

The paradox of fiction /
The paradox of life

Lisa Jacobs

Date: September 12th, 2017

Student number: 0337056

Supervisor: dr. L. de Bruin

Word count: 19513

Scriptie ter verkrijging van de graad “Master of arts” in de filosofie
Radboud Universiteit Nijmegen

Hierbij verklaar en verzeker ik, Lisa Jacobs, dat deze scriptie zelfstandig door mij is opgesteld, dat geen andere bronnen en hulpmiddelen zijn gebruikt dan die door mij zijn vermeld en dat de passages in het werk waarvan de woordelijke inhoud of betekenis uit andere werken – ook elektronische media – is genomen door bronvermelding als ontlening kenbaar gemaakt worden.

Plaats: Amsterdam datum: 12 september 2017

Summary

The paradox of fiction lies in the seeming contradiction that we appear to have real emotions for fictional characters, while at the same time knowing that they do not exist. The several solutions all have their drawbacks, and mainly focus on behavioural and emotion disanalogies between emotions for reality and fiction. However, in focusing on fiction, they overlook the many borderline cases of emotions in real life that exhibit the same disanalogies. The most promising solution – thought theory – claims that the object of emotion is the thought, not the (non-existing) object. As we also deal with reality in a narrative fashion, this means that the paradox actually extends to reality.

Index

Summary	4
1. Introduction	6
2. Denying the belief condition.	12
2.1 Illusion theory	12
2.2 The issue of the existence of fictional characters	14
3. Denying the response condition	17
3.1 Walton's quasi-emotions	17
3.2 Object of emotion	26
4. Cognitivism and the irrationality of fictional emotions	28
4.1 Cognitivist theory of emotion	28
4.2 Radford's conclusion: the irrationality of fictional emotions	32
4.2.1 Rationality of emotion	33
4.2.2 The appropriateness of emotion	36
5. Denying the coordination condition	42
6. The difference between reality and fiction	48
6.1 The truth value	48
6.2 Narrative theory	50
6.3 The implications of make-believe	53
6.4 Truth in fiction	54
7. Conclusion	58
8. References	61

1. Introduction

The paradox of fiction was first suggested in its current form by Colin Radford (Colin Radford & Weston, 1975). It consists of the two seemingly incompatible assertions that a reader feels some form of emotion towards a fictional character while at the same time he is perfectly aware that the character does not exist. That we feel some sort of emotion for fiction seems to be an incontrovertible fact of everyday life; it is at the core of why we enjoy reading, watching plays, and going to the cinema. We cry when our favourite character dies and we cheer for Laurie and Vincent to finally get together on our beloved soap opera. At the same time, we are perfectly aware that our favourite character, as well as Laurie and Vincent, do not actually exist. The problem arises because we assume that being in an emotional state of compassion normally entails having a particular *belief* about the object of one's compassion. Robinson explains it as follows:

“I believe, for example, that Anna Karenina has had a raw deal ~ whereas if I know that Anna does not exist, I cannot believe that *she* has had a raw deal, since there is no “*she*” for me to have the belief about (and I know this).” (Robinson, 2004)

The problem stated above is how Radford framed the paradox of fiction, but its philosophical traces go back further. Hume talks about it in ‘Of tragedy’, noting that we feel emotion about fictional characters, though laying the emphasis more on another paradox: the paradox of tragedy (Hume, 1757). This paradox entails a longstanding debate about how we can enjoy a tragedy. How can we feel pleasure watching suffering and empathizing with it? But at the basis of this paradox lies the same observation, namely, we appear to feel emotion when watching a play.

Of course, one may wonder why this is a paradox in the first place. It seems a fact of life that we experience emotions in response to fiction. How is belief tied into this in the first place? Imagine I tell my sister about the horrific car accident that my friend has had, leaving her horribly disfigured and unable to continue working

as a volunteer for starving African children. My sister is suitably horrified about the story and feels terribly sad for my friend and the poor children she can no longer help. Now imagine I tell my sister it was all a joke and that no such accident has actually happened. My sister will be angry (and knowing her, probably try to hit me), but she will no longer feel for my suddenly imaginary friend.

This example seems to show that our emotional response is irrevocably tied to our beliefs about the subject, both about the circumstances and the existence of that subject. Take away that belief and apparently we are no longer capable of having the emotion.

Eldridge (2003) summarizes the paradox as follows:

1. We are moved by the career and fate of Anna Karenina. (The so-called *response condition*)
2. Anna Karenina does not exist, and we know this. (The *belief condition*)
3. Being moved by the career and fate of a subject requires belief in the real existence of that subject; it is impossible to really care about something that one knows does not exist. (The *coordination condition*)

Historically, there are various ways of dealing with the paradox. I will briefly outline the various approaches in this introduction, before engaging with them more fully in the following chapters. The first approach is to deny the response condition (1.) and go against our common-sense intuition about what we feel when we read or view fiction. The main proponent of this tactic is Kendall Walton (1990) who approaches the paradox by denying the fact that we are actually moved by Anna Karenina, instead saying we “make-believe” when we feel emotion for characters. He claims it is “fictional when we feel sorrow or terror” and what we are actually experiencing are not real emotions but “quasi-emotions”. In a way Hume employs a similar move by arguing that it is not in fact

the career and fate of Anna Karenina that we are moved by, but the eloquence with which it is presented.

“I answer: This extraordinary effect comes from the very eloquence with which the melancholy scene is represented.” (Hume, 1757)

Another approach is denying the belief condition (2.) by saying that we do not in fact know that Anna Karenina does not exist. These types of solutions are commonly referred to as the illusion theories of fiction. This group of approaches started with Coleridge’s “willing suspension of disbelief”, which basically states that while engaging with fiction we actually believe that the fictional characters or events we are engaging with actually exist. Alternatively, we can deny (2.) by saying that Anna Karenina does in fact exist, risking a trip to Meinong’s jungle.

The 3rd premise can also be denied, by removing the belief condition from any sort of definition of emotion, and thus denying that true belief is a necessary condition for real emotion. The idea that belief in the propositional content of the emotion is a necessary condition for experiencing (rational) emotion is at the basis of the generally accepted cognitivist theory of emotion. Thus denying this condition makes it difficult to maintain any sort of propositional theory of emotion. The non-coordination theories that deny (3.) allow the possibility of having an emotional response to a character, while at the same time believing that the character is purely fictional. In this respect the most successful candidate is *thought theory*, which states that we can have an emotion directed purely at a thought, rather than at an object.

Finally, Radford’s solution to his own paradox is that the emotions we feel for fictional characters are in fact irrational. He says that we do have genuine emotional reactions to fictional characters and events that we believe to be purely imaginary, while (tacitly) holding that we should have genuine emotional reactions only to characters and events that they believe to be real. These reactions

are inconsistent and incoherent and as such are irrational. Implicit in this solution is the idea that it is possible for emotions to be rational or irrational and that it is possible for an emotion to be both irrational and genuine. Radford's solution would also seem to imply that all the other theories deal with finding a solution that allows for rational emotions for fictional characters.

The solution to the paradox of fiction is important, because we take emotions to be meaningful; we see them as sufficient reason for acting. If I hit my little brother and my mother asks me why, I can answer: "Because I am angry (with him)." Though my mother and I might disagree about my reasons for being angry or my way of dealing with those feelings, there is no disagreement between us that my being angry was the reason I slapped my brother in the face.

The idea that an emotion is a sufficient cause for acting is also where the special case of fiction is apparently different: if in fact, we do experience real emotions when watching a play, why don't we act on them? If I see a play and feel the kind of anger at one of the characters that causes me to act violently towards my brother, why don't I slap the actor in his face as well? These behavioural disanalogies will need to be explained.

In the end, we read and we watch precisely so we can experience emotions. Isn't the reason we read a romance novel to feel emotional about the development of the relationship between the two lead characters? If we take away the possibility of experiencing real emotions in response to fiction, we take away one of the key ingredients that seems to make fiction such a powerful source in our lives.

The question I will attempt to answer in this thesis is whether the paradox of fiction is really specific to fiction. As seen in the possible approaches for solving the paradox, it seems necessary to say something about what it means to feel an emotion in general, not just specifically emotions in response to fiction. There are

several puzzling features of emotions in this respects, namely their intentionality, their rationality and their appropriateness.

First, I will argue why denying the response condition and belief condition are not viable approaches to solving the paradox. I will continue to argue that thought theory is the best approach to solving it, but that its consequences go beyond fiction. In fact, our emotions in response to actual events are dependent on the same notions as emotions in response to fiction are. I will argue that what we in fact respond to is the underlying coherence of the narrative, regardless of whether it is true or fictional, and we judge the rationality and appropriateness of any emotion by that coherence. This is similar to what Hume states in the above quote: we respond to the eloquence by which the scene is represented. In elaborating on each approach I will thus ultimately argue that the previous approaches have looked at fiction the wrong way around. Each has approached fiction as something separate from what we call reality.

The cognitivist theory of emotion makes the fundamental assumption that we have to belief in the existence of a character of event in order to feel emotion and this causes the paradox to arise in the case of fiction. In arguing this, it relegates emotional response to fiction to being a special instance of emotion, somehow fundamentally different from the emotion I experience when watching the 8 o'clock news. This in fact is a misrepresentation. This makes the problem even more salient and yet more complex, as the problem of the intentional nature of emotion would in that case also pertain to reality, as well as to fiction, making the cognitivistic account of emotion also problematic in previously seemingly unproblematic occasions.

Thought theory allows us to deal with emotion in response to fiction in the same way as emotions in response to actual events. When we no longer consider emotions in response to fiction as a special case of emotion in general, but rather we view it as a paradigmatic case of our experiencing of emotions, it becomes

clear what role the coherence of narrative structures plays in emotion and that we should adapt thought theory to accommodate this central role of coherence in order to have a functional theory of emotion that also explains the apparent phenomenology of emotion itself.

I will argue that it is not the case that we cannot feel emotion for fictional characters or fictional events as we do for actual people or events, but rather that we ultimately perceive the world as 'fictional' to a certain extent and our emotion about the things in it is contingent on that. What I mean by this is that the input we receive, the stories our friends tell, the news we see on TV, the things we experience are embedded in a narrative that makes these events coherent for us, so that we can adequately deal with this information and undertake the right kind of action. Our approach to fiction is essentially the same as we approach reality.

2. Denying the belief condition.

Anna Karenina does not exist, and we know this. (Belief condition)

2.1 Illusion theory

In a way, the simplest way of dealing with the paradox of fiction is to deny that we actually know that Anna Karenina does not exist. Noel Carroll calls these theories the ‘Illusion Theories of Fiction’ (Carroll, 1990). These theories state that the fiction we consume is so overwhelmingly convincing that we can’t help but believe that there is an actual monster on the screen when we are watching Frankenstein. Through the illusion that fiction puts before us, we actually believe that the fictional character exists and thus there is no paradox. An early proponent of such a theory of fiction was Samuel Taylor Coleridge who, in a quote attributed to him, advocated that the consumer of fiction engaged in a “willing suspension of disbelief” (Coleridge, 1885).

