-character from her current situation to the next task she needs to complete and a brief description of what the new goal is. Although this satisfies our requirement that interactive narration relates occurrences only when the user has changed the work’s display in a suitable fashion, this is so minimal a sense of interactivity as to be barely noteworthy. Games with a greater depth of interactive narration would have various possible cut-scenes and dialogues, and the user is presented with the appropriate scene only. This too is still very minimal, but there is more to interactive narration than the cut-scenes. The world of the video game (distinct from Walton’s conception of the game-world) is authored by the video game’s developers, but the work-world (in Walton’s sense) is co-authored by the developers (who fix the setting and various facts about the world, the player-character, the non-player characters etc.) and the user. Where the player is making choices about which lever to throw (assuming each leads down a different potential branch of the story) or which dialogue options to select (even if certain options are there only to reveal non-essential but relevant details of the world of the video game) the external narration is affected by the user’s choices and can reasonably be described as interactive. However, it is clear that the notion of an implicit internal narrator will quickly
run into difficulties as the level of interactivity in a given game increases - whether the goal of introducing an implicit internal narrator is to shed light on aesthetic qualities of the story, it must surely fail to do so when the user is dashing about fruitlessly or exploring the wider world of the video game\textsuperscript{46}. This is particularly clear when we consider that much ‘exploration’ is done when the user is less focused on the storyline of the work, perhaps because he is tired or in need of some respite from the frenetic button-mashing in which he has until that point in time been engaged. These are factors outwith the fictional world and cannot reasonably be considered an integral part of the aesthetic aims of the work. The notion of an implicit internal narrator as storyteller might be taken to be more convincing, particularly when the player’s character makes ‘spontaneous’ (i.e. not by the choice of the player) comments about the situations she encounters. Sometimes these comments might be in order to provide clues to the player as to what should be examined or utilised to solve a particular puzzle, or they might be simply to add ‘colour’ to the proceedings. However, in these cases it is arguable that it is true within the fiction that the character is providing narration, in which case it is explicit. Where it is unclear that it is true within the fiction, we are left with the familiar questions of just who this narrator is, how reliable their narration is, why they are never noticed or acknowledged by any of the other characters (particularly if they are hovering above and behind the player’s character with a ‘camera’ at all times) and so on. Where the game is in first-person perspective we can also ask how the narrator is able to see through the character’s eyes, and if it is suggested that

\textsuperscript{46} The additional story information would be largely, if not wholly, irrelevant, and might well be considered detrimental to the wider story.
the character is the narrator we must question how *they* are able to present what they see and hear to us. In addition given the close relationship between player and character it seems that the player herself takes on at least part of the role of narrator, suggesting that the player is narrating to herself just what she is making fictionally true as she does it, which does not appear to be a useful position to adopt. Even if satisfactory answers could be found, we are still left with a narrator who relates in equal detail events pertinent to the story and aimless, unconstructive episodes that have no bearing or relevance. We must conclude that high levels of interactivity do not of themselves benefit any implicit internal narration that might be present, and indeed risk being detrimental to it; if there is evidence for implicit internal narration it will be in non-interactive elements of the game.

Explicit internal narration can incorporate interactivity with much greater ease, as exemplified by *Bastion*, from Supergiant Games. This game begins with a voice-over describing events that happened some time in the past, and the user controls the central character of the narrator’s story, a boy referred to only as “the Kid”. The population of the player-character’s world have been petrified and have become dead, fragile statues. As the user progresses through the game, she is free to attack these statues and destroy them. Depending on whether she chooses to destroy them, the voice-over narrates appropriately. Some of the statues were dear friends of the player-character and, should the user choose to destroy them, the narrator describes the emotional pain the Kid felt, despite knowing that “they wouldn’t want to be left that way” or excuses his actions with comments along the lines of “he was so angry he couldn’t
bear to look at them like that”; others the Kid had no love for and the narrator comments that he “took grim satisfaction in turning her to dust” and so on. As the user completes (or not) the various tasks, the narrator explains what was achieved or lost as a consequence. Explicit internal interactive narration is, then, relatively straightforward, and external interactive narration at least makes sense as a concept even if in highly interactive games it is not particularly artistically valuable. Implicit internal interactive narration however is a best highly problematic, and if it adds anything to the aesthetic appreciation of a game it is highly likely that the contribution will not be positive.

### 2.2.2 Interactive Narratives

In the previous section on narrative and narration, I claimed that the issues arising from the differences between the level of narrativity of a work and the aesthetic quality of that narrativity apply to at least as great an extent to interactive media as they do to non-interactive media. To see why, we can return to *Heavy Rain* – the story itself is relatively short, involving a small cast of inter-related characters, but it is quite compelling. The role-playing game *Witcher 2* also involves a fairly small cast of main characters (if we disregard optional “side quest” characters which serve pragmatic uses with respect to player-character development but which are irrelevant to the central story of the game) but it is considerably longer than *Heavy Rain* and requires the completion of many individual quests (perhaps roughly analogous to chapters). Whilst players may find each segment of the story interesting and engaging while they are playing through it and enjoy finding solutions to the problems they encounter,
considered as a whole the story is quite meandering and rather less compelling than that of *Heavy Rain*. The visibility of these issues may be greater because there are additional factors that need to be considered when assessing both the level and the success of a game's narrative. Currie's notion of higher- or lower-narrativity is helpful in understanding why: in both of the previous examples there are segments where the level of narrativity switches – it is higher when checkpoints are reached or quests completed, usually occasioning a cut-scene where the player's control over the player-character is greatly reduced or totally suspended, and lower (indeed, where the player's actions can have no effect on the way the storyline progresses, such as in *Crysis 2*, it is fair to claim that it is absent altogether) when the player is engaged in completing the next task. In most, if not all, games to date, an increase in interactivity is usually coupled with a decrease in narrativity, and to a degree this is inevitable - the greater control a player has over a character's actions the greater the scope the player has for having his character act in a way that does little or nothing to advance the story. *Heavy Rain* enjoys greater success in managing the transitions between states of narrativity than *Witcher 2*, possibly because it is more linear in its style of gameplay. The achievement of an interactive work's expressive or communicative goal may also be better served by a shorter storyline given that one's focus is often highest when engaged with the problem-solving or decision-making elements of the work – and for video games with high levels of action or which require great concentration in the segments of lower narrativity players may well use overly long segments of high narrativity as a complete break from the game.
In the preceding paragraphs I noted that where interactive works do not provide meaningful choices for players (in the sense that their choices alter the outcome or at the very least the intermediate content of the story) the segments outwith cut-scenes lack any narrativity; in these cases the narratives of the works should be considered non-interactive. For any work to have an interactive narrative the audience engaging with it must be given some sort of collaborative role in the authorship of the work (or, more accurately, the work-instance). In *A Philosophy of Cinematic Art*, Berys Gaut provides a clear way of understanding how this collaboration can be understood with a description of Jane and her young son Otis:\(^47\): to begin with, Jane repeatedly tells Otis a story she has made up called 'Teddy and the Dragon' to which Otis is initially a non-interactive audience. As he becomes bored by the story Otis demands that changes be made, but Jane is not prepared to do this – teddy continues to enter the forest and meet the dragon, and the narrative remains non-interactive. Over time, however, Jane relents and starts to allow Otis to make decisions on behalf of teddy, improvising new elements of the story as required; at this point Otis is partly responsible for the content of Jane's story, and the narrative has become interactive. Jane goes on to create a video game based on their stories in which a player takes the part of teddy and can fictionally fight dragons or hunt for treasure. Had Jane taken the *Crysis 2* approach, she could have allowed Otis to have teddy wander about the fictional world but refuse to change any of the pertinent details of the story – endlessly repeating “Teddy wanders aimlessly about the beach, but he has a nagging feeling he ought to go back and enter the forest”. Whilst this might be considered to have some highly minimal sense of

\(^{47}\) Berys Gaut, *A Philosophy of Cinematic Art*, pp. 230-1
interactivity it is clearly not enough to warrant calling it an interactive narrative. When creating her videogame, Jane allows the player to choose between entering the forest and going to the beach, and instead of being able to create just a forest and a dragon she doubles the work she has to do: she now needs to create a beach and some treasure as well. Clearly, as the number of meaningful player choices increases the amount of programming required escalates rapidly, and for modern highly complex video games created with limited resources this seems to suggest that there must be a trade-off between the interactivity and quality of the narrative; whilst there are various strategies to limit this trade-off, to an extent it is inevitable. This variability of content and outcome is a distinguishing feature of interactive media and it is already becoming a generally-accepted criterion in critical reviews of such works as audiences increasingly look for “replay value” – being able to begin a new instance of a narrative, make different choices and discover the ramifications they have. As creators and consumers of interactive works add more weight to the importance of satisfying player-choices “the object of aesthetic appreciation [will shift] in part to the story tree, not just appreciation of each individual story line”48.

2.2.3 Conclusions

The more interactive a narrative becomes the less useful the notion of an implicit internal narrator becomes, but explicit internal narration is far more straightforwardly

48 Ibid., p. 240.
understood and realised within video games. Narratives are interactive if they respond to actions the player has taken within the game, but highly interactive narratives are liable to lose the sense of ‘purposeful’ or (directly) ‘goal-driven’ activity that identifies better narratives. It is also likely that the thematic unity of interactive narratives will tend to the more general the more interactive it is. Greater interactivity in narratives also risks expanding the story to such an extent that, taken as a whole, it is less satisfying or communicates the work’s expressive goal less well than narratives lower in interactivity. Replayability is a key factor in assessing the interestingness of interactive narratives and this indicates that both the quantity and the quality of the different branches the story might take are important criteria in judging the aesthetic value of these works.
Chapter 3 - The Paradox of Fiction

When we learn the details of some happy or sad event, often we will experience certain emotions in response. When the events in question have actually happened to real people, these responses are unsurprising - we feel happiness upon hearing of the birth of a friend’s baby, outrage at the news of some injustice and so on. However we often also report having emotional responses to fictional events; we might feel tense while watching the climactic confrontation of a play, or pity for a fictional character who has, within the fictional world of the book we are reading, suffered a tragedy, and we might describe ourselves as having been frightened by the flesh-eating zombie depicted in a horror film. It has been suggested that, since there is nothing really at stake in the exchange, no-one has actually suffered and we are not in fact in any danger from the monster on the screen, we cannot have genuine emotional reactions to such fictional events - or that, if we do, our responses are irrational. I will present some of the arguments that have been made for and against this view, and argue that it is both possible and rational to feel genuine emotions towards fictional situations and characters, but that for non-interactive cases responses of the final type (fear for oneself, jealousy, guilt, shame and so on) are only possible in unusual cases and, should they arise, they are often irrational.

3.1 Emotions and Beliefs

The cognitive-evaluative theory of emotions requires emotions to be about an object: we evaluate a slavering dog as a potential threat or events that have befallen someone as upsetting, and we are frightened of it or pity her. We can frame these evaluations as beliefs (I
am frightened of the dog because I believe it can harm me) and we would expect to be motivated to take an appropriate course of action - if I believe the dog can harm me I will be motivated to avoid it. We can also frame evaluations in terms of imaginings - I am frightened of the dog because I imagine it can harm me; I shall return to this later in the chapter, but it will be helpful to concentrate initially on evaluations as beliefs. What is at issue with regard to emotions directed at fictional entities is this: if we pity someone because we believe that she has suffered and we fear something because we believe it can harm us, how can we hold these beliefs when we know that the objects of our emotions do not exist? Further, if we generally expect emotions to be motivating, why do we not attempt to intervene or flee?

Colin Radford provides an example of emotional response founded on belief:49 a friend relates a horrifying story of the misfortunes suffered by his sister. The pity we feel for her dissipates when he reveals that he in fact has no sister. Carroll notes that there are other examples where we would expect a change in beliefs to result in a change in emotions, such as telling a child afraid of ghosts that there are no such things.50 Carroll’s example clearly accords with our everyday experiences; generations of parents have had to comfort and reassure their children after letting them watch frightening or upsetting films, whether the change in belief is effected by convincing them of the monster’s non-existence or by appending a happier resolution to a fiction (“he apologised and they all became friends in the end”), and adults’ perception of injustice regarding real-world situations vanishes should convincing evidence comes to light that the apparent victim did indeed commit the crime of which he has been accused. However, it is clear that we can say that Radford’s example is not analogous to fiction in that it involves

49 Radford, “How Can We Be Moved By The Fate Of Anna Karenina”, p. 68.  
50 Carroll, The Philosophy of Horror, p. 60.
deception (indeed our discovery of this deception might replace our pity with anger or shame), and this makes the case quite distant from our usual engagements with fictional works.