One of the most powerful objections to the illusion theories of fiction is the one of behavioural disanalogies: people just don’t behave like they would if they actually believed in what is happening. Why don’t I tell Juliet that Romeo is not really dead when she is on the verge of killing herself? Presumably, I would ordinarily do something to stop the suicide that is about to happen right before my eyes of one I empathize with. Or if I see green slime crawling towards me on the television and I actually believe it may kill me, why don’t I run in the opposite direction? Why do I put the pillow before my eyes, but stay put on the couch? It seems obvious that people don’t respond with the same actions towards fiction as they would if they believed what they saw was actually real. If the illusion theories of fiction are correct, there seems to be no reason for me to act differently when watching a play or watching reality unfold before me. The problem of behavioural disanalogies is a recurrent problem when looking for a way to solve the paradox that any theory attempting a solution will need to address.

Another thing is that we would ourselves deny that we truly believe what we are watching. Furthermore, as Carroll points out, the illusion theory is also less applicable to literary fiction than it is to films and plays (Carroll, 1990). Why would mere words make us forget that they are merely a fiction written down, instead of reality? In plays and films at least I get something approaching the ‘whole experience’: images, sound. How could mere words conjure up the same illusion of reality that is so convincing that I forget I am reading a book? This seems unlikely.

Then there is also the point that is illuminated in the so-called paradox of tragedy: we appear to enjoy reading/watching fictions. This, in a way, is an emotional disanalogy similar to the behavioural disanalogies mentioned above. If we actually believed that Dumbledore had just died, how could we still make statements such as: “That was the best Harry Potter novel so far, I thoroughly enjoyed it”? But we can and do make such statements and we do it coherently. Again as much as this is an argument against illusion theories, it is problematic for any account attempting a solution to the paradox. I will briefly revisit the paradox of tragedy in chapter 4, when dealing with Radford’s views.

In Coleridge’s quote the word “willing” is particularly noticeable. It seems to imply that the suspension of disbelief, or the belief in the illusion of the fiction is something that is under our voluntary control. As Carroll rightly points, this doesn’t seem to align with how we normally think belief functions. We cannot just will ourselves to believe that $3 \times 3 = 6$. It seems equally unlikely to think we can will ourselves to believe that we actually live in the fantasy medieval world of Game of Thrones. However, there is a point to the willingness that is in Coleridge’s quote. We can and do speak coherently about how believable a novel or film is. The believability seems to play an important part in how good we judge the piece of fiction to be. It seems far more likely (and in accord with our experience of fiction) that our willingness to ‘believe’ in a fiction is dependent on its eloquence, to borrow from Hume again. As judging the eloquence of a play or

movie or book requires cognitive action, this would require a scenario as follows: I watch a movie. After I judge the quality of writing to be believable, the actors to be lifelike, the special effects to be undetectable, I judge the film to be believable and thus I will myself to suspend my disbelief in the scenes I am watching. This allows me to experience genuine emotion. There seems to be a contradiction already in the idea that I have to find the film believable in order to will myself to believe. But the distinction between finding something believable and actually believing is important and is nicely illuminated by looking into Coleridge's idea.

Also, when we talk about a novel being believable, we do not mean that we actually believe what happened in the book. Rather, we mean something along the lines of "the book formed such a coherent whole, that I could believe in its possibility". We do not actually believe. We believe the hypothetical possibility of it being actual. When we talk about the suspension of disbelief, or another form of illusion theory, we cannot explain this common way of talking about literature.

Believability seems part of the conditions necessary for us to have emotions about fiction, though it certainly doesn't mean that a narrative has to be possible in the world we live in. We judge Harry Potter to be believable, *within the universe* of Harry Potter. We need to be able to relate to the narrative, but that is also true for any news story. We are touched more by a story happening in our neighborhood or in a culture similar to ours, than by something on the other side of the world. I will return to this issue in chapter 4.

2.2 The issue of the existence of fictional characters

Alternatively, we can look at whether it is true that Anna Karenina does not exist. Here we enter the field of the metaphysics of fictional entities. In short, the discussion about the metaphysics of fictional entities circles around the problem that we do make statements about fictional entities such as "Anna Karenina was Russian", even though Anna Karenina does not exist and thus the Anna Karenina

in the sentence is not referring to anything. This makes it impossible to determine the truth value of the statement. A solution to this is to allow that fictional entities do indeed exist.

Accounts using this approach deny the belief condition and thus seemingly clear away the paradox. Some realist views about fiction take fictional characters to be some kind of beings that do indeed have the properties that are ascribed to them in the story in which they figure. One of these views is the idea of the Meinongian non-existent object; an entity that exists neither spatiotemporally nor non-spatiotemporally (Routley & Sylvan, 1980, p. 7). Another view is that of possibilism, which holds that fictional entities do not exist in the *actual* world, but only in some other possible world.

However, the actual existence of fictional entities would not solve the paradox of fiction. It would just change the paradox (for possibilism) to:

1. We are moved by the career and fate of Anna Karenina *in this world* (Response condition)
2. Anna Karenina does not exist *in this world*, but does exist in some other non-actual world, and we know this. (Belief condition)
3. Being moved by the career and fate of a subject requires belief in the real existence of that subject *in this world*; it is impossible really to care about something that one knows does not exist *in the same world*. (Coordination condition)

The problem is that when we accept possibilism, Anna Karenina exists in a non-actual world, which is by definition separate from our own world – the world in which we feel moved. Why would I be moved by an Anna Karenina in some other world, when she does not exist in this world? It is completely unclear how it is possible for some entity in a non-actual world to have causal influence on the way I feel. Also the second part of the coordination condition becomes unclear in this

respect. Do we actually know that Anna Karenina exists in some other world? Because if we don't know – are unsure – the question becomes: can we feel emotion about some entity about whose existence we are unsure? So the question is merely displaced.

Posing new metaphysical entities solely for the purpose of solving the paradox is perhaps a bit extravagant and it does not solve the initial problem, which is that I know that (*in this world*) Anna Karenina does not exist and it is a prerequisite for the coordination condition. There are of course many other objections to this debate, but discussion of these falls outside the scope of this thesis. For now, it is enough that we have established that denying that fictional entities do not exist doesn't help us solve this particular paradox.

3. Denying the response condition

We are moved by the career and fate of Anna Karenina (Response condition)

3.1 Walton's quasi-emotions

Another way of dealing with the paradox is to deny the response condition. In general terms the response condition states that we experience an emotional response towards a certain fictional character or fictional event. In other words, denying the response condition means that we deny that we experience an emotional response towards a certain fictional character or fictional event. It follows that we are not in fact moved by the career and fate of Anna Karenina. This seems to go against common sense: we all know (or most of us do) what it is like to read an incredibly sad book and cry or feel happiness when the hero in the action movie prevails.

The most famous account using this approach is K.L. Walton's (1978). In an interesting move, he says we are indeed moved by fictional characters, but the emotions we experience are only quasi-emotions. In fact, when we read or watch a work of fiction, we engage in an elaborate game of make-believe. He famously uses the example of Charles. Charles is watching a film where deadly green slime is coming towards him, threatening to destroy him. Upon seeing this, Charles exclaims that he is scared out of his mind. Walton denies that Charles is actually scared out of his mind. Instead Walton's account of Charles' experience is as follows:

“Charles believes (he knows) that make-believable the green slime [on the screen] is bearing down on him and he is in danger of being destroyed by it. His quasi-fear results from this belief.” (K. L. Walton, 1978, p. 14)

According to Walton, Charles is playing a game of make-believe, using the images on the screen as props. So the fear Charles experiences is crucially different from that of a person with an ordinary case of fear and Walton thinks the fact that Charles knows that the slime is fictional is good reason to deny that what Charles feels is actually fear. This fictional or make-believe fear is what Walton calls quasi-fear.

Walton explains this by using the example of a girl playing a game of make-believe with her father. The father pretends he is a girl-eating troll, and every time he tries to get hold of her, she runs and screams. She make-believes that she is terrified of the troll coming to get her and so she runs and screams. Similarly, Charles, wanting to be entertained, make-believes that the green slime is actually coming towards him.

In short, Walton has two arguments for his claim that the emotions are merely quasi-emotions and that we are engaging in a game of make-believe. The first argument goes as follows: we only fear what we believe can endanger us. We know what is fictional does not exist and thus cannot endanger us. Therefore, we do not feel genuine fear towards fictional characters or events. The second argument has to do with behavioural disanalogies, already mentioned in chapter 2. What Walton points out is that the behaviour of the girl is not compatible with her actually thinking her father is a girl-eating troll. If she would truly believe that, she would not protest when her mother tries to put a stop to the game, because they are making too much noise. This, according to Walton, is similar to Charles, who stays put in his chair, which certainly he would not if he thought the green slime would actually kill him.

One problem with these behavioural disanalogies is pointed out by Alex Neill (1991). Neill points out that there is an important difference between the girl playing make-believe and Charles. The girl's standpoint in playing the game is *internal*. She is the participator. In contrast, if I were watching the game, as her mother does, my standpoint would be *external*, which is similar to Charles'. As

Neill points out: “From the external standpoint, one has beliefs about what is make-believedly the case relative to the game. From the internal standpoint, one makes-believe.” (Alex Neill, 1991, p. 51). Walton claims Charles’s quasi-fear is caused by his *belief* that it is make-believe that he is threatened. So, according to Neill, the account needs to be modified into Charles’ quasi-fear being caused by *making-believe* that he is threatened by the green slime.

The fact that the girl in the above example is playing make-believe is identified by what she *does*: she runs away, but keeps coming back for more. Charles’ game of make-believe that he is experiencing fear is identified by the fact “that he experiences “quasi-fear,” that he clutches at the arms of his chair, that he utters the odd involuntary shriek, and so on.” So Charles’ supposed game of make-believe, according to Neill, is identified by what he *suffers*. The difference here is that the girl purposefully runs, while Charles involuntarily has his heartbeat rise and flinches when the green slime nears him. So in the first case there is evidence that the response is faked, that there is a game of make-believe going on, while in the second case, it is absent. There is no actual evidence that Charles is making-believe that he wants to escape: he does not run, he does not yell “let’s get out of here”. So there is no evidence that Charles’ quasi-fear is caused by making-believe he is threatened by the green slime.