### 3.2 The Possibility of Fiction-directed Emotions

In his *Mimesis as Make-Believe*, Kendall Walton puts forward the view that when we experience emotional responses to fictions what we feel are not genuine emotions but some sort of quasi-emotions. It is important to note that Walton does allow that the depictions in fictions can cause genuine emotions; Walton describes an audience member, Charles, who watches a horror movie about green slime and claims to have been scared by the slime – Walton suggests that if Charles had a heart condition he might fear the *depiction* of the slime, as the depiction is real and if it is highly effective it might cause Charles to have a heart attack, but that if the slime does not exist (and Charles believes it does not exist) Charles cannot be said to fear the slime itself.\(^{51}\) Walton’s objection is that Charles knows that the slime poses no physical threat to him (isolated as it is within the fiction), and if we consider fear for oneself to involve a belief that one is in danger from something, it is impossible and inaccurate to describe Charles as fearing it – the object of his response does not exist so it cannot affect Charles and Charles cannot affect it, how then can he be said to fear it? Walton suggests that to allow this would be to sever “the normal links between the physical and the psychological”.\(^{52}\) Although we should note that this is only true should Charles believe that what he is seeing is fictional - or rather, that even if Charles were to believe that the slime is real he also believes that he is only seeing a depiction of it, and that a mere depiction cannot harm him (heart conditions aside), it is fair to grant the typical movie-goer such awareness. However, that Charles believes he and the slime cannot affect one another is not enough to claim that he does not fear it, rather, we can only argue that


\(^{52}\) Walton, *Mimesis as Make-Believe*, p.196.
he cannot rationally fear it. Further, if we consider emotions other than fear it should be clear that the absence of the possibility to affect or be affected by an object does not prohibit feeling an emotion towards it.

Take, for instance, real situations that are distant from us, whether temporally or geographically; we can do nothing to aid Russian submariners trapped underwater but we would not wish to characterise our pity for them as fictional or make-believe, even if we come to learn of their plight after their death.

Walton dismisses the idea that Charles has 'suspended his disbelief' as this seems to suggest that in some way Charles is not definitively sure that the slime does not exist - that part of him believes it exists and another part believes that it does not - and it does not seem plausible that Charles could feel anything less than certain about this without taking some sort of action (calling his family/ the police etc.). This is not an accurate characterisation of suspending one's disbelief; it is not that there exists some uncertainty about the veracity of what one is seeing, but that one chooses to set aside the belief that it is fictional. I do not consider appeals to the suspension of disbelief an adequate line of response, however, as not only is the notion of choosing to have or not to have a belief problematic (consider being threatened with death if you do not choose to believe something you know not to be true; even in cases where people claim to choose to believe in someone’s honesty or fidelity, it is more accurate to say that they choose to act as if they held that belief), it should be clear that, as Radford puts it, no matter how good the set, the lighting etc. "we are never unaware that we are watching a play". Nor can Charles be said to suffer from momentary lapses in his grasp of reality that are too short for him to feel a compulsion to take an action as not only is he scared throughout the film, this

---

53 Radford, “How Can We Be Moved By The Fate Of Anna Karenina”, p. 72
approach also fails to explain how we might pity Anna Karenina or admire Superman, as these are not obviously induced through sudden, momentary cinematic technique\textsuperscript{54}. We should also note that emotions such as these may last long after the work itself has ended – we can imagine a child who read a Superman comic and whose apparent admiration for him continues undiminished for many years – and the 'momentary lapse' approach cannot account for this.

Walton goes on to note that Charles's responses are not subdued or lessened by the part of him that suspects that the slime is fictional, and suggests that it is similarly difficult to explain why this should be the case. Although Walton raises this point in support of his denial of his characterisation of suspension of disbelief it is also problematic for his account, as he concedes that Charles does have certain responses we would expect to observe of someone experiencing fear. Walton argues that this is not a “gut feeling”\textsuperscript{55} that might cause an individual to act contrary to how they know cerebrally they should act, as Charles takes no action to remove himself from the source of his supposed fear. The changes he does experience are automatic reactions that are not under his control, and “[o]ne doesn't have reasons for things one doesn't do”.\textsuperscript{56} Whilst I concede that one may not have reasons to sweat in hot conditions or shiver in the cold, there are certainly explanations for such physiological responses, and we can account for Charles’s raised pulse and white knuckles by suggesting that although he does not have a belief that he is in danger, a primitive part of his brain recognises certain sensory inputs as signs of potential danger and reacts accordingly, putting his body in the same state as if he really were in danger. We can therefore say that Charles is at least in a genuinely fearful mood, even if the belief that he is in danger would be irrational.

\textsuperscript{54} Walton, \textit{Mimesis as Make-Believe}, p.199.
\textsuperscript{55} Walton, \textit{Mimesis as Make-Believe}, p. 198.
\textsuperscript{56} Walton, \textit{Mimesis as Make-Believe}, p. 199.
As Walton’s assertions rest on the notion that to fear the slime Charles must believe he is endangered by it, he examines possible counterarguments that might undermine this view – can there be fear without a connected belief? Walton cites the following example from Patricia Greenspan’s *Emotions and Evaluations*: Frances was once bitten by a dog and now feels fear whenever she is near any dog, even a toothless, harmless old dog called Fido. Since she is aware that Fido does not pose a danger to her, Greenspan suggests that it is difficult to explain her fear as requiring a belief that Fido is a threat. Although Walton allows that exactly what constitutes a belief is far from precisely understood, his principal objection is that Charles does not act in a way that betrays a feeling of fear, whereas Frances does by trying to avoid Fido or running away from him. For Walton “[f]ear is motivating in distinctive ways, whether or not its motivational force is attributed to cognitive elements in it … Fear emasculated by subtracting its distinctive motivational force is not fear at all”⁵⁷. Walton distinguishes between the emotions experienced by Charles at the movie, by Frances fleeing from Fido and by Aaron, a man afraid of flying steeling himself before take-off, as different types of fear, suggesting that “[t]o assimilate them would be to emphasize superficial similarities at the expense of fundamental differences”⁵⁸ and that we should consider genuine emotions as belief-desire complexes. What can fall under the umbrella of ‘fear’ could therefore include something that motivates action in a similar way to the belief that one is in danger. It is unclear exactly what one must be motivated to do for Walton to consider the feeling to be genuine fear, or how fear’s motivational force is distinctive – as I shall return to later, extreme fear may not motivate you to do anything, it may simply overwhelm and paralyse you, and you might be motivated to act in a way typically associated with fear without the belief that you are in danger, screaming and clutching the

---

restraints of a roller-coaster ride you are enjoying, for example. Walton concedes that it is not readily apparent what motivational force should be associated with grief.

But if Charles does not genuinely fear the slime, how can we describe what he is experiencing, and why does he describe his experience as fear? Walton's position is that Charles is engaged in a game of make-believe in which he fictionally fears the slime. He compares him to a child, Timmy, who is playing a game in which his father is a scary monster chasing him from room to room – although he may be laughing at the same time, Timmy is fictionally scared by the fictional danger. Charles too is fictionally scared by the slime, by the danger in which it is fictional he is in his game; Walton describes him as both actor and object within his game, “a reflexive prop generating fictional truths about himself” similar to an actor playing himself within a play. Timmy presents us with a middle ground between Charles and a stage actor, in that Timmy’s fear may be partly performance, like an actor’s, and partly the quasi-fear that Charles feels.

Since he could fail to react (behaviourally, if not physiologically) but still experience the same quasi-fear when the monster attacks, Timmy’s fictional fear depends on his mental state, not just on his performance. Walton suggests that Charles's behaviour is similar in that he does not perform to show others his fear; his behaviour does not account for his (fictional) fear. However, that his behaviour is not a performance, or may not be in line with what one might ordinarily expect to see of someone experiencing fear, does not seem to be enough to qualify his experience as fictional, or quasi-fear – after all, people facing extreme real-life situations may react in unexpected ways; consider crossing the road and two steps across seeing a car.

59 Walton, Mimesis as Make-Believe, p.242.
60 Walton, Mimesis as Make-Believe, p.242.
bearing down on you – whilst you undoubtedly feel fear (and consider the car to be a danger to you) it is plausible that you might not take any action to avoid the impact, your fear having rooted you to the spot. It might be suggested that there is still a motivation to flee, but that it is overridden in some way, perhaps by the sudden rush of adrenaline. However, reflecting on a near-miss I experienced as a child leads me to believe that this is not the case - as I recall, the closest thing to a formulated desire I had was for the car just to stop, and whilst some people’s retelling of such experiences may include assertions such as “I wanted to dive out of the way, but I just couldn’t move”, still others will not. The best that can be argued from anecdotal reports like these is that sometimes the agent will fail to recognise the motivating force of their fear, but often people will describe their perception of time “slowing down”, they may be able to recall with remarkable lucidity details of the few seconds of the event, or of “my life flashing before my eyes”. It is at the very least highly suggestive that many accounts do not include any reference to fleeing or jumping out of the way, and if we persist in maintaining that the motivation was simply overlooked it is difficult to view a motivation that can pass under our attentive radar as very motivating. Alternatively consider the actions of those awarded the Victoria Cross – whilst there can be little doubt that they were aware they were in danger and felt fear for their safety their behaviour is anything but flight or an act of self-preservation, and it would be difficult to argue that fear for their own safety is what motivates them to place themselves in further danger. Again, anecdotal reports often include references to putting the danger they faced “out of their minds”. So an absence of behaviour normally associated with fear cannot be enough to distinguish real fear from fictional fear.

Walton asserts that what constitutes real fear is this quasi-fear (physiological reactions) plus the belief that one is in danger, and that the belief that something is dangerous is what qualifies it to be the object of that fear; since Charles’s belief that the slime threatens him is fictional so is his
fear. Walton agrees that Charles's quasi-fear could have been generated in a different way, without the fictional belief that he is in danger, but maintains that there is a (fictional) motivational pressure to flee – he is “fictionally inclined to escape the slime, even though he is not actually so inclined”.\textsuperscript{61} This fictional truth is apparently what makes his “psychological role” in the game so realistic. The levels of his fictional fear depend on the levels of his quasi-fear, and Walton suggests that Charles “follows the waxing and waning of his “fictional fear” by looking within himself, much as he would follow the progress of actual fears.”\textsuperscript{62} Whilst I can imagine feeling the physiological effects Walton labels quasi-fear before I realise that they are due to some creeping worry about a situation, and that after so reflecting I come to understand my bodily sensations have been brought about by fear of this situation, it seems peculiar to suggest that an audience member needs to check whether her heart is pounding in her chest or her palms are clammy to determine that the film she is watching is getting progressively more frightening. If anything, it is arguable that as the level of quasi-fear rises, such introspection becomes ever more difficult.