Another point is that Walton does not explain the behavioural disanalogy he himself notices. If Charles is fictionally afraid of the green slime analogous to the girl being fictionally afraid of the girl-eating troll, it still doesn’t explain why Charles doesn’t run and scream, just as the girl did. If he is engaging in a game of make-believe, why doesn’t he extend it to a game where running and screaming are involved? Is it because it is implicitly part of the rules we accept when we watch a movie? This means we would know that when engaging in a game of make-believe with our father, we know it is part of the set of rules of the game that we can run away, while when watching a movie, we know we are to stay put. But if that is the case, how come we do respond to being scared by a movie by for

example grabbing hold of our girlfriend's hand and putting it in front of our face? It seems that these rules are rather opaque and dependent on context. This means we are asking a lot of poor Charles. On one hand he must act according to implicit and opaque rules that nevertheless he manages to obey, while his own first-person report would state that he is actually afraid and not engaging in a game of make-believe (unlike the girl running away from the girl-eating troll). So his actions are not only not in line with his own first-person report, they also follow context dependent and implicit rules that govern his responses.

Walton points out that we need not be aware of the rules of the game in order to play it, and that for example the viewer of a horror movie needs not be critical enough to be able to recognize in himself that his emotion is make-believe. However, according to Walton, Charles can be made to realize that he is quasi-afraid, rather than actually afraid. Charles' "realization that fictionally he is afraid is based largely on introspection... Mistakes seem out of the question (assuming he has mastery of the relevant principles of make-believe)" (A. Neill & Ridley, 2013, pp. 246-247). Säätelä points out that this is exactly what is the problem here: Charles could still claim he is afraid, thus *not* coming to this realization through introspection, and the "relevant principles are those that Walton's theory imputes to him" (Saätelä, 1994, p. 29). Carroll rightly thinks it "strains credibility" to think we are not just unaware of some of the rules, but we are unaware of playing a game altogether (Carroll, 1990, pp. 74-75).

Something that is generally overlooked in the literature looking at the paradox of fiction is that these behavioural disanalogies are also present in reality. When we watch the news and see footage of someone being in a car accident, we don't get up and help, as we would when we saw it happening out of our window. This seems so obvious, we don't consider the reasons behind it. *Of course* we don't get up: we know it is happening somewhere else and we can't help. But the same goes for fiction. When I read about a car accident, I don't get up, because I know it is happening elsewhere – in this case a fictional elsewhere – that I cannot reach.

But what about if I see a murder happening in a play? I could easily get up, run on stage, and stop one actor from stabbing another.

Walton explains this through “quasi-emotion”, but this is an unnecessary move. I could still feel real emotion when watching the murder take place, yet have the (true) belief that the murder is happening elsewhere: it is happening within the fictionalized narrative I see on stage, of which *I am not part of*. My world and the world on stage don’t overlap and I know this. This is why I know my actions would be futile. In a way, even if I were to run on stage to stop the murder, I know that the outcome doesn’t change. Just as when I stop reading a book because my emotions are so strong about something I fear might happen, I know the story doesn’t stop. My stopping reading does not change the narrative that has already been written down, and I take that as an obvious fact about fiction. It is there despite my action, simply because my actions have no consequence in the world in which the story takes place. Precisely because it is fiction, it is evident that any action is futile. When watching a news story about a horrible drought and the ensuing famine in Africa, I might be moved to start collecting money (notice how I am not moved to try to put food into the television). Here my action might have causal effect, and I know this. In fiction my action won’t have causal effect and I know this. Unlike what Walton says, I am not a participator in a game of make-believe, when reading fiction. There is no participation on my part possible, because the narrative I am consuming takes place in a different world than I am in. Furthermore, the narrative is already set, and I know it. It is not within my capabilities to change the outcome.

This could also explain part of the paradox of tragedy, mentioned before. We can enjoy a tragedy, despite the suffering we witness and the emotions we might feel, because our watching has no actual bearing on the narrative of the tragedy. No action is required of us, because no action *can* be done. This is also true for actual events and helps explain why we turn on the television every night to watch the 8 o’clock news, despite the fact that it confronts us with all the suffering in the

world. It is suffering that has already happened, and suffering that has happened in a place outside of our (perceived) causal influence. I would venture we would much less enjoy watching a news story featuring live images of something suffering a heart attack right in front of our door, with bystanders unskilled in CPR just watching, and remain sitting on our couch, despite knowing that we are capable of performing CPR. And we would much less enjoy it *because* it is within our reach to change something about that situation. We can no longer be spectators on the couch, it would require us to get up. Then imagine the same news story with live footage of the same thing happening on the other side of the globe. We would be horrified. We would desperately hope that someone does something. And then, when the person would have died, we would turn off the television and comfortably eat our dinner.

According to Walton the quasi-fear Charles experiences is a state that comprises both physiological aspects (such as increased adrenaline, with its physiological consequences) and psychological aspects (the sensations of those physiological consequences). Walton first points out that “it is partly the fact that Charles is in a state of quasi-fear, the fact that he feels his heart pounding, his muscles tensed, etc., which makes it make-believe that he is afraid” (K.L. Walton, 2014). But then the question is what this account does to help solve the paradox. It just postulates an extra kind of emotion that in its physiology is identical to the actual emotion. What then does the difference actually consist in? If it is merely that this is the type of emotion that we experience in a game of make-believe, as opposed to the type of emotion we experience in response to ‘real’ situations, then we are begging the question. The same goes for pointing out that it is the causality that differentiates between the two: a ‘real’ emotion is caused by a true belief and a ‘quasi-emotion’ is caused by a make-believe belief.

In a footnote, Walton himself acquiesces that perhaps the purely physiological aspects of Charles’ reaction are not part of the quasi-fear, and so quasi-fear should be understood as referring to only the psychological aspects of Charles’ condition.

But if we leave out the physiological aspects of emotion from the state Walton calls quasi-fear, we are still lacking an explanation for this physiological response. It cannot be a reaction in response to Charles' true belief that the green slime is coming towards him, so it must be a direct reaction to the imagery of the green slime. But if the reaction is direct and not mediated by a cognitive process, aren't we moving close to a feeling-centred theory of emotion? This also leaves unexplained what role the physical response Charles has to the green slime plays in his game of make-believe. It seems that Walton acknowledges that the physiological aspects are prior to the quasi-emotion and thus prior to the game of make-believe Charles engages in. But how would he then describe this physical reaction? It seems incontrovertible that the physical response is at least part of what we deem an emotion to be.

Some philosophers have pointed out that there is a certain "startle effect" in films: an involuntary reaction to visual stimuli that is immediate and hence does not depend on any existence belief about what we are seeing. As the startle response is so direct, it seems unlikely that any cognitive process is at the basis of it. Simo Säätelä makes use of one of Walton's own examples to make this point. Walton quotes Charles Darwin putting his face against the glass of the cage of a puff-adder in the zoo "with the firm determination of not starting back" if the snake were to try and bite him. However, "as soon as the blow was struck, my resolution went for nothing, and I jumped a yard or two backwards [...]. My will and reason were powerless against the imagination of a danger which had never been experienced" (K.L. Walton, 2014, p. 260). Where Walton accepts that Darwin in making-believe that the snake is dangerous, Säätelä rightly points out that it is hardly this fictional fear that makes Darwin jump back. Rather it is an involuntary reaction (Alex Neill, 1991; Saatela, 1994). The same "startle effect" could be applicable in the case of Charles, who is startled by the green slime on the screen. Alex Neill notes that "as reactions, these facts [...] do not require an explanation in terms of making-believe, any more than they do in terms of belief proper" (Alex Neill, 1991).

Despite that this startle effect could perhaps explain the emotion of fear in response to fiction, it cannot explain our experience of other emotions such as grief, pity, exultation, etc., as Walton also point out (K.L. Walton, 1990). It seems undeniable that what we consider a full-fledged emotion comprises both the physiological as well as the psychological aspects of emotion. So a theory of emotion should be able to account for both of them. Walton's account falls short in doing so.

Another important objection against Walton's theory is that it does not seem to agree with the phenomenology of what we experience in the face of fiction. Because we do believe we feel genuine fear when we watch a scary movie, or genuine pity when Anna Karenina feels driven to commit suicide. One issue is the question of free will, as Noel Carroll points out in *The Philosophy of Horror* (Carroll, 1990). If the fear we feel is only make-believe fear, just as the child feels for her girl-eating troll of a father, it would seem we would be able to turn it off and on at will. After all the child is capable of 'stepping out' of her make-believe fear to beg her mother to let the game continue. But it does not seem as though we can turn off our fear as easily when engaging with fiction. Glenn Hartz shows this through the following example:

"My teenage daughter convinces me to accompany her to a "tear-jerker" movie with a fictional script. I try to keep an open mind, but find it wholly lacking in artistry. I can't wait for it to end. Still, tears come welling up at the tragic climax, and, cursing, I brush them aside and hide in my hood on the way to the car. Phenomenologically, this description is perfectly apt. But it is completely inconsistent with the Make-Believe Theory, which says emotional flow is always causally dependent on make-believe. [H]ow can someone who forswears any imaginative involvement in a series of fictional events respond to them with tears of sadness?" (Hartz, 1999)

It is also important to recognize that apparently we don't engage emotionally with just any fiction. There are interpersonal differences and contextual circumstances, but generally it seems we agree that some fiction is easier to engage emotionally with than other fiction. Many people cried or at least felt very sad when Dumbledore died in the Harry Potter books, but presumably not many people felt fear when watching the actual film *The Green Slime*. As Carroll points out: "if the response were really a matter of whether we opt to play the game, one would think that we could work ourselves into a make-believe dither voluntarily. But there are examples [of fictional works] which are pretty inept, and which do not seem to be recoverable by making believe that we are horrified. The monsters just aren't particularly horrifying, though they were intended to be" (Carroll, 1990, p. 74). On the other hand, I could not easily will myself to not be affected when watching a horror movie, like a switch I can turn off or on. The girl-catching game can be stopped when the girl in question is tired, but it seems unlikely that we are able to do the same when watching the exorcist. So there seems to be an obvious disanalogy between the paradigmatic games of make-believe that children play and the supposed game of make-believe we play in the face of fiction.

Alex Neill thinks Carroll misunderstands Walton's argument to mean that Charles only *pretends* to be afraid. According to Neill, on Walton's account Charles can actually believe he is afraid. He argues that "Walton labels Charles' physiological/psychological state "quasi-fear" to mark the fact that what his feelings and sensations are feelings and sensations *of* is precisely what is at issue." He continues saying that on Walton's view we can actually be *moved* by works of fiction, but is make-believe that what we are moved *to* is *fear* (Alex Neill, 1991, pp. 49-50). Here we return to the problem stated above: if this is the case, what is then actually the difference between a quasi-emotion and a 'real' emotion if the physiological state is identical and the subject believes the emotion to be genuine? The only difference seems to consist in that the object of the 'real' emotion is real and of the quasi-emotion is fictional. But this is the problem we were aiming to

solve, and so Walton is begging the question. He posits a new type of emotion which only difference from what we would normally consider an emotion is that it deals with fiction. Then he argues that because it deals with fiction, it must be different from normal emotion. And if I am just generally unaware of playing make-believe and I feel as though I am actually afraid, and I am not aware of the rules and cannot will myself to be affected or disaffected, what does the actual difference between make-believe emotion and actual emotion consist in? Isn't it then a merely logical postulation, only different in name, that solves the problem?

3.2 Object of emotion

Another important issue is that in Walton's account it is Charles who fears for his own life when the green slime is coming towards him. This seems different from Charles pitying Anna Karenina or him rooting for Harry Potter. In the former case, he is the object of the emotion, in the latter ones the object of the emotion is the fictional character. Usually, when we experience fear while watching a movie or reading a book, we do not fear for ourselves, we fear for the character in the horror picture who unwittingly goes into the house in which we know the axe murderer to be hiding, or the lovely couple who are soon to be threatened by a tornado rushing their way. We do not believe that the axe murderer will kill *us* or the tornado will swipe *us* up, rather it is the characters for whom we fear. It is not unthinkable that Charles' fear when he sees the green slime coming towards him on the screen is not because it is coming towards *him*, but rather because it is coming towards the character for whom he feels and through whose eyes he is now watching the slime creep towards him (let's call him John).

Hence there is no need at all for Charles to make-believe that the slime is coming towards him or that it is threatening to kill him. To use Walton's terminology, Charles would be making-believe that he fears for John, who (within the fictional operator) actually believes that the slime is coming towards him and that it is threatening to kill him. But it makes even less sense to assume that Charles would

be making-believe that he fears for John than for himself. The analogy with the make-believe game of the girl and her father completely disappears. The girl is *actually* making-believe that the girl-eating troll will eat her. She is not making belief that the character of the girl, that she is playing, is on the verge of being a girl-eating troll meal. That would be introducing another fictional layer where it is not necessary and furthermore, does not have any explanatory power.

The same goes for the pity we feel for Anna Karenina, albeit in a slightly different way. We feel pity for Anna Karenina because of what has happened to her within the fictional scope. According to Walton, I make-believe that I feel pity for Anna Karenina. Here Anna Karenina is indeed the object of my emotion, unlike Charles. But there is an important difference. To use Neill's example: suppose that you walk right through a child's game of making-believe that the mud-pies are actual pies and you step right on a carefully crafted mud-pie. Johnny starts to cry, because you have destroyed his make-belief pie. He is make-believedly distressed (Alex Neill, 1991, p. 51). But the pity you feel in actual pity, because you pity Johnny, feeling his make-believe distress. There is no reason for you to make-believe the pity: the object of Johnny feeling make-believe distress is real after all (plus the fact that you caused it). The same goes for Anna Karenina. I don't have to make-believe my pity. Anna Karenina fictionally experiences all sorts of woes. My pity is *external* in the sense that it was external when I pitied Johnny as opposed to Johnny's *internal* distress within his game of make-believe. Walton, in his example of Charles, unfairly displaces the object of the emotion from the fictional character to the consumer of fiction.

Of course, this still leaves us with the question: how is it possible that my pity has as an object a character who experiences awful things within a fictional scope, or in other words: how can it be that I pity something that does not exist?

4. Cognitivism and the irrationality of fictional emotions

Before dealing with the coordination condition in the next chapter, it is important to elaborate on the underlying theoretical assertion of the coordination condition of the paradox of fiction. Because this assertion stems from the cognitivist theory of emotion, I will start this chapter by providing a global background of this theory, the objections and its main rivals.

Then I will return to the paradox of fiction and its treatment by Colin Radford, who initially stated the paradox. I will deal especially with his solution of the paradox, namely that the emotions we feel for fictional characters are real but irrational. I will discuss what is meant for emotions to be rational or irrational, or for them to be appropriate or inappropriate and what this says about the cognitivist theory of emotion.

4.1 Cognitivist theory of emotion

The cognitivist theory of emotion classifies emotions as cognitions, as they are mental states in which the subject is cognizant of some object of that emotion. The emotion process involves the manipulation of information or of the stimulus, making it a cognitive process. These theories have in common that in order to have an emotion, a propositional attitude is needed in the form of a belief or evaluation. Emotions can be seen as special kinds of judgements (Carroll, 1990). Cognitivism is sometimes called judgementalism, because according to this theory a necessary part of experiencing an emotion involves having the belief or judgement that the object of the emotion instantiates a certain evaluative property. For example, to say 'I fear that dog' would necessarily involve the judgement or belief that I am in danger from that dog. From this the commitment to existence-beliefs follows, as the belief of being in danger from the dog seems to entail the belief that the dog exists.

Cognitivist theory came about as a reaction to feeling-centred theories of emotions. Such theories, such as the James-Lange theory (James, Haupt, & Lange, 1922), take bodily sensations to arouse certain feelings which constitute emotions. In other words, emotions are feelings caused by physiological conditions. One of the main criticisms on such feeling-centred theories is that they cannot deal with the supposed rationality of emotion. I will return to this criticism later in this chapter when discussing Radford's solution to the paradox.

Another major criticism is that they do not adequately deal with the intentional nature of emotion. Feeling-centred theories assimilate emotions to bodily sensations. However, we speak coherently of an emotion being unreasonable or inappropriate, and by extension, sometimes reasonable or appropriate. As Deigh explains: "anger can be unreasonable when it is misdirected; fear can be irrational when its object is innocuous. By contrast, it would be gibberish to describe a bodily sensation, a toothache, say, as unreasonable." (Deigh, 1994). So this rationality, or reasonability is part of emotions, which cannot be captured by feeling-centred theories. On any cognitivist account this is unproblematic: emotions are the logical outcomes of the beliefs that together form them, so describing them as irrational or rational can be done by assessing the underlying beliefs as rational or irrational. This means that there must be some property of the intentional object that constitutes the rationality of emotion that it causes. I will return to this point shortly.

The problem cognitivism sees in feeling-centred theories is that they seemingly do not recognize the intentional nature of emotion; the idea that emotions are about something or directed at something. It seems clear there is an intentional nature to emotions. When I am afraid, I am perhaps afraid of falling off the Empire State Building; my fear is directed at the height I am facing. When I am gleeful, it is about the fact that I finally finished my thesis. In other words, emotions are always about something; they have intentionality that cannot so easily be ignored within a working theory of emotion. Bodily sensations are just that, they are a

collection of activated neurons recognizing the warmth of the sun or the cold of ice, but they are not *about* the sun or the ice, they are not directed at them. Thus they are not intentional states.

For the cognitivist, this intentional character needs not to be aimed at an object *per se*. I can be afraid of a snake, because I think it is venomous and will bite, when in fact it is a completely harmless snake, who is actually rather afraid of me. The belief that the snake will bite is enough for me to experience the emotion of fear, as long as I truly believe that the snake is venomous.

However, the idea that intentionality is a necessary part of emotion, creates a problem for cognitivism as well. It seems that not all emotions have intentionality. We seem to be able to be anxious or depressed or happy, without being anxious or depressed or happy *about* anything. Perhaps we could concede that typically emotions have intentionality, though it is not a necessary condition for them. Cognitivists usually reply to this objection by either deeming these emotions without intentionality as not actual emotions but as ‘moods’, or by attributing a subtle or suppressed intentionality (Deigh, 1994). In the case of emotions we have in response to fictional characters or events, this objection is not particularly relevant, as the whole paradox arises precisely because there is an intentional object and it does not exist.

Another common objection to cognitivism is the ‘fear of flying’ objection. This objection states that propositional attitudes are neither necessary nor sufficient for the existence of an emotion, as I might well be aware that flying is the safest means of transport and at the same time be mortally afraid of actually stepping into an airplane (Stocker & Hegeman, 1996). The problem here for the cognitivist theory of emotion lies in the discrepancy between the intentional content of the belief and the action that follows from it, and thus in the apparent irrationality of the action of actually stepping into the airplane. This is a more interesting objection when thinking about the paradox of fiction as there seems to be a

similar discrepancy at work when Charles is watching the green slime come towards him.

Cognitivists might respond that this example merely shows that the propositional content of emotion and the propositional content of belief are different. It doesn't show that emotions have no propositional content at all. Nonetheless, it still seems the case that in the example of fear of flying the emotion cannot be assimilated to the underlying belief, which is problematic for cognitivist theory. This is also what could be the case for emotions in response to fiction.

However, the objection is strange when looked at more closely. The content of the salient belief that I have when I fear flying is not that flying is not the safest means of transport *in general*, but that I believe that *I* could die when I fly. More to the point it could be that it is a particularly gruesome way of dying, where I slowly choke to death when falling 10 km from the sky. Alternatively, it could be that perhaps I have the false belief that when I drive, I am safer than the safest means of transport. The actual chance of me dying does not feature heavily in my belief. I can fear brain cancer more than obstructive lung disease, with epidemiology arguing against me, but my fear is not about the epidemiology of these diseases, but about something else, which in this case might be the perceived results of getting brain cancer, versus a lung disease I know relatively little about. So the underlying belief is actually compatible with the emotion experienced.

This does call into question to what extent the belief, which according to cognitivism is a necessary condition for emotion, needs to be a true belief and furthermore it invites us to judge the underlying belief in terms of coherence, consistency, rationality and appropriateness. Cognitivism seems to propose that the belief underlying the emotion meets these criteria in order to be 'real' or 'a true emotion'. Through Radford's own solution I will now look at this idea.

4.2 Radford's conclusion: the irrationality of fictional emotions

What does it mean for emotions to be rational? And a further question, what does it mean for emotions to be appropriate? Radford starts his illumination of the paradox of fiction with an example:

“Suppose then that you read an account of the terrible sufferings of a group of people. If you are at all humane, you are unlikely to be unmoved by what you read. The account is likely to awaken or reawaken feelings of anger, horror, dismay or outrage and, if you are tender-hearted, you may well be moved to tears. You may even grieve. But now suppose you discover that the account is false. If the account had caused you to grieve, you could not continue to grieve. If as the account sank in, you were told and believed that it was false this would make tears impossible, unless they were tears of rage. If you learned later that the account was false, you would feel that in being moved to tears you had been fooled, duped.” (Colin Radford & Weston, 1975)

He says this example shows that “...the possibility of your being harrowed again seems to require that you believe that someone suffered.” Which is what gets us into trouble: “If this account is correct, there is no problem about being moved by historical novels or plays, documentary films, etc. For these works depict and forcibly remind us of the real plight and of the real sufferings of real people, and it is for these persons that we feel. What seems unintelligible is how we could have a similar reaction to the fate of Anna Karenina, the plight of Madame Bovary or the death of Mercutio. Yet we do.”

He then systematically tackles the response condition, the belief condition and the coordination condition, arguing for and accepting all of them. In his argument for accepting all the conditions, his background assumption about rationality shines through. For example, Radford demonstrates the need for the response condition by stating: “It would seem then that I can only be moved by someone's plight if I believe that something terrible has happened to him. If I do not believe that he has not and is not suffering or whatever, I cannot grieve or be moved to tears” (Colin

Radford & Weston, 1975, p. 68). We can interpret this statement as him saying that we can only be *rationally* moved by someone's plight if we believe that something terrible has happened. Otherwise, our grief or our tears are not rational.