Noël Carroll provides further objections to Walton's position in his \textit{The Philosophy of Horror}; most straightforwardly Carroll asks what emotional state Charles is in if it is not fear, and if for Charles the state of quasi-fear is more or less indistinguishable (both physiologically and in terms of Charles's own reports of his feelings) from real fear, why shouldn't we call it fear?\textsuperscript{63} More importantly, if the realisation that he was experiencing quasi-fear prompts Charles to engage in a game of make-believe, Carroll points out that it is unlikely Charles would be able to give anyone a clear account of the rules of this game.\textsuperscript{64} For Walton, however, such rules can

\textsuperscript{61} Walton, \textit{Mimesis as Make-Believe}, p.246.
\textsuperscript{62} Walton, \textit{Mimesis as Make-Believe}, p.246.
\textsuperscript{63} Carroll, \textit{The Philosophy of Horror}, p. 69.
\textsuperscript{64} Note that Walton's position is that it is the game of make-believe that prompts the quasi-fear, not the other way around. Carroll's point remains, however, if it is entirely unexpected that the slime should
simply be implicit – like a subconscious extension of the rules of his childhood games of make-believe – and Charles therefore need not be able to provide any such account. However, Carroll further suggests that Charles is unaware that he is playing any such game at all, and asks how we can engage in a game of make-believe without choosing to do so – what is there about the way we watch non-fiction that differs from how we watch fiction? We can take this objection further by imagining watching something that may or may not be fiction: the original marketing campaign for *The Blair Witch Project*, for example, claimed that the footage was real, and the opening credits of *Fargo* claim that the events depicted in the film were based on a true story. It is easy to imagine two cinema-goers watching *The Blair Witch Project*, one of whom is under the assumption that what is being depicted is based on fact and the other believes it to be entirely fictional, who have identical reports of their emotional responses on leaving the cinema. Even if Walton’s theory were in some way true, it seems helpful only in terms of its attempt to deal with the paradox of fiction, not for the identification or understanding of emotional responses – if anything, by introducing a whole new realm of emotions and games we are not even aware we are playing Walton’s creation seems to be a cumbersome and unwieldy framework that falls foul of Occam’s razor; Charles too would deny that he was pretending (he would claim that the fear felt real to him) and anyone observing Charles would most likely concur.

Carroll also asks whether we can engage in such pretend emotions at will\(^{65}\) – it seems just as unlikely that I can choose not to be frightened by a good horror film as I can choose to be afraid during an awful one. Although Carroll suggests that Walton may simply have misdescribed the

---

\(^{65}\) NB Walton doesn’t require this, as he allows for spontaneous imaginings and responses in chapter one of *Mimesis as Make-Believe*.
situation, and Charles is in fact fearing for one of the characters in the film, I do not think that argument stands — Walton is clear that no small part of his concern is with the ontological gap between objects within the fiction and objects (people) in the actual world, so his focus on fear for oneself is quite justified. If we were to examine fear for fictional characters, Carroll suggests Walton might insist we can only (genuinely) fear for those we believe to exist, and notes the apparent arbitrariness in this position, pointing out that “a great deal of what we would pretheoretically call emotional responses are to fictions,” and noting that it seems unlikely Walton’s position could hold for all emotions, given that one can be sexually aroused if a suitable subject is described, depicted or even imagined — and it would seem most peculiar to describe this arousal as make-believe. With disgust Carroll is on even surer footing — from stomach-churning stories told over dinner that might make one lose one’s appetite to a scene in Creepers in which a girl falls into a pool of decomposing bodies, there seems to be nothing pretend about our feelings of revulsion and even if there were, they again seem indistinguishable from the real thing, regardless of our knowledge that what we are seeing is fictitious. If disgust does not require existence-beliefs then it seems we must at least grant the possibility that fear can arise without them too; Carroll suggests that if we imagine our hand getting mangled in some sort of industrial machinery we will shudder. Walton might describe this as a physiological response of quasi-fear, but as Carroll notes, his only explanation as to why this must be the case seems to rest on the presumption that it simply cannot be real fear, so it must be something else.

66 Carroll, The Philosophy of Horror, p. 76.
67 Presumably Walton’s position would be that we pretend to fear for the character.
68 Carroll, The Philosophy of Horror, p. 76.
For Carroll there is “nothing amiss”\(^6^9\) with the notion that fictional characters can elicit genuine emotions from an audience. Carroll attempts to explain how what does not exist can affect what does exist, and how this relationship might dissolve the apparent paradox of fiction. Rather than claim what we feel are not genuine emotions, or that we somehow fail to believe that the characters are fictional, he attacks the notion that genuine emotion must be founded on a belief that the object exists. Instead, he suggests that emotions can be generated by entertaining a thought.\(^7^0\) If belief is the assertive entertainment of a proposition, thought is the non-assertive entertainment of a proposition, such as when we feel fear at the thought of falling from a cliff despite not believing that we are about to do so. Carroll holds that we do not fear the event, but the content of our thought.\(^7^1\) This tallies with our strategies for reducing the intensity of emotional responses we have when watching a film, such as distracting our attention from the screen or thinking about something far removed from what is being portrayed. For Carroll, these are attempts to keep ourselves from the thoughts of what is being shown. Since we know they are just thoughts, and not beliefs, not only do we have an explanation for why we do not rush from the cinema or call the police, we can also avoid the charge of irrationality from having responses to mere thoughts. Indeed, Carroll goes as far as to suggest that it may have been our capacity to generate emotions from thought-contents that led to the creation of fiction in the first place.\(^7^2\) There seems to be nothing to recommend quasi-fear over genuine fear arising from thought-contents; however, since Carroll's theory allows us to feel genuine emotions towards fictions, it has more to recommend it than Walton's.

---

\(^7^0\) Carroll, *The Philosophy of Horror*, p. 80.
\(^7^1\) Carroll, *The Philosophy of Horror*, p. 80.
\(^7^2\) Carroll, *The Philosophy of Horror*, p. 83.
Alex Neill believes that we can have genuine emotional responses to fictions and in “Fiction and the Emotions” he asks whether we can feel real pity for a fictional character. Like Radford, Neill also dismisses the notion that I pity real people that are ‘brought to mind’ by the fiction. Neill also suggests that describing my response as a ‘state of feeling’ or an imaginary or make-believe emotion does not address the issue that if we cannot believe things to be true of a fictional character (“Emma Woodhouse was handsome, clever and rich”) because the character does not exist, it seems we cannot be said to have properly understood the work. To deal with this apparent paradox, we can add the qualifying term ‘in the fiction’ – “while it is not true that Winnie [Verloc] had a miserable time, what is true is that it is fictional that she did”.73 If we can thus hold beliefs about fictional characters, our responses might be genuine emotions. Neill goes on to dismiss the suggestion that these are “insincere” or “putative” emotions, as the belief that (it is fictional that) Emma Woodhouse had those qualities is genuine. It is not the beliefs that are fictional, rather it is the contents of the beliefs - the reader believes that, in the fiction, Emma is handsome. The problem may be that we can only pity objects that actually exist – we may believe that (it is fictional that) Winnie suffered, but this does not allow us to pity Winnie as she does not exist; that “beliefs about what is fictionally the case lack the causal power to move us”.74 Neill denies this, asserting that Boruah’s position that beliefs about fictions are “mere recognitions” that something is the case lacks force, as we can construe a belief in a real situation as a “mere recognition” that something is the case and neither is obviously more effective as a cause of emotion.

Neill suggests that appreciators adopt a ‘perspective’ - “seeing things from another’s point of view”.75 a belief concerning actual people and events has emotional causal power because I can

74 Neill, p. 3.
75 Neill, p. 3.
see (to at least an extent) what being in their position might be like. Why can we not similarly adopt the point of view of a fictional character? Neill notes that an author's ability to successfully show us their character's perspective (or their characters' perspectives) on the fictional world is often a standard of evaluation of the work, and that we do not consider the rabid dog's perspective when it charges toward us, nor the monster's when we watch a horror movie – “not all emotions ... stem from adopting another's perspective”. We must therefore be aware that it may not be the case that all emotions can be treated similarly – two examples are fear for oneself and jealousy, as we do not (or at least are not licensed by the author to) believe that it is true within the fiction that we ourselves are endangered and we cannot believe that something a fictional character has is rightfully ours.

Rather than adopting someone else's perspective, both emotions rely on the subject “standing in a certain sort of relation to the object of the response, a relation that cannot obtain between the inhabitants of different ontological “worlds””. We should take care to note, however, that interactive media can provide a causal relation between these worlds; this is a key difference between interactive and non-interactive fictions, and, I shall argue in the following chapter, this relation will alter the range of emotional responses it is possible for an interactive work to elicit.

Neill goes on to look for necessary or characteristic features of pity that would be missing when considering fictional scenarios. Given we can clearly have bodily reactions to fictions, they do not depend on believing the events to be actual; Neill does concede that these reactions are often quite different from those resulting from actual situations, but notes that responses to fiction may well be more intense than responses to certain actual events. Moving on to distress

---

76 Neill, p. 4.
77 Neill, p. 4.
as a feature of pity, Neill suggests that perspective offers a solution to how it might be possible to feel positively about feeling negative emotions (e.g. enjoying a tragedy) by pointing out that “our attention may have more than one object”\(^78\) so the object toward which we feel positively may not be the same object to which we feel negatively – we may feel sad about the events portrayed but happy about how they were portrayed, the quality of the performance.

Desire is the last characteristic of pity Neill examines, suggesting that it is too simplistic simply to claim that in actual events we desire to prevent or halt the suffering, and that in fictional events we lack this desire altogether because we know that any suffering we see is purely fictional. When watching a performance of *Romeo & Juliet*, we might desire that Mercutio’s death is avoided, while at the same time desiring that the play follow the original plot, i.e. we desire that Mercutio die and not die. Neill goes on to describe how our desire for Mercutio to die might better be explained as being aware that, should he not die, we would not be watching a performance of *Romeo & Juliet* but something else, and points out that we can have conflicting desires about actual situations also (e.g. suffering through bereavement as a necessary step towards recovery). Neill then considers a different side to the desire component of pity – the desire to take action. This would be problematic if it were a necessary component, as we do not generally desire to take action regarding a fictional plight, and Neill suggests this may even be impossible due to the “ontological gap” – not only would it be logically impossible, even if we could “we know nothing we could possibly do would *count* as helping”.\(^79\) If I can easily end someone’s suffering but I am not motivated to do so it seems reasonable to conclude that I do not pity him. However, this desire cannot be a necessary component in all cases of pity, as we have seen in the submariners example above: watching reports about the plight of the crew of

\(^78\) Neill, p. 6.
\(^79\) Neill, p. 8.
the Kursk, whether live on the news or months later in a documentary, we are equally unable to help but we would surely characterise our emotions towards them as pity. We may well desire that there were an action we could take (or could have taken), just as we could desire that there were some action we could take to assist a fictional character. Neill suggests that we might believe that it is fictional that Anna Karenina existed and suffered, and we might desire that fictionally her situation had been different such that she did not suffer, and that this desire is in no way fictional – we have actual desires about fictional objects. Between our beliefs and our desires, then, we have satisfied the requirements of feeling pity, even for fictional characters. However, there is now the possibility that, if we wish that the outcome of a fiction were different, we in effect wish the fiction were differently written. Neill tackles this by suggesting that a desire that fictionally there had been a different outcome is not the same as desiring that the plot had actually been written differently, remarking “my desire to lose weight is not in effect a desire to stop eating iced buns; I may desire to lose weight without having any desires with regard to iced buns at all”.\textsuperscript{80} A desire for Anna Karenina not to suffer the fate she does is not inconsistent with the lack of desire that the novel itself had been written differently; we can pity Anna but recognise that a happy ending would have resulted in a work of lower literary quality. Neill returns to his notion of adopting different perspectives when responding to a fiction – when I focus on the story aspect of \textit{Anna Karenina} I have no desires concerning other aspects of the novel such as plot, pacing or language, likewise when I focus on it as a novel I do not (or am unlikely to) have desires concerning Anna. If this is the case, and it seems reasonable to grant that it is, we do not have to face the dilemma of wanting the same story with a different outcome. This also explains why our responses are often much shorter to fictions than to actual events – after the performance/novel is over, or even multiple times during it, we shift perspective and our attention switches with it; indeed “the ability to make this [shifting] an

\textsuperscript{80} Neill, p. 10.
integral part of our experience of the work rather than an annoying distraction is one criterion of mastery of the art of fiction”.  

3.3 The Rationality of Fiction-directed Emotions

Colin Radford allows that we may feel genuine emotions towards fictional characters and situations, but claims that we are, in at least some sense, irrational to do so. Returning to his example outlined above in which a friend tells us of the unpleasant happenings in his sister’s life only to reveal he in fact has no sister, Radford may well be correct to claim that in this case any feelings we might have for the sister would vanish. However, it seems to me that the fact we learn that we have been duped is a critical factor in this example, and one which is absent in most cases of engaging with fictions. Radford states that “the possibility of your being harrowed again seems to require that you believe that someone suffered”, but if I imagine a horrid fate befalling someone real, but about whom I know nothing else of any substance – a 'blank' or generic person – I can find what I imagine deeply unpleasant. It is not clear how imagining the same situation with a 'blank' or generic person about whose existence I have no beliefs would fail to give rise to the same emotions (or something indistinguishable from them), or how imagining a fictional person undergoing the same trauma could fail to do so, and indeed Radford goes on to allow that “[w]e shed real tears for Mercutio”. His concern rather is to explain why we might feel for characters that do not exist.