Radford's conclusion is that "our being moved in certain ways by works of art, though very 'natural' to us and in that way only too intelligible, involves us in inconsistency and so incoherence". He argues this on the grounds that he accepts that we are in fact moved by Anna Karenina's fate, that existence beliefs are necessary to be moved by her fate, and that such beliefs are lacking in the case of us reading the novel. In other words, he accepts all three conditions and thus the paradox and solves it by stating that the only conclusion is then that the emotions we do feel are not rational.

4.2.1 Rationality of emotion

In accepting our emotions for fictional characters to be irrational, Radford highlights another assumption behind the cognitivist account of emotion, namely that emotions are or should be rational, or on associated notions, coherent and consistent. For Radford's response to be properly understood, we first need to clarify what it means for an emotion to be rational. The aim of this chapter is to discuss through Radford's work this notion of rationality and the pertinence to this thesis. Radford accepts all three premises of the paradox, which we shall not do and I will not go into his arguments for accepting them.

The question of rationality is in fact a normative question. We talk about a rational emotion when it results in the appropriate response in a certain situation, the norms for which are inseparable from social norms. When we talk about the causal chain of emotions we need to be clear about which part we mean when we talk about rationality. First in the chain is the belief, which could be rational or irrational according to some yet to be determined standards. Next is the cognitive evaluation of the belief, which could also be subject to irrationality. Then there is

the emotion the belief gives rise to. And ultimately the emotion leads to action, rational or otherwise. Radford places the irrationality at the point in the cognitive process between belief (that we know Anna Karenina does not exist) and emotion.

Emotion is one of the main mechanisms by which we direct and maintain attention. In doing so it aids us in our decision process by rendering a relatively small number of options salient, which helps us create a framework in which to make rational decisions. An important question here is whether we are to think of emotion as an aid to rationality or if we can think of emotions being rational in and of themselves.

As mentioned above, there are several steps at which the causal chain in which emotion plays a part can be judged rational or irrational. The first step is belief. Rationality implies that one's beliefs conforms to one's reasons to believe. To be able to apply the judgement of whether an emotion is rational or irrational, a cognitive element (such as belief) as part of emotion is required, which can be consistent or inconsistent with reason, or on a more instrumental note, with the actions we take. When we take emotion to be purely a physiological response to a situation, without a cognitive element such as belief featuring in it, it becomes incoherent to speak of irrationality. For example: when I am angry at my boyfriend for leaving his clothes on the floor, how can we judge the physiological response I have – the biochemical reaction that is autonomous, so beyond by conscious control – as rational or irrational? It merely *is*, no further judgements are applicable to it. Only when a cognitive element is added – e.g. I have the belief that just last week we agreed after a long fight that he would pick up his clothes from now on – it is appropriate to judge my anger as rational. As such, a cognitivist theory of emotion at least allows for the judgment of rationality or irrationality.

To assess emotions for rationality in themselves, we have to look at the instances where we would judge emotions to be irrational. If I say my friend is irrational in

his anger at me for being late, what I am really saying is that the anger he feels is inappropriate. Rational emotions in this sense are very much context related and have an intrinsic intentional character. My friend is angry *about* something, and in deciding if his anger is rational, I take into account the emotion and the intentional context. This is also the intentional character that the cognitivist theory tries to capture. A cognitivist theorist will say that a reasonable emotion is one whose constituent propositional attitudes are reasonable (De Sousa, 2014). But this is only a necessary condition, not a sufficient one. At least part of the appropriateness of an emotion is the step from belief to emotion. It seems rational for me to be angry at my boyfriend for flaunting the agreements we made. It seems less rational for me to be so angry I cannot work for days, because I'm so full with it. There needs to be some proportionality to the emotion, or in other words, the underlying beliefs and the emotion need to make a coherent whole. When thinking about emotion in fiction, this means it could be rational to cry for Dumbledore when he dies, but less rational to go in deep mourning for 3 months and only wear black. Radford is on the right track when explicating the rationality of emotions, but he misplaces the irrationality in them being a response to fiction. Instead we need to look at the appropriateness of the emotion, which I will do shortly.

One step further down the line we look at the behavioural consequences. Coherence and consistency in the sphere of belief are notions closely associated with rationality as well as the outcome of an action. Optimisation of this outcome is seen as rational, whereas non-optimal action – action that is non-optimal in light of the available background information – is deemed to be irrational. This is interesting in relation to the idea of the behavioural disanalogies we display in response to fiction. As we noted in chapter 4, there seems to be a behavioural disanalogy in consuming fiction. When we are watching *Deep Impact* and the flood is rolling towards us and we do not get up from the couch and run, which presumably we would be wise to do if there was an actual flood, but as discussed in chapter 4, the actual emotion of fear when seeing the approaching flood is not

for ourselves, it is for the small figures of the father and daughter standing on the beach, about to be drowned. So contrary to what Walton's claims, there is no behavioural disanalogy and this cannot serve as an argument for why our emotions are irrational.

4.2.2 The appropriateness of emotion

When we speak of the appropriateness of emotion, rather than rationality per se, the concept of coherence and consistency are much easier understood. The emotion itself needs to be coherent to the context it presents itself in and needs to be consistent over time, at least within the person experiences that emotion. This means that the underlying beliefs need to be coherent and consistent in order for the emotion to be appropriate and rational. So rationality extends further down the causal chain of emotion and coherence and consistency are standard by which the causal chain is judged. This also shows that to a great extent the appropriateness of emotion is tied to social norms.

Emotions have so-called correctness conditions. Fear is correct if and only if what is feared is dangerous (Goldie, 2009). Of course, as I showed before, the second part of this correctness condition is problematic, as we have seen in the case of phobic fears. We apply an extrinsic normativity to emotion in the sense that we seem to assume there are absolute standards by which we can say kittens are not dangerous and thus my fear of kittens is irrational, even though I personally would disagree. This makes the emotion experienced dependent on the object and thus problematizes emotions felt about fictional objects, though it could be argued that as we, as a society, unproblematically describe fictional emotions as real emotions, on the basis of their phenomenology, we then unproblematically accept that those emotions are 'correct' or rational. The idea of correctness conditions is part of the intentionality of emotions as mentioned before.

Joyce argues convincingly that while phobias are a paradigm case of irrational emotion, they are not, as Radford argues, similar to the emotions we experience when reading (looking at) fiction. The irrationality of phobias is placed by Radford at the supposed irrationality of emotion, because it lacks consistent beliefs: the person mortally afraid of flying knows it to be the safest means of transportation. Joyce argues that this person does in fact have the belief the plane is going to crash and it is this belief that is irrational, as it is held in the face of solid discrediting evidence. Whereas in the case of reading fiction, Radford argues there is a lack of consistent beliefs. This means that the presumed irrationality of phobias is not an argument for the presumed irrationality of emotions held in the face of fiction (Joyce, 2000).

Cognitivism appears not to be committed to emotions being necessarily rational. It requires an emotion to have constituent propositional attitudes, but they don't need to be reasonable per se. It posits that there is a cognitive process, presumably in the form of some sort of judgement. In the above example where I am angry at my boyfriend for leaving his laundry scattered on the floor, cognitivism will not deny that what I am experiencing is an emotion. First we have my physiological response to the context (me stepping on my glasses because I didn't see them with a men's T-shirt on top of them) and there is a judgement or belief about this context (taking into account last week's fight). So far cognitive theory has no problem in identifying my anger as emotion. Presumably it is even rational emotion. But the rationality (or justifiableness) of the emotion becomes questionable when just yesterday in a mood of defeatism, I told him I would pick it all up, because he was just never going to do it. And it becomes even more unclear when he believes (out of some sort of religious upbringing) that in a relationship, a woman is responsible for all household chores, whereas I believe men and women should share their load. From his point of view, my anger is irrational (I should have known when entering into a relationship with him that this was how it was going to be), while from my point of view my anger is very

rational indeed, thank you very much. It is not clear that cognitivism would deem my anger in these border case of rationality as something else than emotions.

As mentioned before, one of the presumed benefits of cognitivism above feeling-centred theories of emotion, is its capacity to deal with the supposed rationality of emotion. As we have seen above, cognitivism does provide a framework for evaluating the rationality, and rationality or appropriateness do seem an important part of what an emotion is. On this theory we have emotions, which have a cognitive element, allowing us to judge the rationality. Cognitivism doesn't deny the existence of irrational emotions, and as such, Radford can judge our emotional response to fiction to be irrational, while at the same time preserving the coordination condition and with it, cognitivism.

Radford, by accepting the coordination condition, changes it from an existential condition to a normative condition: "We do not *rationally* feel fear unless we believe ourselves (or someone actual) to be in danger" (Radford & Weston, 1975). In doing so Radford essentially changes the paradox to:

- 1.a. We are *rationally* moved by the career and fate of Anna Karenina
- 2.a. Anna Karenina does not exist, and we know this.
- 3.a. Being *rationally* moved by the career and fate of a subject requires belief in the real existence of that subject; it is impossible really to *rationally* care about something that one knows does not exist.

So ultimately Radford does reject the response condition and changes it into: We are *irrationally* moved by the career and fate of Anna Karenina. This makes it clear what the importance of the concept of rationality of emotion is when trying to decide how to tackle this paradox. However, the paradox in the formulation above still stands and the problems with denying the response condition as mentioned in the previous chapter are just as problematic for this rephrased response condition as they were for the original condition.

Along the way Radford makes some interesting observations. One of them is the seeming difference between the emotion we experience when a friend dies, and the emotion we feel when Mercutio dies on stage. We cry a bit when Mercutio is murdered, but otherwise still enjoy our evening at the theatre, whereas when our friend dies our grief is of a whole different magnitude. As Radford says: we are moved by Mercutio's death, but do not grieve him. The point he makes is that, even though the order of magnitude is different, we do indeed feel for Mercutio; we have real emotions, and real tears, despite their irrationality.

This is the emotional disanalogy as I mentioned in chapter 2 as expressed in the paradox of tragedy that Hume had already pointed out:

“It seems an unaccountable pleasure, which the spectators of a well-written tragedy receive from sorrow, terror, anxiety, and other passions, that are in themselves disagreeable and uneasy. The more they are touched and affected, the more are they delighted with the spectacle; and as soon as the uneasy passions cease to operate, the piece is at an end.” (Hume, 1757)

There are a couple of problems with the distinction Radford makes between our emotion to a real-life event and our response to a fictional one. For one, it is not clear at all that in this example both events are comparable. A friend that I have known for years, is not a stage character that I interact with for some hours on one night. My response to the death of a vague acquaintance could very well be the same as my response to Mercutio's death in terms of order of magnitude. When we watch the evening news, we can be moved, but do not grieve (if we need to make this distinction). So the difference could very well consist in our degree of separation from the person/character, rather than in the fact that one is fictional and one is not.

Another aspect he shows through a pithy remark of O.K. Bouwsma, who says: the cigarettes and chocolates go in our mouths too, and we may mutter, if not to each other, then to ourselves, 'How marvellous! How sublime!' and even 'How moving!' (Bouwsma, 1965). This is of course a point that in philosophy is

researched through the related paradox of tragedy that Hume mentions in the quote above: How can we feel a painful emotion, like grief, when watching a play, and still enjoy our evening? Without going into that specific debate, it is important to ask if this is truly such a problematic combination of emotions. When I go to a family member's funeral, say my grandfather, we can go out to dinner with our family and, despite grieving my grandfather, also enjoy being together. Perhaps all this points to is that humans can experience multiple emotions at the same time, or at least at shortly subsequent times. Furthermore, is there truly a difference between fiction and reality in this respect? 'Why do people enjoy watching scary movies?' is in effect the same question as 'why do people watch the evening news?', 'why do people watch reality TV?', 'why do people slow down to watch the traffic accident on the other side of the road?'. It is an interesting question, but not a question particular to fiction in the least.

Another interesting way of looking at fictions is that of fiction as an instructional endeavour, allowing us to experience emotions. Feagin, when talking about the pleasure of tragedy, poses that the pleasure we take in it is actually pleasure about the 'meta-response' we are having. I pity Anna Karenina and then feel subsequently pleased at myself for having the appropriate emotion (Feagin, 1983). This meta-response is part of the instructional character of emotion that Joyce also sees as a large part of the value of art (Joyce, 2000). But it could well explain why we also enjoy watching the car wreck on the highway in real life, or at the very least enjoy satisfying our curiosity as to why the traffic is slowing down so much, but also slowing down to see exactly how many cars are turned upside down. Here again we encounter the importance of having the *appropriate* emotion. I will return to this issue in chapter 6 when I will investigate the difference between experiencing emotion in response to fiction and in response to the evening news or some other narrative that we view as real.

To return to what the underlying idea of the appropriateness of an emotion consists of, this becomes clear in the case where I try to persuade my friend that

his emotions are irrational. For example, he is deadly afraid of stepping in the car with me when I am driving. As his reason he states that he thinks I will drive off a bridge. I will try to persuade him by pointing out that his belief that I drive off bridges is unfounded, as so far I have never even come close to driving off a bridge, only having been in fourteen car crashes so far, none of which involved bridges (although we can argue about how wise it is to bring up the latter point). In other words, I challenge the beliefs on which he causally bases his emotions, pointing out that those beliefs are false and thus irrational and hence his emotion is irrational. Here we also see that it is also dependent on the person propounding the belief. We can have a thesis filled with discussion on who is irrational in this discussion, just as we could about the discussion between me, my boyfriend and the laundry. Is there some external standard that dictates the amount of car crashes I can cause before I am no longer allowed to call myself an excellent driver?

So what *is* the intentional object that our evaluation of appropriateness is then aimed at? Just as we cannot talk about appropriateness of a biochemical reaction that is autonomous, we cannot talk about appropriateness of the intentional object, in the form that cognitivism wants us to. They are committed to the intentional object merely existing, making a judgement about the appropriateness of its existence nonsensical. It is the meta-response or thought about the intentional object that allows us to evaluate appropriateness. Cognitivism thus misses an important step in explaining how it is we have the judgements of rationality and appropriateness. When it is no longer Anna Karenina, or my sister, that is the intentional object, but rather the meta-response or thought about Anna Karenina or my sister, this judgement is much easier to explain. I will return to this issue when discussing thought theory in the next chapter.

5. Denying the coordination condition

Being moved by the career and fate of a subject requires belief in the real existence of that subject; it is impossible really to care about something that one knows does not exist. (Coordination condition)

The most common approach to the paradox is to reject the coordination condition, which states that being moved by a fictional character or event requires belief in the real existence of that subject. In the previous chapter I discussed how Radford, in this attempt to show that emotions in response to fictional characters are, though real, irrational and undesirable, actually changed this condition into one including rationality of emotion. In this chapter I will show various ways of dealing with this condition, and show that thought theory is the most promising and can actually illuminate why the paradox is so important to our understanding of emotion, not just in fiction, but also in general.

One option in rejecting the coordination condition is to deny that any higher-order cognition is necessary for emotions. This is the route taken by feeling-centred theories of emotion. Without any cognition and thus intentionality the ontological status of the object of the emotion becomes irrelevant. The problem with this I mentioned in chapter 4 when dealing with various theories of emotion.

Another, much more promising, option is the one that modifies cognitivist theory so that belief is no longer a necessary condition for the experience of emotion. ‘Thought theory’ holds that the contents of one’s thoughts can cause emotions. The difference between emotional response to reality and to fiction is that in the case of fiction we ‘mentally represent’, to use Peter Lamarque’s terminology (Lamarque, 1981). We do not actually have to believe in the existence of the fictional character or event, rather, in the case of fiction, the intentional object of our emotion becomes the mental representation or thought of it. In making this

move, thought theory effectively denies the coordination condition, while upholding the cognitive component of emotion.

Thought theory also poses that, although our emotional responses to actual characters or events may require belief in their existence, there is no reason to think that our emotional response to actual characters or events are a paradigm case for emotional response in general. The emotional response to reality is on this account a special case of a general theory of emotion, rather than serving as a model for it. In this sense, thought theory presupposes that our emotional response to fiction is real (and rational) and moves the emotional response to fictional characters and events to the centre of their theory. It reverses the role of reality as the paradigmatic case of emotion and fiction as a special case.

Carroll makes the useful distinction between thought and belief: “Thought here is a term of art that is meant to contrast to belief. To have a belief is to entertain a proposition assertively; to have a thought is to entertain it nonassertively. Both beliefs and thoughts have propositional content. But with thoughts the content is merely entertained without commitment to its being the case; to have a belief is to be committed the truth of the proposition.” (Carroll, 1990). Thought theory often uses (phobic) fears as an example. Carroll gives the example of standing over a precipice. We might entertain the thought of falling and have a tingling feeling of fear, though we do not necessarily believe that the chances of our actually falling are very great. We could be standing on the Grand Canyon Skywalk, where the floor is made entirely of glass and the ground is 600 meters below and experience this feeling of vertigo, though we are quite secure in the constructional prowess of the engineers who build the skywalk. We are “not frightened by the event of our thinking of falling, but by the *content* of our thought of falling”.

One of the main critiques of thought theory is that there is no real object of the emotion (Lamarque, 2003). The fear we experience is according to thought theory caused by the thought of slime, but the fear is not *of* the thought. It is the thought

of monsters that I am scared *by*, not the thought of monsters that I am scared *of*. The content of the emotions is ‘fear of the slime’ (Levinson, 1996; K.L. Walton, 1990). One could say the object of the fear is the imagined green slime and we speak of the intentional characterization of the fear (Lamarque, 1991). Or say that although it is true that the Egyptians worshipped Osiris, there is nothing such that the Egyptians worshipped *that* (Lamarque, 2005). Joyce discounts this discussion, mainly because it cannot so easily be extended to other emotions. When we pity, what is then analogous to *by/of*?

We can ask what exactly we are sad *about*. Radford asks the following questions in order to highlight the mysterious nature of our emotional responses to fiction: "We are saddened, but how can we be? What are we sad *about*? How can we feel genuinely and involuntarily sad, and weep, as we do knowing as we do that no one has suffered or died?" (C. Radford, 1977, p. 77). In this sense thought theory is still liable to the accusation of irrationality. The blatant irrationality of both believing and not-believing that Anna Karenina is real is dissolved, but in doing so it has made us liable to a new kind of irrationality, namely that we are moved by mere thoughts.

Joyce remarks that strictly speaking embracing thought theory means that “one does away with the locution that the thought *causes* the emotion; rather the thought, in part, *constitutes* the emotions. The thought causes the physiological agitation, and the obtaining of both these events, causally connected, *is* the emotion” (Joyce, 2000). Joyce argues that this is unproblematic to the extent that it is also a problem that a pure cognitivist theory of emotion faces. By saying that my being in danger is what caused me to be afraid, what I actually say on that account is that the danger caused physical agitation and this in combination with my belief that I was in danger is what constitutes the emotion.

It is interesting to examine the background assumption in this accusation, which is that in ‘true’ (rational) emotions we are not moved by mere thoughts, that it is

something actual that moves us. This is also what thought-theory in general holds. However, this is entirely unclear when examining other situations where I would arguably be experiencing real rational emotions. Let's go back to the false story that I was telling my sister in the introduction, about the horrific car accident that left my friend disfigured. If we presume I was actually telling the truth, it is uncontroversial that the emotions my sister would feel would be real. But there is no actual event she is responding to, merely my recounting of the event. If I am watching the news, I respond to the framing of the narrative by the news reader.

But what if green slime would actually be threatening the world. I would see it coming into my street and feel good old genuine fear. What is this fear really about? Green slime is too limited as an answer, because in and of itself it is unclear what it is about green slime that is capable of inducing fear. What I am actually frightened about might be the pain I will feel when I am covered by it, death itself, or the weird tickly effect it turns out to have. But these are at that point in time all still hypotheticals. The green slime is yet a good 200 meters away, so the things I am actually frightened about at least 8 seconds in the future. As pointed out earlier, I have a stimulus, a belief, and then the emotion. And as the above examples show, it is at the point of the cognitive evaluation of the belief that the cognitive basis for the propositional notion of emotion is laid down.

In fact, most of our emotions are about hypothetical situations rather than actual situations, e.g. fear is usually an emotional response to what *could* happen, rather than what is actually happening. When a burglar enters my house, I am not afraid per se of the fact that there is someone in my house who ought not to be there. Instead I am afraid of the hypothetical events that would follow the burglar's actions. For example, I fear that the burglar would take my computer, which has all the drafts of my philosophy thesis on it, or I fear he has a gun and will use it on me, or perhaps I am afraid he will permanently move in with me. The rationality of my fear seems contingent on how likely it is for such a hypothetical event to

happen, but the fear itself is consistently directed at an event that is at that point in time not actually happening.

This does not only apply to fear. When I am happy to have finally graduated, it is because I think I will now get a better job or have time to devote to something else, both future hypothetical events. Of course I can also be happy about having achieved something, which seemingly is an event in the past. Goldie talks about this as narrative thinking, which is in fact what we do. When we talk about hypothetical situations we are narratizing the situation. But that would result in all hypothetical situations being 'fictional' to a certain extent (Goldie, 2012). I will return to this in the next chapter.

Also, we are capable of feeling emotion about memories, including false memories. For example, I get very emotional thinking about my grandfather who has passed away and how we used to play by the swimming pool. However, my grandfather does not exist anymore (after all he passed away), so I am feeling emotional about someone who does not exist. One could also say that what I am actually feeling emotional about is the memory of my grandfather or the thought about it, as thought theory would propound. But what if this memory is not an actual memory? Rather than actually remembering playing by the pool, I have seen a photograph of me and my grandfather doing so. Thus I know it happened and slowly over time I remember it not as a photograph but as an actual memory. Then the memory does not even properly exist. Is it then impossible for me to feel emotional about the scene, just because I falsely believe that I remember it happening?