---

81 Neill, p. 11.
82 Radford, p. 68.
83 Radford, p. 70.
In “Rational Fear of Monsters” Richard Joyce outlines a defence contra Radford of the rationality of feeling emotions towards fictions, in which he distinguishes between the paradox of the possibility of emotional responses to fiction:

i. We fear for ourselves only if we believe ourselves to be in danger; we fear for others only if we believe they actually exist and are in danger.

ii. When we watch a horror movie, we do not believe ourselves, or anyone actual, to be in danger.

iii. We are sometimes frightened when watching a horror movie.

and the paradox of the rationality of emotional responses to fiction:

iv. We are not rationally frightened unless we believe someone actual to be in danger.

v. We do not believe anyone actual to be in danger when we watch a horror movie.

vi. We are sometimes rationally frightened when we watch a horror movie.

noting that Radford's approach is to deny (vi). Joyce attempts to show that such reactions are rational inasmuch as they are desirable or instrumentally valuable, building on Lamarque's 'thought theory' in which we are frightened by our thoughts.

Like Walton, Lamarque believes that the audience engages in an act of make-believe when watching or reading a fiction, and they do this, at least in part, by holding relevant appropriate thoughts in their minds such as 'the girl is in danger'. As this is make-believe the audience does not believe the propositional content of those thoughts, but with "adequate imaginative attention,
the thought may produce the emotion of fear – real fear”. Lamarque distinguishes between being scared by something and being scared of something, and between thoughts as 'brain events', which have causal properties such as prompting fear, and thoughts as 'representations'; in this way we are scared by the thought as a brain event, we are scared of the thought-content.

Joyce provides a reply to Radford's question “Why is it false to say that it is because the moving picture of the slime looks so real and horrible that, when it suddenly appears, I involuntarily flinch, recoil, thrill in fear and horror?” by suggesting that we are frightened by both the image and the thought – fear may be caused by a combination of things and it seems reasonable to place importance with both the image of the slime on the screen and the thought in the appreciator's mind. With its focus on the thought Lamarque's theory can easily be applied to film, literature and even to daydreaming, as there will always be a thought involved in any emotional response therein.

Joyce also notes that in (i) a belief is required in order for an emotion to be present – it is part of what constitutes an emotion – whereas in Radford's (iv) the presence of a relevant belief is simply what makes the emotion rational; if emotions can arise without a belief this is a departure from the 'judgementalist' version of the cognitive-evaluative model of emotions, i.e. that our emotions result from forming judgements or beliefs about something (Greenspan’s example of Frances’s fear of Fido provides a case where there is no judgement that one is in danger - Frances lacks the relevant belief, so by (iv) her emotional response of fear is irrational). Since Radford does not provide an alternative account of what an emotion might be considered to be, Joyce suggests modifying the standard cognitive-evaluative theory of emotions. Taking Carroll's position that experiencing an emotion involves having certain physiological responses

---

84 Joyce, p. 211.  
85 Radford (1982), p 261 quoted in Joyce, p. 211.
in addition to “an evaluative belief that p … or the thought that p”,\textsuperscript{86} i.e. the thought is one part of the emotion and the physiological responses caused by the thought are the other, Joyce notes that this makes the relation between thought and emotion constitutive rather than causal and in doing so we can no longer simply say we are frightened \textit{by} thoughts. However, he objects to “[t]he ‘frightened by’/’frightened of’ locution” itself, pointing out that there are not similar relations for emotional responses such as ‘pity for’, and arguing that whatever emotion we feel the thought-contents remain “the most salient explanatory item”.\textsuperscript{87}

Joyce attacks Radford’s comparison of emotional responses to fiction to children scared by stuffed tigers or the arachnophobe’s fear of a spider they know to be harmless, noting firstly that we do not generally label emotional responses to fictions silly or irrational as we do the other cases. Whilst he concedes that ‘generally being the case’ is insufficient to show something’s rationality, it is enough for it to be reasonable to question the validity of these analogies and Joyce points out that, in being able to call the arachnophobe irrational, we display some capability to differentiate between rational and irrational emotions – if someone were to remain distressed and frightened days after watching a horror film we might well consider that irrational, and if we had no response whatsoever we might at the very least be considered unusual. Joyce suggests that the reason we call the arachnophobe’s fear irrational may not be because a certain belief is missing – the belief may even be present – instead, we do so because “their emotion is obsessive: it is beyond their control and self-destructive”.\textsuperscript{88} If we are more in control of emotions experienced through fictions Joyce believes we should credit them with rationality according to their instrumental value:

\textsuperscript{86} Joyce, p. 213, original emphasis.
\textsuperscript{87} Joyce, p. 214.
\textsuperscript{88} Joyce, p. 215.
A person has interests, and if there is something she can do to help satisfy those interests, we may say that she has an 'objective reason' to do that thing. When she is justified in believing herself to have an objective reason, then she has a subjective reason. She is rational insofar as her actions are guided by her subjective reasons.89

Joyce believes emotional responses to fictions to have great value when they are under our control, and suggests that if this is the case they might be considered rational. In order to consider their instrumental value emotional responses need to be thought of in terms of actions; not the emotion itself but the having of that emotion is what is being examined. If emotions required beliefs they could not be considered actions inasmuch as we cannot control the beliefs we form that provide the foundation for emotion. However, Joyce's suggestion was that we allow emotions to be caused by thoughts, and it is much easier to see how we can choose to think about something, or not, and how we might control the strength or power of that thought. For Joyce, “entertaining thoughts of Anna or Dracula are as much actions, are as much subject to the norms of practical rationality, as entertaining guests for dinner”.90 We might feel more or less than we are expecting to at a given film or reading a particular book, but the strength of the feeling is unlikely to be very much greater than I would normally expect from watching films or reading books in general, and that this is outwith our control may be due to the physiological responses from entertaining the thought being beyond our control. Although it is easy to imagine situations in which fictions overwhelm us emotionally, Joyce believes that the majority of the time we are in fact in control of our emotional states – that we allow the stirring music or

89 Joyce, p. 215.
90 Joyce, p. 217.
the unfolding drama to draw emotions from us, and usually we can quell them more or less at will; we label as irrational those people who wail and howl uncontrollably. Joyce presents us with an example of two movie-goers to make clear what he means by 'control': the two audience members watch the same film and have the same outward behavioural responses (a few tears, say), after the movie the first reports that she was, indeed, a little sad but the second reveals that he was overwhelmed by sadness and that he cannot shake the feeling even now the film has ended. Joyce suggests that we would call the second movie-goer's responses 'out of control', and that this “creates a strong presumption in favour of the claim that the ordinary cases of pity, sadness, and so on, felt in response to fiction are not uncontrolled”.\(^91\) Note that, should our emotional responses far exceed our expectations, we have not necessarily become irrational: returning to the dinner party analogy, Joyce suggests that such cases are similar to inviting a risky guest who we justifiably believe is often fantastic company but occasionally very rude. If it is rational to invite her, it is no less rational should she behave badly in the event.

Joyce notes that this case involves another's agency, the guest's, and suggests the following principle for cases that involve only one agent:

\[\text{vii. If a person performs action } \Phi, \text{ believing that } \Psi \text{ will likely result – and no other agency is involved – then, if } \Psi \text{ does result, that person is deemed not merely responsible for } \Psi, \text{ but to have performed the action } \Psi.\]\(^92\)

If I turn the oven up knowing the cake inside is likely to burn, and it does, then I am deemed to have performed the action of burning the cake. If we apply this principle to emotions, we choose both the thought we entertain and the physiological responses we expect this to elicit;

\(^91\) Joyce, p. 218, original emphasis.
\(^92\) Joyce, p. 218.
we can view the whole emotion therefore as an action, and as such judge its rationality on practical terms. As previously noted, where emotions are founded on beliefs we have less, and possibly no, control, and these emotions cannot be judged actions. However, Joyce contrasts these cases with sitting and imagining in as much depth and detail the death of a loved one, which seems much more like an action – and one which would be deemed irrational “if the whole process fails to serve one's purposes”. However, this is directly contradictory to Joyce's dinner party example; an actor imagining the death of his father in order to generate suitable physiological responses for the scene he is playing seems to be acting rationally whether or not his imaginings succeed in enhancing his performance. If the scene were meant to be cheery and the actor believed his imaginary scenario would make him sad he would certainly be acting irrationally, but Joyce’s claim that mere failure to achieve one’s objective makes one irrational should not be the test here - it should instead be the reasonableness of one’s expectations that the action should deliver one’s goal, and indeed this is borne out in his later (viii).

Applying this principle to beliefs also causes Joyce some problems: if $\Phi$ is the taking of a pill that would cause $\Psi$, that one would hold a false belief, Joyce states that it is undesirable that believing the false belief should be considered an action. This reluctance is surely uncalled for – we did not want to consider beliefs actions because they were largely outside of our control but in this example we have an unusual measure of control over belief-formation, so an unusual conclusion - that the holding of this belief is an action - should not be so surprising. Indeed, insofar as actions require some measure of choice, we have chosen to undertake the action that will lead to holding the false belief, so by (vii) Joyce should be committed to holding that, at least in this scenario, holding a belief can be considered an action. The far greater problem for Joyce

93 Joyce, pp. 218-9.
94 Joyce, p. 219.
is that, regardless of the instrumental value of holding the false belief, the belief itself cannot be said to be rational. So whilst Joyce has shown that we may be instrumentally rational if we choose to have emotional responses to fiction, this does not allay Radford's chief concern, that such emotions are nonetheless themselves irrational.

Joyce moves on to consider how we might determine an agent's rationality in performing an action, and suggests the following principle will help distinguish between two senses of one's rationality:

viii. If a person, P, performs action α, justifiably believing that β will likely result, and justifiably believing that β will serve his ends, then (I) α is a rational action, and therefore (ii) P is \((\textit{ceteris paribus})\) rational.\(^{95}\)

If α is the taking of the belief-forming pill and β is the formation of the factually false belief, P is rational in one sense and irrational in another; applying this to emotional responses to fictions, if we believe that experiencing certain emotions will be of some sort of benefit to us and we further believe we can take action to provoke those emotions (i.e. engaging with a fiction) we cannot be said to be straightforwardly irrational for taking that action. Note, however, that the emotion itself remains irrational.

This proposal admits of other criteria by which P may be considered irrational – this is most plain in the case of beliefs: since we form beliefs to help us correctly interpret the world around us by accruing accurate data, forming beliefs in stark opposition to available information is at the very least counter-productive, and at worst dangerous.

\(^{95}\) Joyce, p. 219.
We can also look for other criteria for rationality when $\beta$ is an emotion. Radford's condition (iv) states that we must believe someone actual to be in danger to rationally fear for him, and is drawn from examining the irrational fear of the arachnophobe who believes the spider in front of him is not dangerous. Joyce correctly points out that there may be a sense in which the arachnophobe believes the spider is dangerous, just as someone afraid of flying might well believe the aeroplane will crash; it is their irrational beliefs that cause them to be called irrational themselves.\textsuperscript{96} Joyce allows that an arachnophobe might confirm that he knows the spider is not dangerous, but argues that this scenario may not be analogous to other phobias: “[d]oes the taphephobe know that he is not going to be buried alive, or does he irrationally believe that he is? Does the pathophobe know that disease is not dangerous, or does she irrationally believe that she’s about to fall ill?”\textsuperscript{97} and if this is the case it seems not all phobias are irrational due to failing (iv). Joyce suggests that lack of control and counter-productiveness might be better criteria by which to judge a response's rationality: if we fear an actual object which actually endangers us it may be out of our control but it is productive in that it helps keep us safe. If the actual object does not endanger us and we uncontrollably fear it (the phobic response) the fear is counter-productive, either in that it does not succeed in keeping us safe, or that it wastes energy or prevents us from exploiting a potential food source. Joyce concedes he cannot explain the 'purpose' of emotions, but this position does have intuitive credibility in the case of fear, anger (the protection of oneself or one's family, property, rights etc.) and even envy could serve such purposes as encouraging competition or drive. It may be harder to find acceptable 'purposes' for emotions such as sadness, pity and grief, but intuitively it seems we should still be able to identify situations in which they are out of control and counter-productive as irrational. For emotions that are counter-productive but under our control, whether we will be rational in

\textsuperscript{96} Joyce, p. 221.
\textsuperscript{97} Joyce, p. 221.
feeling them is determined by whether we are justified in believing that feeling them will be of instrumental value.