Examples like these show how the scope of what the paradox is about is actually much broader than just fiction. And there is no clear reason why the emotions in response to fiction are actually different from the emotions in response to actual events, based on the idea that fiction and actual events are fundamentally different. Both provide us with a narrative or thought that gives rise to an

emotion. That thought, for all intents and purposes, is true, and its truth-value is independent of the truth-value of the event that evokes the thought.

One could argue that by making this move, we merely displace the problem from the causal chain of fictional character or event to emotion, to the chain from fictional character to thought. How is it that a fictional entity gives rise to a true thought, in the form of: “Sherlock Holmes has died when he fell off the waterfall.” Doesn’t that thought require belief in the real existence of Sherlock Holmes, analogous to how the coordination condition is formulated above? It seems obvious from the hypothetical situations mentioned above that this is not the case.

As a solution to the paradox, thought theory provides a useful option, but at the same time it makes it all the more unclear, what the actual difference between emotion for fictional characters and for actual people consists in. Or more to the point, what the difference between fictional characters or events and actual people/events actually consists in to the point of where they can cause a thought leading to an emotion. I will deal with this in the following chapter.

6. The difference between reality and fiction

6.1 The truth value

In the previous chapter, the question of what we feel emotion *about* popped up. This is a problematic question, but it is a general problematic question, as when watching a news report, or hearing a (true) story that someone tells us. When I fight with my boyfriend about the laundry, am I angry about the laundry, or about him not doing it? How can I be angry about something that didn't happen? This must mean that my anger has to do with the thought or idea that it should have happened, not with something actual. The same goes for me fearing my job interview in the near future. That job interview is not yet actual, it is yet to happen, so how can I fear it? We unproblematically accept that we do experience these emotions. This means that any philosophical difficulty we have with emotions about thoughts involving fictional characters, is just as much a problem for any emotions about thoughts about the future/the past or anything that is not happening right now. As with the behavioural disanalogies, this shows that the fundamental difference between fiction and reality is not quite as clear-cut as we presumed.

So far in the available literature, there has been little discussion about what the actual difference between reality and fiction consists in. However, it highlights an important assumption that the paradox of fiction relies on, namely that fiction *is* fundamentally different than reality. This assumption makes sense, as the difference is what allows the accepted account of emotion to be unproblematic in dealing with actual situations, leaving it only problematic for the presumably special case of fiction. The difference also seems trivial, because there is of course the difference in the underlying truth value. Fiction is not truthful in the sense that it is not actual in our world. Reality is truthful in the sense that it is actual. There is a difference in existence value. The question is how important this difference is with respect to our capability to feel emotion.

Let's engage in a thought experiment. Assume I am reading a book. It is about the second world war, and it is horrifying. It is non-fiction, or so I think, so according to any account of emotion, there is nothing standing in the way of my emotions be classified as real, when reading it. After all, I believe in the real existence of the characters and events in that book. But what if I were reading that same book, but knew it was fiction? It would mean that in that case my emotions would become paradoxical. Or even more problematic, what if I were reading the book without being informed in advance whether it is fiction or non-fiction. Would I respond differently? According to Walton we would. In the former case of believing it to be fiction, I would engage in a game of make-believe, while I would feel actual emotion if I knew it to be reality. But what about situations about the latter, where I am unsure? How are we to know which one to 'choose'?

Do these in-between cases matter? Generally speaking, isn't it clear when something is fiction and when something is not? I would argue there are in-between cases abound. A novel being based on an event that actually happened, propaganda, a carefully cut news story showing the view of the reporter, thus always eliminating facts that show the "whole picture". The example I started with: me telling my sister a story about a car accident, with my sister only realizing it was not true after I told her so. In fact, we fictionalize reality every day when relating a story, leaving out bits and pieces, magnifying or exaggerating other parts. When thinking about it, it is rare outside of direct experience that a story, a news item, a report is an exact representation of reality. The rise of "fake news" has brought that to our attention.

Perhaps then what is crucial is the intention of the author, whether he intends it to be fiction? A news item is intended to represent the truth and we know this. The problem illustrated by the above thought experiment remains however, because we do not always know beforehand what the intention is. In Western Europe we consider Russian news items on the greatness of Putin as fiction, while in Russia

they would seem to consider them as actual news, while regarding our reporting of their perception as Western propaganda. That would mean a Russian would be “allowed” to feel actual emotion towards the same item a Westerner would not.

The only way to save Walton’s account is to say that the format of the book or movie in itself triggers us to engage in a game of make-believe, pre-emptively, perhaps only being convinced to actually engage when it is verified that what we are reading is reality, not fiction. But this means we would approach many stories this way. Also it has the problem of either retrospectively changing the nature of emotion from ‘quasi’ to ‘actual’, when we come to realize what we are reading is true -which seems complete nonsense - or alternatively having experienced merely quasi-emotion when we could have experienced real emotion (and we would have, had we known what we were reading to be non-fiction). And wouldn’t this problem then apply to all stories we read, including newspaper articles? It would be an uncomfortable conclusion if we only experience quasi-emotion in response to newspaper articles. One could argue that it is the context of a newspaper that triggers us into not playing a game of make-believe, but one can imagine the issue of fake news arising. Does our emotion retrospectively become quasi emotion? Or can we only feel real emotion when experiencing something in real life?

6.2 Narrative theory

Peter Goldie says that “our lives have a narrative structure - roughly speaking, they comprise an unfolding, structured sequence of actions, events, thoughts, and feelings, related from the individual’s point of view.” (Goldie, 2000, p. 4). This idea starts with Amelie Rorty, who first noted the narrative structure of emotions and the cruciality of them for the development of critical thinking. (Rorty, 1988, pp. 295-297). Narrative theory starts from the assumption that in order to understand our everyday experiences, we use narrative as a basic strategy. Goldie defines a narrative as a representation of a sequence of events. This is done from a

certain perspective or perspectives and it gives coherence, meaningfulness and evaluative and emotional import to the events related. In chapter 4 I discussed the role of rationality, coherence and the appropriateness of emotions. Narrative theory makes it clear what the underlying structure is that allows these notions to play a role in our evaluation of emotions.

In his book “The mess inside” Goldie consistently demonstrates how narratives help us in everyday life to make sense of our surroundings. (Goldie, 2012). He points out there is an epistemic gap when we, as an audience, are watching the detective attempt to find the killer, while we have known all along who has done it. This epistemic gap is a sort of emotional equivalent to the behavioural disanalogies that fiction seems to provoke. But this gap makes sense when thinking about the perspective of the audience. The excitement and fear we feel when the protagonist unknowingly goes to a cabin in the woods with the killer are in that case emotions felt for the fictional character. This is similar to feeling fear and excitement when watching a news report about a volcano about to erupt, while people in the village at the foot of the mountain are unaware and go about their lives as normal. Obviously *I* am not fearing the volcano (I have yet to spot one in the Netherlands), rather I fear for the people living in the village. Goldie calls this divergence of perspectives ‘dramatic irony’. This idea of ‘dramatic irony’ is also applicable when thinking about the past, where you now know things about your past that you then, in the past, did not know. (Goldie, 2012, p. 28).

This dramatic irony is what I dealt with in chapter 3, when discussing the behavioural disanalogy that Walton found problematic. We are just as much an audience when watching a news report as we are when watching a play, because we perceive our causal influence on the narrative that is happening as non-existent, precisely because we are an audience. This does not mean the emotion we feel is not real, it merely means that the set of actions available to us is limited, because our causal influence is limited.

It goes outside the scope of this thesis to delve much further into the discussion between hard-core narrativists and those with a more in-between position, such as Goldie. However, what is interesting, especially in light of chapter 4, is that narrative theory in general points out that narrative accounts of events have three distinguishing points from mere causal accounts, namely coherence, meaningfulness and evaluative and emotional import. Coherence in this respect means that “it reveals, through the process of emplotment, connections between the related events, and it does so in a way a mere list, an annal, or chronicle, does not” (Goldie, 2012, p. 14). Meaningfulness is a connected notion to coherence and means on one hand that the *internal* perspective of the characters makes sense (i.e. is coherent) and on the other hand the *external* perspective of the narrator, which are the thoughts or feelings on why the narrative was related. This borders on Feagin’s metaresponse that I dealt with earlier. What is interesting is that Goldie, in his account of narrative in everyday life is recounting a condition that in the particular case of fiction we would use to judge the “eloquence” of the work, if we may borrow from Hume. The evaluative and emotional import for Goldie is the emotional import to people internal to the narrative (Goldie, 2012, p. 16). He makes it clear that this coherence is what allows us in combination with meaningfulness and emotional import to experience emotions.

On this account, there is not much difference between the dramatic structure of actual events or that of fictional events. The same notions – for Goldie coherence, meaningfulness and emotional import, or on my account coherence, consistency and appropriateness – play a part in experiencing emotions, whether fictional or actual.

Goldie rightfully argues that not everything we do is planned in an author-like way, which does apply to fiction. Furthermore, to identify life with the life that is being narrated means losing the distinction between representing and what is represented. This is similar to what I argue here, which is not that reality is a

narrative, but rather that the way we deal with reality, in the sense of experiencing emotions about it, is often as though it is a narrative. This means that there is no fundamental difference to how we deal with fictional narrative or actual narrative.

6.3 The implications of make-believe

To return to Walton, a big problem for the arguments for his account is that, when carefully considered, they actually have a much wider implication than merely for emotions experienced in response to fiction. His first argument is based on the premise that we can only fear what we believe can endanger us. An interesting case in this respect is of phobias. In the ‘Fear of flying argument’, it is argued that we can fear flying, despite knowing it to be the safest means of transportation. I dealt with this argument as an objection to cognitivism in chapter 4, but for now it is interesting to notice that on Walton’s account, what we experience is not real fear, but quasi-fear, because we do not believe flying can actually endanger us, because we know it to be the safest means of transport. This has the interesting consequence that anyone with a phobia should be able to quit playing the game of make-believe and thus stop being afraid at will. It seems unlikely that sufferers of phobias (and the army of psychiatrists treating them) would agree to that.

But also in less extreme situations, it seems possible to feel fear, while I know myself to be safe. Take Walton’s own example of Darwin jumping back, that he treats as an example of make-believe. If that too is make-believe, then a great good deal of our emotions are make-believe. Everything I experience in a theme park is make-believe. But also the fear I experience when watching a bomb attack in Syria on the news. It is hard to argue that I believe that bomb attack can endanger me personally. Or on a sliding scale: a suicide bombing in Paris. One could argue that my fear is not directed at that particular event, but rather at the resulting raised possibility of such a hypothetical bombing happening in Amsterdam, where I live, and that such a hypothetical situation is in fact close to a fictional event, as it is not actual in the sense that it is directly threatening to me.