This instrumental value may simply be the pleasure of reading a good book, but the cognitivist view of art places value in art's ability to further our understanding of ourselves, others and the wider world, and Joyce suggests that our comprehension of a work is developed through/ aided by emotional engagement, arguing that had Tolstoy simply written 'Passion can be destructive' instead of *Anna Karenina* we would have been denied the opportunity to gain the experience the novels provides and we would in all probability fail to learn the lesson:

Merely being told that a certain kind of relationship will end in disaster has little impact on me. Watching a close friend go through the whole drama brings it home rather more. Going through it all myself, experiencing the hope and agony, will, if I am sensible, teach me not to do it all again in the future.\(^98\)

Joyce also discusses Susan Feagin's theory of the value of tragedy, in which our pleasure viewing a tragedy is described as a "meta-response"\(^99\) in which satisfaction is derived from feeling the appropriate emotional response to the fiction. There may also be satisfaction in being able to experience fear under control, perhaps to learn to better control fear in the future. Presumably different appreciators may have diverging opinions on what an appropriate response is, and they may differ from the response the artist intended to be appropriate; the responses to the original edit of Ed Norton's brutal encounter with Jared Leto in David Fincher's *Fight Club*, for example, were whoops of encouragement and excitement. Fincher re-edited the

---

\(^98\) Joyce, p. 223.

\(^99\) Susan L. Feagin, 'The Pleasures of Tragedy', quoted in Joyce, p. 224.
scene and included reaction shots of members of the crowd rapidly becoming first disconcerted and then appalled at the beating, and the audience responded in similar fashion. Perhaps we are satisfied no matter what our emotional response, just so long as we believe it to be appropriate. Joyce concludes that through practice, reflection and exposure to a number of varying situations fictions allow the spectator the possibility of moral growth, suggesting that a theory “that entails that such a person would be irrational in purposely pursuing such highly valuable ends is an odd one indeed”.

Although Joyce has demonstrated that we may have good reasons for wishing to be moved by fictions we are still faced with the concern that, regardless of their benefits or instrumental rationality, our responses might nevertheless be irrational in and of themselves. There should be little doubt that Charles’s fear of the slime is indeed irrational; it is certainly unwarranted. We can say that it is unwarranted because it is not true within the fiction that Charles (or the audience generally) is endangered by the slime or, to frame the situation in more Waltonian terms, Charles is making-believe in a way that is not prescribed by the author or by our general conventions and fictive practices - with most films we do not make-believe that we ourselves are a part of the events on the screen, that we occupy the position of the camera or that we are one of the depicted characters. If we were to do so we would be faced with problems similar to those of the implicit internal narrator - why does nobody notice our presence, how do we teleport from place to place or backwards and forwards in time, how do we float in mid-air, shrink to microscopic size or enter the dreams or thoughts of other characters and so on. Regardless of the philosophical debate regarding narrators we should be able to agree that the reason audiences do not ask these questions every time they visit the cinema is that it is not our

100 *Fight Club*, 1999, Director’s commentary.
101 Joyce, p. 224.
general fictive practice to make-believe that we ourselves are a part of the fictional world depicted, and if Charles does so then his "intrusion into the work-world of the film",\textsuperscript{102} as Stephen Davies describes it in his "Responding Emotionally to Fictions", is "inappropriate".\textsuperscript{103} Although we may be authorised to place ourselves within the game-world it is fair to say that in most cases this is a departure from what is intended or usually prescribed. For the purposes of aesthetic appreciation, Davies points out that such imagining or making-believe:

\begin{quote}
would lead one to miss the work's narrative structure as a whole and the place of particular events within this, and to miss also the affective relevance of the attitudes expressed in the work to its contents, since it is only those who are external to [the fictional world] who can acknowledge these.\textsuperscript{104}
\end{quote}

Building on this, we can say that it is not true within the fiction that the events described or depicted happen in such a way as to provide instrumental value to others, whether in terms of a 'moral' or as a basis for reflecting upon in order to gain insight into 'the human condition' or some political or philosophical point of view, or for the purpose of entertainment, whether in terms of comedic coincidences or misunderstandings or simply a satisfying story-arc. If Charles does not make-believe that he is part of the work-world and yet still describes himself as feeling fear for himself, he has mistaken either fear-for-a-character or the startle response for fear-for-himself; if Charles were to deny that he is mistaken then his response echoes Joyce's depiction

\textsuperscript{102} Davies, "Responding Emotionally to Fictions", p. 277.
\textsuperscript{103} For Walton it is permissible for Charles's game-world to include himself and a fear at least of the representation of slime, but as Davies notes (p. 278) Charles's fear of the slime is "fictional within the game-world", not actual.
\textsuperscript{104} Davies, p. 273.
of the second of his two movie-goers whose emotions are out of his control and, as such, irrational or "not cognitively warranted".\textsuperscript{105}

If we are as separated from Anna Karenina’s plight as Charles is from the threat of the slime, can our feelings of pity for her be cognitively warranted? Davies makes a convincing case that we can, by drawing on emotions that concern the real world, but whose objects do not currently exist or cannot be influenced or altered.\textsuperscript{106} Our feelings of regret and remorse combine recollections of events which have already happened with counter-factual scenarios that we would rather have taken place. We use our imagination to picture predicaments of which we have no direct knowledge (just what was happening on board the Kursk, to recall our earlier example) and we imagine what we or others could have done differently and the improved outcome that might have resulted. Looking to the future we may feel hope or foreboding, “emotions that target what does not exist and may never happen”,\textsuperscript{107} and again it is our imaginings that are at the centre. Though we are unable to interact with these imaginings we nonetheless feel emotions towards them. Although it must be acknowledged that wrapped up in these emotions there may well be beliefs about what could happen or could have happened, the scenarios are imaginary and as such fictitious, and if we imagine situations in the distant future, say for example millions of years hence when the Sun begins to swell and incinerates the Earth, we may believe that humanity will by then either have moved to another planet or become extinct, but still be able to imagine people dying never having enjoyed lush forests and plentiful water and pity them. With regard to Anna Karenina then, we can imagine Anna and the events that befall her, and our imaginings cause us to have emotional responses. We can say that we correctly imagine when our thoughts are in line with what is true in the fiction, or that we are

\textsuperscript{105} Davies, p. 281.
\textsuperscript{106} Davies, p. 270.
\textsuperscript{107} Davies p. 270.
prescribed to make-believe what is true in the fiction; if we were to use the novel as a prop in a game where a poor girl inherits a fortune and lives happily ever after, or we imagine that Anna does not die but goes on to lead a contented and satisfied life, we would have incorrectly imagined at least in terms of our general fictive practices.

3.4 Conclusions

I have argued that we routinely have emotional responses to imagined scenarios, whether they are real events from which we are geographically or temporally separated, fictional events that might or might not occur to real objects at some point in the future (or even to objects that only might exist at that point in time). We also routinely have responses to purely fictional situations involving characters we know to be fictional, which are often described as genuine emotions. Although describing responses in the last group as ‘quasi-emotions’ is suitable for capturing how we might feel fear or other emotions which have a distinct accompanying physiological response, when applied to emotions without such physiological reactions I have suggested they are indistinguishable from genuine emotions and as such the distinction does not appear to be helpful or informative - this is perhaps clearest when we consider hope and amusement: when we hope for a happy outcome we appear to be engaged in a game of make-believe that uses facts about the real world as props for imagining the future, and when we watch or read a comedic film or novel we play a game in which something funny happens. In the former case it would appear that hope, being founded on our imaginings, is always and only a quasi-emotion, in the latter our emotional response (and, potentially, the physiological response of laughter also) seems not to differ in any significantly appreciable way\textsuperscript{108} from the responses we would have if the events happened in the real world. I can see no benefit to labelling certain

\textsuperscript{108} Although the magnitude of the response might differ, it would not always be stronger or weaker in either an imaginary world or the real world, so it is fair to disregard this potential difference.
responses quasi-emotions beyond the worry that they are irrational but, as I hope to have shown, there is no hindrance to our holding beliefs about what is true within the fiction, nor responding emotionally to imaginings. When our emotional responses are based on what we have imagined, they are rational if our imaginings are correct, which is to say we have imagined what is true in the fiction, and if they remain broadly under our control. Even if they go far enough beyond our control to be described as irrational, there may be good instrumental reason for choosing to experience them (although as I have noted this does nothing to change the response’s irrationality). Fear for oneself in relation to imagining a possible real-world future is straightforwardly rational (again, assuming the fear does not escape one’s control and - importantly - does not ‘spill over’ into the real world to an unwarranted degree\(^{109}\)) and may serve considerable instrumental purpose, but fear for oneself in relation to imagining a fictional world can only be rational if it is true within the fiction that one is in danger. There are comparatively few non-interactive works in which it is fictionally the case that you are part of fictional world, and the limitations of the media are such that it is probably fair to say that most attempts at this fail to elicit a significant response in audience members - point-of-view shots in cinema are rarely, if ever, convincing, and, from my own experience at least, novels written in the second-person such as Tolstoy’s *Sebastopol in December* are no more powerful as props than those written in first- or third-person. Even in these cases, however, we should note that for the fear for oneself to be rational, it should be contained within the “it is fictionally the case...” locution.

\(^{109}\) Our fear that is prompted by imagining the approaching car hitting us should stop us stepping out in front of it, but it should not prevent us from crossing the road once the car has passed; imagining a terrorist attack in Baghdad should not dissuade us from holidaying in Torquay, although it might rationally prevent us travelling to Iraq.
Chapter Four - Emotion in Interactive Media

That we have emotional responses of some sort to works of art is uncontroversial. Often the artist intends audiences to feel a particular emotion or range of emotions when appreciating a piece, in order simply to add to the overall experience, to convey her feelings about what is being represented or to guide the audience to a particular point of view about a subject, and in some cases the appreciator having an emotional response is the very purpose of the work. Where a particular emotional response is desired, if an appreciator feels a different emotion (or even no emotion at all) it may be argued either that the artist and his work have failed or that the appreciator has failed to understand the piece properly. There are two areas where I believe interactive works provide particular interest in terms of the emotional responses they can elicit, however: they should be able to provoke any of the emotional responses we might expect from a non-interactive work, but they have at least the potential to amplify those responses by drawing the audience into the work more directly than non-interactive works often can; I will also argue that some interactive works can elicit certain emotional responses that I have argued are typically beyond the reach of non-interactive art.

4.1 Non-Narrative Works

We will begin by examining interactive works with little or no narrative content; for convenience I shall refer to these works as non-narrative, but it should be noted I
include in this group works which have a minimal narrative content. This narrative may be in terms of something pictorially represented which changes over time or when users are in a particular relation to it, musical works which it might be argued express some sense of narrative, or works liable to prompt audiences or users to construct their own narrative using the work as a prop or ‘jumping-off point’. It should also be noted that works of higher narrativity share the properties of ‘non-narrative’ works in this context which allow for increased user responses, but they have additional features which will be examined separately later in this chapter.

One example of an interactive work which might be argued to elicit stronger emotional responses than a similar, non-interactive piece is Scott Snibbe’s *Boundary Functions* (1998). The work comprises a platform above which is mounted a camera and projector which are connected to a (hidden) computer. As members of the audience walk onto the platform the computer uses the camera to note their positions and uses the projector to display lines between them that divides the space between them. These lines are dynamic, moving and changing as audience members move around the platform, and as participants step on to or off the platform lines are added or removed as necessary. Each participant has their own space, the area of which grows or shrinks as the number and positions of participants changes, and each participant is cut off and isolated from the others. We can imagine a non-interactive version of the piece, *Observed Boundary Functions*: an image or video of individuals separated into cells in the same sort of way; even if this work were to provoke the same sort of feeling of isolation it is likely that, by directly involving members of its audience - actually isolating
them in at least a certain sense - *Boundary Functions* will elicit stronger personal responses. This is because of differences in the ways in which the audience (or users, for the interactive version) relate to each work. In the non-interactive version, the audience might be moved if they imagine themselves in the place of one of the depicted participants. There are two ways in which this imagining could take place: an audience member might make-believe that she herself was part of the work, replacing one of the other people depicted, or she could imagine actually being one of the depicted people - of having the observable properties of one of the individuals. The imaginative leap in this second case may be greater or lesser depending on who is shown and in whose place she imagines being - a similarly-aged, short, slim woman or an older, tall and heavy-set man. Imagining what it would be like to be trapped in a such cell, in either sense, she might further imagine what desires and feelings she would have. If she is imagining actually being one of the depicted people the desires and feelings she goes on to imagine will be informed by the actions of that person and the emotional state she believes him or her to be in - if his arms are outstretched as he tries to bridge the boundary between him and a child (which we could reasonably believe to be his) and his face is contorted in a way consistent with those experiencing pain and loss, our audience member will imagine herself having the properties of having a child, of wishing to be with that child and of feeling great sadness at her inability to be so\(^\text{100}\). If she imagines herself within the work, she is likely (at least initially) to imagine sharing the

\(^{100}\) Whilst a discussion of the neurological processes involved in ‘mirroring’ the feelings of others we observe lies far beyond the scope of this thesis, it is a commonplace enough occurrence that (without additional reasons to do otherwise) we smile when we see others smiling, wince when we see someone slip and fall, and otherwise share in at least a limited way in the emotions of those around us.
emotional state of those depicted as we are psychologically predisposed to ‘mirror’ the states of those around us, but she need not - she may equally well imagine that she is much happier with the situation than those depicted. Even if she begins by imagining herself experiencing similar emotions as the others she might go on to imagine herself having very different responses, either purposefully (perhaps asking herself “What would it be like if I were happy to be isolated?”) or through discovery (noticing that although she attempted to imagine feeling sad, she in fact observed a feeling of peace and contentedness).

In the interactive version of the work however, our audience member does not need to imagine being isolated, she actually is isolated, and rather than imagining what her emotional responses might be she is able simply to observe them. Certainly imagination still plays a significant role in the interactive version - regardless of the lines on the floor there is no physical impediment to her reaching her children or communicating freely with other users or non-user audience members that are observing from off the platform. However, by being placed within the work she experiences it herself and not solely through an imaginative act, and her emotional responses come directly from that experience without being filtered by what she chooses (or indeed happens) to imagine.

This is, of course, not to say that all interactive works will generate stronger responses simply because they have interactive features - Sam Taylor-Wood’s Still Life (2001), for example, is a non-interactive digital piece in which a classically-composed bowl of fruit
is displayed on a video screen and, through time-lapse recording, decays in front of the spectator’s eyes; we can imagine an interactive version of this work in which a detector is used to determine the user’s proximity to the screen and as he approaches the video spools forward; there is no apparent reason why adding in this interactivity would amplify the emotional responses of its audience - it might well detract from them, since they are less able to see the earlier states of the fruit or the way in which the decay progresses. Returning to our imagined Observed Boundary Functions, it is clear that this is not the only way a non-interactive piece might convey a sense of isolation and other works might be more effective than both our imagined piece and Snibbe’s original. I contend, however, that by involving audience members directly interactive works have an immediacy and personalness that can be difficult for non-interactive works to achieve. To flesh this out a little further, we can say that at least some non-interactive works require an intermediate step absent from similar interactive works - in our static version of Boundary Functions, the audience member views the work and sees that the individuals portrayed are separated from one another, imagines herself being in the work (in either mode) and being similarly isolated, and comes to feel a sense of isolation. In the interactive version of the work it is the participant herself who is separated, and as she realises she is cut off from the other participants she experiences this sense of isolation. Indeed, for some participants the feeling of isolation may even arise before they come to appreciate just why this feeling has been prompted.

It should be noted that changing the work in this way may cause different responses altogether; by linking the participant’s proximity to the extent of the decay he might feel a sense of responsibility for the decay that would be absent from the original work, or perhaps a sense of helplessness arising from the notion that, no matter how quickly he attempts to get close to the fruit, it will have decayed when he reaches it.
However, it is important to be clear that immediacy and personalness do not equate to simple magnitude. When our audience member views *Observed Boundary Functions* she might imagine herself actually being in the work, but as noted above she might instead imagine being one of the people within it. By imagining that she has children and is separated from them, and that she shares the apparent emotional state of the man she imagines being, it is reasonable to believe that her emotional reaction will be stronger than if she were to imagine actually being within the work. Of course, she may have a strong predisposition against children that reduces her emotional response but we should be confident in asserting that *ceteris paribus* having a representation of the physiology of the emotional state of the depicted character will assist her in more fully mirroring imaginatively the emotions he appears to be feeling. Her participation in the interactive version would make it harder to construct these imaginings as she would have to disregard her own reactions in favour of those arising from her imagination - she would have to imagine that she was no longer involved with the work, but her imaginary counterpart was. In addition she would not have the depicted man’s observable anguish and suffering to use as a prop, she would have to rely solely on what she imagined she might feel. We can therefore say generally that by actually involving us in the work, some interactive pieces raise obstacles to imagining being someone else experiencing the work that are absent from non-interactive pieces, and by removing some of the prompts we might use imaginatively they might in fact limit the magnitude to which we feel emotional responses that are less relevant to our real-world situation.
The feeling of guilt can be elicited by some works, but as we have already noted, in narrative non-interactive works the only way to provide a rational account for feeling guilty is through some sort of abstraction - we should not feel guilty for the way one character treats another so feelings of guilt must arise via some additional mechanism, most likely that we are reminded of having treated someone in a similar way, or having failed to act to prevent a similar fate. This sort of abstraction must also be present in non-narrative works, whether or not they are interactive. As an example, let us recall the previously-discussed *Panda Eyes* (2010) by Jason Bruges, which involves its users by means of the rotating models of pandas which turn to face them. A similarly themed non-interactive work might convey the same ideas of responsibility and failure to act to prevent extinction and endangerment of species: an appreciator must understand what the work is trying to convey, and then reflect on what they have or haven’t done to help the situation. If he feels that he should have done more, he might feel guilty in response to the piece. The interactive nature of *Panda Eyes* provides the piece with additional resources to reinforce its emotional impact. Alfred Leete’s famous recruitment posters featuring Lord Kitchener are a well-known example of a work whose central character’s eyes appear to follow appreciators as they move, and is comparable to *Panda Eyes* in at least that respect. However, this illusion is essentially lost when one breaks eye-contact with the depiction of Kitchener, and it seems unlikely that viewers’ emotions were still engaged to a great degree when they turned away. In contrast, *Panda Eyes* relies on the participants’ physical location, not the focus of the eyes, and once a participant is aware that the pandas are tracking her movements she will be aware that they are continuing to stare at her even as she walks away. She will likely even be
aware that there is a limit to the range at which the piece can track her, which may well add further content to the piece’s message - the idea that time is limited and, if she simply walks away now, the chance to act will be lost. In addition, if there are other users present she will observe a faction of the pandas turning to face them, reinforcing very explicitly the idea that everyone is accountable. So here we find a work whose message is made clearer by the involvement of other users in a way considerably more appreciable than in a similar non-interactive work.

One area where non-interactive art might be said to avoid the sorts of abstracting or intermediate steps I have suggested are often present are abstract expressionist works from artists such as Mark Rothko, Jackson Pollock and Wassily Kandinsky. Works in such styles are often said to be “experienced” by the audience, in the sense that the artists have tried to reduce the amount of conscious processing involved in appreciating a piece. This is achieved by excluding recognisable images such as people, fruit and so on, in favour of abstract shapes, and the artist’s goal is to have an effect on the viewer at a subconscious level. Whilst there is certainly something to be said for this line of argument, it seems to me that there are important differences between the way works of interactive art and abstract expressionist works achieve this immediacy, and what this immediacy is able to convey. In order for abstract expressionist works to have an effect on their audience’s subconsciousnesses the audience members need to be in what we might describe as a receptive state of mind. We can describe this state of mind as being one which is free (as far as is possible) from any preconceptions about the artist or the piece and from any conscious effort to interpret or analyse the work, so
that the appreciator can be as receptive as possible to the mood or feeling the piece might convey. Whilst this might be described as a rather passive state of mind, it should be clear that it requires an active effort to achieve it, and that if these works are encountered casually appreciators are less likely to ‘experience’ them fully, if at all; certainly such works are often dismissed by appreciators who are frustrated by questions such as “what’s it supposed to be?” and “what’s it supposed to mean?”.

Bearing all this in mind, it would appear that the responses abstract expressionist works are able to generate might be limited in several ways. Firstly, it is easy to imagine audiences feeling a sense of calm, wonderment, excitement, passion and so on when viewing an abstract expressionist piece. It is much more difficult to imagine a work that would elicit such feelings as pride, guilt or personal culpability - feelings that we might term “reflexive” in that the object of the emotion has some sort of direct connection with the appreciator (we feel pride or guilt for actions we have or have not taken, for example). Another way to understand this idea of ‘reflexive emotions’ is to compare it to reflexive verbs in French - je m'appelle, je me lave etc. - which require a reflexive pronoun to indicate that the subject and object of the sentence are the same. Fear for oneself would be considered reflexive in the sense that I envisage, fear for the character about to be (fictionally) consumed by the slime is not, and as we will see later in this chapter, interactive works are able to achieve these sorts of reflexive emotions without the need for abstraction in the way *Panda Eyes* or its non-interactive equivalent would. Although works do not need to be interactive in order to achieve this sense of immediacy, it seems we have grounds to argue that non-interactive works are more
limited in the scope of emotions that can be so immediately elicited. Secondly, when an
abstract expressionist work is ‘experienced’ it is generally held to be a personal
experience, in the sense that the emotions one appreciator may feel might be similar to
or different from those felt by another appreciator experiencing the same work. Whilst
interactive works must admit of the possibility that some participants could have
different emotional responses it seems clear that the artist of an interactive piece is
more able to direct the sorts of responses she desires an audience to feel than the artist
of an abstract expressionist work; interactive works do not need to trade off this
direction against the immediacy of the appreciators’ emotional responses.

However, this claim requires careful unpacking: as we saw with *Boundary Functions*,
the reactions of a user of a non-narrative interactive work come more easily from his
personal experience of the work as placing him within the work constrains to some
degree his ability to imagine himself as someone else experiencing it. We must
therefore recognise that although non-narrative interactive works may allow the artist a
greater guiding role in terms of the emotions she wishes to elicit than non-interactive
abstract expressionist works, this guidance is limited. The same person who feels
isolation and sadness at *Observed Boundary Functions* may experience peace and
contentment when participating in *Boundary Functions* due to the greater involvement of
what they bring to the piece, and if we imagine an interactive version of a Rothko work,
the users would in part determine its display, reducing what little control over their
emotional responses the artist below that which Rothko had.
If we could place the sorts of works discussed on a graph indicating the immediacy of response (the lack of need for abstraction) and the control of the artist over the types of response elicited, it might look something like this:

It should be noted that a user engaging with *Boundary Functions* might also create a minimal narrative in which it is true within his participation in the work that, regardless of his efforts, he is unable to dissolve the boundary between him and his children; this narrative is a candidate for a rational feeling of guilt without the need for abstraction. It is a weak candidate, however, as the work dictates that the boundary cannot be dissolved and if we believe that ‘ought’ implies ‘can’ we should conclude that a feeling of guilt is misplaced.
Lastly, interactive works do not require their audiences to put aside preconceptions or enter into the receptive state of mind previously described in order to experience them (as we have seen, they often require and encourage a very active role from the participant in order for the scope of the interactivity to be revealed). In this way we can perhaps suggest that they are more immediately inclusive and approachable - they can be appreciated, at least to an extent, even when encountered only casually in a way that abstract expressionist works often cannot.

I have argued that non-narrative interactive works provide easier access to personal emotional responses by virtue of placing the user directly within the work. The directness or immediacy of these responses can be achieved by non-narrative non-interactive works, but most successfully in abstract impressionist works which are limited in the scope of emotions they can reliably be expected to elicit and which require appreciators to be in a particular, receptive state of mind. The direct involvement of users in interactive works can be an obstacle to a user trying to imagine someone else’s experience of the work, and the lack of direct involvement in non-interactive works can make it easier for the artist to provide prompts with which to direct the audience’s reactions.

4.2 Identification and the Player-Character

Newer games have embraced and continue to embrace advances in technology – for example the shift from subtitles or speech bubbles to spoken words and then to
professional actors, the mapping and modelling first of people, and now athletes, martial artists etc. and even the faces of professional actors to capture realistic movements and expressions, and – in at least a few examples – greater attention being paid to creating better conceived and executed plots and storylines. These advances help to improve the level of realism within games - not only how realistic things look, but also how realistic the characters sound (both how they speak and what they say) - but they also have a more far-reaching effect: they make it easier for video game developers to borrow from the established aesthetic practices of other media to encourage players to identify with characters within the work in much the same way as they do with those other media. Instead of crudely-rendered hearts floating above the hero’s head and a few bleeps and chirrups from the speaker we now have swelling orchestral movements, detailed facial expressions and spoken dialogue, scenes of high action swap between close ‘over-the-shoulder’ shots and wider viewpoints for smoothly-animated ‘set-pieces’ and so on.

Video games are also becoming ever closer to cinematic works in their use of ‘atmospheric’ sounds. Sometimes the player will have control over these, as in, for example, the Grand Theft Auto series where the player can tune car radios to a number of different stations, each playing a different type of music, and select the one most appropriate to his or her mood. In this case the music is diegetic - it is true in the fiction that the player-character is listening to this music. It is also true in the real world that the player is listening to this music too. Players might even choose music to improve their aesthetic experience of the game – perhaps by playing opera for performing the
bonus stunt tasks which play out in slow motion – or even to modify their behaviour – when earning money as a taxi driver a player might listen to reggae in order to reduce the temptation to drive aggressively (which would cause the passenger to bail out of the cab), or pounding drum and bass when trying to reach destinations quickly or pursuing quarry. Where the soundtrack is not within the player’s control (particularly when it is non-diegetic) it is often used in the same way as in film to encourage or heighten emotional or psychological states – relaxing music for general exploration, suspenseful music at points when a foe might suddenly appear without warning and dramatic, fast-paced music when under attack.

It should not be controversial to suggest that our responses to events that befall the characters in video games are likely to become closer to our responses to such happenings in films the more closely the videogame depictions of those characters and their situations resemble those we see on screen; the less laughable the scripts and deliveries of lines the more likely we are to respond appropriately to what the characters say and the less ludicrous the plots the more likely we will engage with the narrative seriously.

As video games continue to be able to more closely mirror films, whether live-action or animated, they will more often and more naturally be held to and judged by similar aesthetic criteria. The way we relate to and identify with characters in video games will also tend towards the way these processes occur in films. The more fully described characters are within a work the more likely we are to identify with them, at least in the
sense of caring to some degree what happens to them, and this can be seen in both films and video games. Compare the deaths of a significant character who has through little or no fault of her own become caught up in a villain’s plans to that of one of the villain’s minimally described henchmen; we are far more likely to pity the former than the latter. What is of particular interest in the case of interactive fictions, however, is the ways in which players identify with the character they control, their player-character.

Tavinor notes that there is a sense of identity between players and their characters, as players will say “I’m the one with the red hair”, “I’ve just been shot” and so on - “the character is the player’s fictional proxy in the game world.”112 Players generally have the most input in the characteristics of the player-character in role-playing games as the player-character is the centre of the fiction; players are often able to flesh out both the physical properties and set and level of abilities that their proxy fictionally has and the fiction is largely played out in terms of increasing those abilities and developing that character’s individual narrative. Other types of games may provide players with a predefined player-character but, if choices the player makes during the course of the game open up certain avenues from the set of potential story lines and close off others, the player-character of one player’s game world can differ, sometimes significantly, from that of another player who has made different choices (going beyond relatively trivial differences such as the route taken to a particular destination or the time taken to complete a particular goal), particularly when players are given choices to act in moral or immoral ways - my player-character might be amoral and happy to kill anyone who

112 Tavinor, p. 70
stands in his way, yours might act as ethically as possible and negotiate his way around obstructive non-player characters. Player-characters can also be very minimally described, particularly where the fictional world is presented from a first-person point of view. Games such as Sid Meier’s *Civilisation* and Maxis’s *SimCity* cannot be said to provide a player-character as such (despite *SimCity* describing the player as Mayor – Tavinor calls this a “partial fictive gloss that … characterizes their manipulative role”), as the player directs the course of a nation’s or city’s progress over many hundreds of years.

Davies remarks “whether [we are] tempted to talk of [our] participation or the participation of the fictional character [we control] is likely to depend … on the extent to which [we are] like, or identify with” our player-character. This seems particularly relevant to role-playing games (where both detail and player choice are usually strong) as in these games the bulk of the narrative comes from the actions of the player, rather than a prescribed plot. However, even when the player-character is greatly dissimilar from the player (perhaps he is sociopathic, racist or amoral) the player is nonetheless apt to mix language about the real and fictional worlds when describing events in the game (“I shot the mutant” etc.) so Davies’s account is not a suitable explanation for players’ relationship to their in-fiction counterparts. In addition, as noted above many games provide highly minimal accounts of the player character - it might be argued that in those cases the minimal descriptions allow the player to imagine the player-character

---

113 Tavinor, p. 73.
114 Davies, p. 275.
sharing most of her attributes. This too seems at variance with anecdotal accounts of players’ experiences: they do not believe that it is true in the fiction that the player-character shares their attributes. What seems instead to be happening is a certain blurring of the gap between the real and fictional worlds that allows players to ‘project’ themselves in some way into the player-character and experience the fictional world vicariously through him. But just how does this blurring take place? Certainly it is the case that I am able to make certain things true within the fiction and, adapting Joyce’s (7) from the previous chapter, if I perform an action in the real world (pressing a button on the control pad) believing that a something will be made fictionally true (my player-character shoots a mutant), then I suggest I can be deemed to have (fictionally) shot that mutant.

The player-character is usually the centre of the primary source of epistemic information about the fictional world; if the game is presented in first-person perspective the player gains knowledge through the ‘eyes’ of the player-character, if it is presented in a third-person perspective the window the player has onto the fictional world is the ‘virtual camera’ through which events are presented. Maps, health meters, ammunition counts etc. can be placed on the screen to provide additional information in a non-diegetic way, but these devices are so commonly employed, and players so used to absorbing the facts about the fictional world they display that this does not appear to have much

---

115 It is important to note that we might identify with our player-character in the same sense of ‘identifying’ used to describe how we relate to non-player characters, or characters in a film, play or novel. What is key to our understanding of our relationship to player-characters is however the sense of identity - that in some way I am the player-character.
impact on their identification with their player-character, they are simply built in to the players’ imaginative projection into the game-world. Tavinor notes the head up display in *Half Life* and the reddening-with-injury view in *Call of Duty: Modern Warfare* as examples of diegetic information displays,116 and it is arguable that if this trend were to become highly widespread, players would in time find non-diegetic information to be a distraction from their projection into the fictional world.

Sounds can also provide information about the fictional world, particularly when they are used to alert the player to the approach of a hostile character or to the fact that such a character has become aware of the player-character’s presence. As gaming technology has advanced so has the level and usefulness of the audio information – in *Doom*, for example, nearby creatures could be heard and would screech or grunt if the player-character was spotted, but the sounds were mono and so contained no information as to the location of whatever was emitting it, and were either loud or quiet (indicating quite close or very close); modern games leverage stereo or surround-sound speakers to provide direction information and can change in volume with much finer gradation so as to better communicate proximity, or to alert players to the direction from which arrows or bullets are coming. This too helps players make their imaginings more complete and correspond more accurately to the player-character’s perceptual and epistemic set.

Players can also gain epistemic information about the fictional world through tactile feedback from modern controllers, whether about the surface on which they are driving

116 Tavinor, pp. 75-6.
or as a way of alerting the player that their player-character has suffered some sort of an injury, either by means of a sudden sharp shake to indicate, say, a bullet strike, or a tremor to indicate failing health that increases in magnitude in line with the deterioration of the player-character’s condition. Again, this kinesthetic feedback helps to make the player’s world correspond more closely with the player-character’s. Such stimuli could also be used to indicate the player-character’s trembling whilst creeping through a cave (or, indeed, as indicative of their mounting anger at the taunts of another character) in order to encourage or heighten similar emotions in the player. This trend seems set to continue - in the recently released *Mass Effect 3* developed by BioWare, a camera and microphone can be utilised in the control of the player character. Instead of using buttons to highlight and select the desired response to a non-player character the player can simply read the dialogue aloud, and in the case of questions requiring a yes or no response the player can nod or shake her head and the camera will detect this. Advances such as these are very illuminating for our purposes - what in effect they achieve is to make it true both in the fiction and in the real world that the player and player-character have said a particular phrase or non-verbally communicated assent or dissent to a proposition. The game even goes so far as to allow doors to be opened by the player raising and pushing forward her arm - an action that she has carried out in the real world and her player-character has fictionally carried out in the game-world. We can imagine further progression along these lines if we think of articulated gauntlets which provide resistance and tactile feedback to communicate to the player in a highly direct and physiologically immediate way the effort required to push an obstacle out of their way. The player’s exertions in the real world would mirror the fictional exertions of
the player-character. As this mirroring of real and fictional worlds increases, the experiential gap between the worlds will reduce, and it should therefore be no surprise that players identify ever more strongly with their player-characters, and can imaginatively project themselves ever more fully into the fictional worlds with which they engage.

4.3 Narrative Works

As videogames continue to evolve, there seems to be greater weight placed on the emotional responses developers expect the characters and plots to elicit; Tavinor notes in *The Art of Videogames* that he and friends found themselves taking actions based on how they had come to feel about particular non-player characters in some of their games, and draws a distinction between emotions a player might feel when succeeding or failing at a task and about the fate that might befall his or another character; 117 I might feel sad when my character dies because I have failed to defeat the monster, but there is a significant difference between feeling emotions towards the death of a character within the fiction and towards my failure to press the right buttons at the right time in the real world.

Just as with non-narrative works, it is often the case that the artist of a narrative work intends her audience to feel particular emotions and, should they fail to do so, it might be said that the artist or the piece have failed, or that the appreciator has failed to

understand the work properly. As we have discussed, however, in interactive narratives emotional responses can be intended to serve various additional purposes: for example, to guide the player’s actions, to influence the choices he makes, or to help - or indeed hinder - his ability to play the game well. For example, fast-paced music might be used to amplify the intensity of a racing game in order to help the player stay focused and alert to oncoming obstacles or difficult sections of track, feelings of sympathy towards a non-player-character may provide the player with motivation to choose one course of action over another, feelings of fear and panic may cause a player to waste scarce ammunition, shoot an innocent bystander or fail to act with the level of skill necessary to defeat a particular enemy. Tavinor goes on to suggest that “[f]aced with a rich decision space in which we need to act, emotions not only focus our attention, but also help to bias the choice over options so that efficient decisions might be made”;\(^{118}\) given his problems panicking when confronted by malevolent monkeys this claim seems unusual – indeed, advanced games might be used to help train individuals to overcome fear, confusion and panic in order to respond effectively in times of crisis, and to achieve this they would need to be designed to bring out these unhelpful emotions and demonstrate how much of a hindrance they can be.

Tavinor quotes neuroscientist Antonio Damasio’s statement that “[a]ll 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

\(^{118}\) Tavinor, “Videogames and Interactive Fiction”, p. 38. Tavinor goes on to acknowledge that panic achieves the reverse result.
about the life of an organism, its body to be precise, and their role is to assist the organism in maintaining life”.119 This is clearly not the end of the story, however, as this would seem to preclude many altruistic actions; aside from the usual extreme examples such as people throwing themselves on grenades to protect others, there are various others that illustrate the point further - one such example might be the shame attached in some nomadic cultures to refusing to share food and water. It may be argued that the underlying motivation is still some form of “potential” self-preservation, insofar as the logic is presumably that one may be in a position to share supplies today but in desperate need of them at some point in the future, and it is in everyone’s wider interest to attach a powerfully motivating emotion such as shame to the refusal to do so. It is interesting to note, however, that there are countless examples of individuals acting for the benefit of their species or local group (however construed) whether they are ants or bees defending their queen, or bears or swans defending their young, and there is clearly an evolutionary advantage to this type of behaviour so it seems it may not be simply a question of intellectually planning for one’s own potential future benefit or survival. There are also various examples of emotional responses being directly harmful to an organism: fear may cause the failure to act to preserve one’s life - the proverbial “deer frozen in the headlights” - or to act in a way that exacerbates a life-threatening situation - you may be aware that the best course of action is to drop to the floor and roll, but the panic of finding your clothes alight may cause you to flail about in terror. We can see parallels in various video games, either in cooperative games where one player sacrifices her character in order for her team to prevail, or in single-player

titles such as *Deus Ex: Human Revolution*, where the player is ultimately faced with choosing between courses of action with the potential to lead to civil war, or sacrificing oneself in order to destroy the relevant evidence in the hope of preventing it.

Having examined some potential purposes of emotional responses, we can now turn our attention to the responses themselves. We can divide the emotional responses to narrative works into two broad categories: those directed at the work itself and those directed at the characters within it. Responses to the work itself are relatively philosophically unproblematic for non-interactive works, as they are directed at real objects or their qualities; in works of mixed media, cinema for example, audiences might have emotional responses directed at the images they see or the soundtrack they hear. As Tavinor notes, with interactive narratives such as videogames, participants could respond not only to the images and soundtrack, but also to the very playing of the game, i.e. they can be frustrated at the difficulty of a game, proud of having completed a game or having solved a challenging puzzle within it, excited or bored by the pacing of a game or sequences within it. Tavinor expresses concern with these emotions however, since “understanding just why a gamer is successful or unsuccessful in a game demands that we refer to their fictional activities” and goes on to suggest, with reference to a particularly troublesome foe within a game he had played, that “[w]hat annoyed [him] when [he] played *Devil May Cry* was not merely that the game was so difficult, but that the fiery spider monster was so tenacious”. Despite going on to lay

---

out a complex framework through which he hopes to explain emotional responses to fictional characters and events Tavinor does not return to address specifically his worries with a player feeling elation or frustration at their ability to deal with challenges set within the game. These worries are not as well-founded as Tavinor believes them to be, and a more straightforward explanation for them can be given. Tavinor’s key concern is that we “refer to … fictional activities”; it should be clear, however, that this conflates two quite different statements about what is happening. That this is confused is brought out further by his next statement, that seems to merge both the real difficulty of the real game and a fictional trait of a fictional character within it. I suggest that a better way both to describe and to understand the situation is preserve the distinction between what is real and what is fictional; we can describe the responses a participant has to the difficulty of a game or a task within a game by saying that the player is elated or frustrated by what he has really caused to be fictionally true. In order to win the encounter with the fiery spider monster, the player must make it fictionally true that his character within the game hits the vulnerable spots on the monster enough times to do sufficient damage to kill it. The player does this by taking real actions - moving his mouse and pressing buttons on his keyboard - at appropriate times - when the game’s (real) display alters in such a way as to indicate that it is fictionally true that the hero and the monster stand in a relation to one another that would allow a successful attack. It is fictionally true that the fiery spider monster is a worthy and troublesome adversary for the character; it is really true that the mechanics of the game are such that it is difficult for the player to press the right buttons at the right moments a sufficient number of times (to make it fictionally true that the hero defeats the monster).
Although Tavinor accepts that we can hold beliefs about what is true in the fiction, he
denies that these beliefs are what cause emotions - instead he suggests that it is
simply our imagination that can give rise to emotional responses and claims that to
believe something is true in the fiction “must mean, of course … that we believe certain
things are imaginary or to be imagined of a fiction”.\textsuperscript{122} Tavinor’s move then, is to make
belief in a proposition unnecessary for the generation of emotional responses, all that is
required is that we imagine something to be the case as “what we imagine is often just
as capable of causing emotions as what is believed”.\textsuperscript{123} Tavinor goes on to outline how
this process can be applied in the case of video games: the participant must first enter
into something like a Waltonian game of make-believe using the game as a prop. Then,
as she progresses through the game, elements of the game cause an emotional
response in the player - Tavinor suggests many of these responses are “\textit{subdoxastic}, in
that they are not mediated by higher cognition, lack an inferential structure, and rely on
more direct sub-cortical pathways in the brain”;\textsuperscript{124} in other words, they do not require any
significant reflection on what is perceived. Tavinor gives examples of disgust at the
images in survival-horror games, and the startle response resulting from the sudden
and unexpected, such as explosions or monsters leaping out at the player’s character
without warning. Tavinor does not limit himself to what we might call “knee-jerk
reactions”, however, arguing that emotions can also be caused by more complex
imaginings that combine what the player perceives on the game’s display with what she

\textsuperscript{122} Tavinor, \textit{The Art of Videogames}, p. 136.  
\textsuperscript{123} Tavinor, \textit{The Art of Videogames}, p. 136.  
\textsuperscript{124} Tavinor, \textit{The Art of Videogames}, p. 137.
imagines to be true of the fiction - the display changes in such a way that she understands indicates that a monster is approaching her - and this combination causes her to feel fear. The players becomes aware that she is feeling fear and this awareness “can then contribute to the cognitive appraisal of the situation”\textsuperscript{125} i.e. if she is mistaken about what she imagined was about to happen (what she imagined was about to become fictionally true) - the monster does not appear, it was a shadow that she misinterpreted - she can attempt to regain her composure, if it is indeed a monster either she will use the state of heightened awareness resulting from her fear to deal with it successfully or she will panic. The final step in this process is when she “integrates this emotional response into [her] game world by characterizing it in the terms offered by the fictive scenario [she is] playing in”;\textsuperscript{126} this appears simply to be appreciation of the relative intensity of the emotion and reporting this either to others or to herself.

Whilst this theory has attractive aspects to it, it relies heavily on the subdoxastic responses such as shock and revulsion to justify the removal of belief from the arousal of emotional responses. Tavinor’s claim that this process scales to encompass “highly cognitive states” is suspect, insofar as it relies on the “realization of what is imaginary”\textsuperscript{127} to explain why the player’s fear is transferred from her knee-jerk reaction to the barely-perceived shifting shadow to the mutant that is revealed. It is surely the case that this “realization” amounts to nothing less than the formulation of the belief that the mutant fictionally poses a threat to the player’s character. Tavinor himself describes his experience of System Shock 2 as having lead to bouts of panic that caused him to

\textsuperscript{125} Tavinor, The Art of Videogames, p. 138.
\textsuperscript{126} Tavinor, The Art of Videogames, p. 139.
\textsuperscript{127} Tavinor, The Art of Videogames, p. 137.
make mistakes; whilst drawing parallels with Walton's framework has been helpful in setting out his position, it is unlikely that he felt he was pretending to panic, and if his behaviour was such that his state was indistinguishable from genuine panic it seems Walton's denial of genuine emotional responses to fictions is as unappealing in cases of interactive media as it is to non-interactive media.  

As I have previously argued, we should be comfortable with holding beliefs about fictional worlds, and with imagination being able to elicit emotional responses. There should therefore be no problem allowing players of video games to have the same emotional responses to the fictional characters of their games that audiences have to characters in non-interactive works. What is different about interactive works is that the player can make certain things true within the fiction. What I have termed “reflexive” emotions such as pride and guilt should not be surprising then, as they are based on what we have really made fictionally true. The most interesting of these “reflexive” emotions for this discussion however are those of fear for oneself and jealousy. From our discussion on the blurring of the real and fictional worlds there should be little surprise that I hold such emotions to be possible. As to whether they are rational, the situation remains somewhat unclear. It is undeniably the case that it can be true within the fiction that my player-character is in danger, or that a non-player character has something which I believe ought to be my player-character's. However, if we wished to suggest that it is fictionally the case that the item in question should be mine, or that I

---

128 In The Art of Videogames (p.141) Tavinor suggests that although Walton may be wrong to deny that our responses to fictions are genuine emotions, it would be accurate to describe our “relationships to fictions” as fictional.
am in danger, we would have to rely on my sense of identity with my player-character. There can certainly be instrumental value in my having these emotional responses, as they can help me focus my attention correctly in order to achieve my goals within the game, and my real-world goal of completing the game. Of course, they can also cause me to panic and make mistakes, but the instrumental value to me remains if I have a reasonable belief that experiencing them will help me deal with them more effectively in the future. This instrumental value transfers to the real world even more directly if we consider simulators for military or emergency-service personnel - by gaining experience in a safe environment they will be better equipped to deal more effectively with similar real-world situations. As I have illustrated, technological advances make my sense of identity with my player-character profound, and this trend will doubtless continue. In the case of jealousy, it is fair to say that I can rationally believe that a fictional item should fictionally be in the possession of my player-character. However, this seems to be a case of feeling jealousy on behalf of my player-character, in a way analogous to believing that a real item should be in possession of my child, spouse or friend and feeling jealous on their behalf. Although this is at variance with a strict definition of the word jealousy, the feeling is indistinguishable from jealousy strictly understood, so I suggest that it is possible to feel jealousy rationally with regard to interactive fictions. Furthermore, there is a strong sense in which my instantiation of the player-character is mine and as a consequence so are his possessions, and it seems reasonable to suggest that it can be the case that I can actually possess a fictional item - an example of how we can make sense of this comes from *World of Warcraft* by Blizzard entertainment. Within the game’s fictional world, items can be traded between players
and there are some items which are extremely rare and useful within the fiction in some way. When players come into possession of these items, they have on occasion traded them to others for real money. This is a clear example of a fictional item being treated as a real good; it actually belongs to one player, who sells it - in the real world - to another player. There have also been reports of items being “stolen” - a player’s login credentials are compromised, and the “thief” trades the rare item to his own player-character. Given the difficulty involved in obtaining such items, it is rational for players who have been “robbed” in this way to feel a sense of anger, loss and other emotions we would expect of someone who had had a non-fictional item stolen from them. It could be objected that the pertinent detail in these cases is that the item has become the possession of another real-world person, and that it would be irrational to have these feelings if the “thief” were a fictional character. However, if I can really possess a fictional item in one case, I must surely be able to own it in the other; if I can really possess a fictional item then there appear to be no ground for denying the rationality of feeling jealousy with regard to a fictional item.

The case of fear for oneself is similarly problematic. Certainly my identification with the player-character might explain why I describe feeling fear for myself, but it appears again to be a case of feeling fear for another - my player-character - and it would be objected that I am either mistaken or irrational. One reply would be to suggest that I am in real danger of losing my progress within the fiction, and I experience this as fear for myself. The fact remains however, that I am not in danger, it is only my fictional counterpart. If we were to imagine that it were possible to inhabit the body of another
real person, and that person were in danger, I may well feel fear for myself if I was so engaged in the situation that I had forgotten that I personally was in no actual danger. However, I would be forced to concede that this fear was irrational given that my real body was safe from harm. I am forced to conclude than genuine fear for oneself when engaged with an interactive fiction remains irrational, although I suggest that when players describe feeling fear for themselves they are simply misreporting what they have experienced. They may feel for on behalf of their player-character, or fearful that they will lose their progress within the fiction, but claims of feeling genuine fear for oneself are either mistaken or irrational.
Conclusions

I hope to have described what works should be considered fictions, and how we can better understand the notion of what it is for something to be fictional. I have offered a minimal definition of what it is for something to be a narrative and described how we might assess the qualities of narratives. I have argued that implicit internal narrators are not necessary, and shown how explicit internal narration can be achieved in an interactive work. I have described how we should understand and apply the term ‘interactive’ to works of art, and what would be required from interactive narratives and narrators. I have suggested that it is both possible and rational to feel most, but not all, emotional responses when dealing with non-interactive works of fiction, and argued that even when certain emotional responses can be shown to be instrumentally useful, the responses themselves remain irrational.

I hope to have shown how interactive works, both narrative and non-narrative, can elicit stronger emotional responses than non-narrative works, and to have explained how the ability of an artist to control or direct just which emotions we feel is reduced in interactive works. After explaining how the sense of identity players feel with their characters can become extremely strong I have argued that the range of possible and rational emotional responses is greater for interactive works than it is for non-interactive works, but that feeling genuine fear for oneself remains irrational.
Bibliography