In chapter 3 I showed that the behavioural disanalogies from Walton's second argument are also present when watching the news, which on his theory would make these emotions also quasi-emotions. If the majority of emotions would thus be labelled as quasi-emotions, even emotions Walton himself would presumably label as 'real' rather than quasi, the theory misses its mark.

6.4 Truth in fiction

When talking about the underlying truth value of fiction versus reality, it is important to note that I speak here of truthful in a specific sense. There are different scopes here under which the truth operator functions. For example, I can say it is true within the fictional scope that Sherlock Holmes solved crimes, just as I can state that in the fictional scope it is false that he was a small fat man. On a wider scope these statements both pertain to fiction and as such are by definition untrue (of reality). There was never such a man as Sherlock Holmes, so I cannot make true or false statements about him. What is salient to take from this discussion is that the problem of referring starts from our judgement that it is appropriate in everyday language to make such statements as those about Sherlock Holmes.

There are a couple of peculiarities about fiction. For one, as Lamarque states, we take certain things to be true about a fictional world, often making inferences beyond what is explicit in the fictional narrative. For example, Arthur Conan Doyle doesn't explicitly tell us that Holmes has a kidney or a working blood stream, but we assume that he does have these, given that he is a normal human being, rather than a robot or Martian (Lamarque, 1981).

Walton's coins the terms Reality Principle and Mutual Belief Principle to explain why we can make such assumptions in fiction. The Reality Principle starts from the idea that the fictional worlds we imagine are as close to the real one as

reasonable, given the fictional nature of the story. This explains why we can coherently make statements about Holmes' possession of a blood stream. The Mutual Belief Principle takes into account that the cultural context of the creator, but can be used to account for our ability to accept fictional worlds full of sorcerers and swords (K.L. Walton, 1990).

This judgment of appropriateness about the possibility of making these statements is at the basis of the judgment on the truth-value of such statements about fictional entities. This is because we can also make inferences outside of the scope of the fictional canon, and have a coherent opinion about these statement and whether or not they are likely to be truthful. E.g. "Harry Potter is an admirer of Sherlock Holmes" could be truthful, whereas "Arthur Conan Doyle's Sherlock Holmes is an admirer of Harry Potter" seems plainly false. Also these judgements of appropriateness on making statements about fictional characters are similar to the judgement of appropriateness we make about feeling emotions towards fictional characters. It is appropriate to mourn Dumbledore, but not to be happy when he dies. The similarity occurs because both judgements rely upon the coherence of the underlying narrative. This is also true for narratives that *are* grounded in reality. Here it is useful to again think about Goldie's conditions for feeling emotions.

It is hard to feel emotion when a (fictional) character does something completely 'out of character'. In the popular script writing book "Save the cat", Blake Snyder strikingly calls this Double Mumbo Jumbo (Snyder, 2005). It means that "audiences will only accept one piece of magic per movie. It's The Law. You cannot see aliens from outer space land in a UFO and then be bitten by a vampire and now be both aliens and undead". What I mean by it in this context is that the narrator cannot just pull a *deus ex machina* and expect the audience to remain emotionally engaged with the characters and story he is telling. Double Mumbo Jumbo – Anna Karenina did not die when she threw herself in front of the train, because suddenly Tolstoy tells us she was an undead zombie after all – or a *deus*

ex machina – just when the train threatened to hit her Superman appears and stops the train just in time – disrupt the internal logic of the narrative that the reader was engaging with, making the story incoherent and inconsistent. Coherence and consistency are necessary conditions for experiencing emotion. Unlike what the coordination condition states, it is not a necessary condition of emotion for us to believe in the existence of the intentional object of our emotion, rather it is a necessary condition for us to have a thought with the causal power to allow an emotion.

In chapter 2, when discussing the illusion theory, I pointed out that when we talk about a novel being believable we actually mean something like “the book formed such a coherent whole, that I could believe in its possibility”. Similarly, when talking about whether an emotion is appropriate in response to fiction, we take into account the coherence and thus the believability of the fiction. This does not mean it is believable when nearer to our reality, rather this coherence is related to the aesthetic judgement of what makes a book or a movie ‘good’. Fear when watching the actual film *The Green Slime* is likely judged inappropriate as the film is considered ‘bad’.

We tend to sooner judge reality as coherent than we do fiction, for the simple reason that it is reality and thus confirms to the basic rules that we are used to (such as: wizards do not exist, killing is morally deviant, time functions linearly, etc). This is also visible in the role morality plays in our ability to experience emotion in response to an event. This becomes clear in Tamar Gendler’s work on imagination and the so-called “Puzzle of Imaginative Resistance” (Gendler, 2000). It is “the puzzle of explaining our comparative difficulty in imagining fictional worlds that we take to be morally deviant.” Gendler uses an example by Walton to show the type of morally deviant world she is talking about: “In killing her baby, Giselda did the right thing; after all it was a girl”. Her point that we have trouble imaging this, does not mean it is impossible for us to feel the

emotions appropriate to that fictional situation (joy for Giselda having done the right thing), just that it is harder.

This puzzle, just as the paradox, does not only apply to fictional worlds. We have difficulty empathizing with soldiers that killed Jews in concentration camps in Nazi-Germany or FARC rebels that kidnap and kill in Columbia, because we take their actions to be morally deviant. Interestingly, we become capable of feeling emotion for them when their story because properly contextualized in a more or less fictional narrative that makes their actions coherent and consistent for us based on the context in which they happened. It is just that in fictional worlds, morally deviant worlds tend to be more common than morally deviant acts – let alone worlds – are in reality. Or we are more commonly presented with morally compatible narratives as the type of news stories that we are presented with are preselected by some form of organization, whether being the news organization or the journalist? It is not unthinkable that morally deviant stories are filtered out due to a form of “moral bias”, that selects only stories that are morally compatible as relevant or newsworthy, precisely because as an audience, it is hard to feel emotion for those stories, thus making us less interested.

The difference between reality and fiction with respect to the underlying truth value is only important for the solution of the paradox if we accept the coordination condition in its original form. If it is denied, as thought theory does, it becomes irrelevant, because there is no longer a premise that refers to the truth value. The intentional object becomes the thought, and here we are not questioning the ontological status of a thought per se.

However, as we have seen above, this means that there is no longer a clear difference between reality and fiction and means that the problems that the paradox of fiction highlights are general problems for any theory of emotion. When fiction is no longer a special case of feeling emotions, it becomes clear what the role of coherence and consistency is in feeling appropriate emotions.

7. Conclusion

Reading and watching fiction is one of the most enjoyable activities known to mankind. We pity Anna Karenina, we cry for Dumbledore and we fear green slime. The paradox of fiction poses such a paradox because phenomenally it seems obvious that we do experience emotions. Also, isn't the point of engaging with fiction to experience emotion in the first place?

In the preceding chapters I have argued that by rejecting the *coordination condition* through thought theory, it effectively becomes unimportant that there is a difference in underlying truth value between actual events or characters and fictional events or characters. Furthermore, this difference in truth value is in many cases unclear, as in the case of "framed" news reports, when reading a book without knowing whether it is truth or fiction or when hearing a story that someone tells us.

The behavioural disanalogies that are used as an argument against illusion theories of fiction and that Walton uses to advance his theory of make-believe are in fact not specific to fiction, but are also present when watching the evening news, because the behavioural disanalogy does not depend on our belief in the existence of the character or the event, but on our perceived causal influence, which in many cases is limited. We have an *external* perspective as an audience as a result of this. Just as something in South-East Asia is outside of my perceived scope of influence, so is a pre-written down narrative in a novel or a scripted play.

The emotional disanalogy that the paradox of tragedy highlights is also not specific to fiction. We enjoy documentaries about serious subjects, we voluntarily watch news reports and we curiously slow down for traffic accidents, just as we install ourselves on the couch to spend an evening reading a sad novel.

Radford points to the irrationality of our responses to fiction. Closely connected are notions of coherence and consistency and appropriateness. He claims to accept all three premises of the paradox, while in fact changing the coordination condition into stating that it is a necessary condition for being rationally moved to have a true belief in the existence of the intentional object. The coordination condition thus becomes a normative condition instead of an existential one. Rationality is based on appropriateness, which in turn is based on consistency and coherence, or coherence, meaningfulness and emotional import if we follow Goldie. The judgement of rationality is just as problematic for everyday cases in which we would unproblematically name our emotions as real (and thus presuppose them to be rational).

Thought theory, rejecting the coordination condition and instead stating that we are moved by thoughts, allows for a clearer theoretical role of the above notions as narrative theory also shows. It also allows us to see that in fact, many, if not all, of our emotional responses in every-day life are in fact not about actual situations but about hypotheticals, meaning that they are in fact fictional to the extent that they are not actual. Narrative theory describes these hypothetical situations through the use of narrative again highlighting the importance of the coherence of the underlying narrative in our capability of experiencing true emotions. Ultimately, the difference between reality and fiction, when reduced to the underlying truth value, becomes irrelevant when we deny the coordination condition with the help of thought theory.

In conclusion, the paradox of fiction is usually taken to be a paradox specific to fiction. In this thesis I have argued that the problem it points out is in fact a problem for any kind of emotion we experience, regardless of whether it is in response to fiction or in response to reality, and that this problem results from the specific cognitivist account of emotion that is generally accepted. If we avoid this problem by accepting a form of thought theory and thus denying the coordination condition, we accept that the scope of the paradox is much wider than just that of

fiction. There is no fundamental difference between fiction and reality that is relevant when using thought theory as an altered version of the cognitivist theory of emotion.

When we no longer consider emotions in response to fiction as a special case of emotion in general, but rather we view it as a paradigmatic case of our experiencing of emotions, it becomes clear what role the coherence of narrative structures plays in emotion and that we should adapt thought theory to accommodate this central role of coherence and consistency in order to have a functional theory of emotion that also explains the apparent phenomenology of emotion itself.

Ultimately, our emotional responses to actual events are dependent on the same features as emotions in response to fiction. So rather than it being impossible or paradoxical for us to experience emotions about fictional characters, we perceive the stimuli for most of our everyday emotions as narratives, and our capability of feeling emotion at all is contingent on that. We enjoy the movies and we love books, just as we love life.

8. References

- Bouwsma, O. K. (1965). *Philosophical Essays*: University of Nebraska Press.
- Carroll, N. (1990). *The philosophy of horror*. London: Routledge, Chapman and Hall, Inc.
- Coleridge, S. T. (1885). *Biographia literaria, or, Biographical sketches of my literary life and opinions : and two lay sermons*. London :: Bell.
- De Sousa, R. (2014). Emotion. *The Stanford Encyclopedia of Philosophy*. Spring 2014 Edition. Retrieved from <http://plato.stanford.edu/archives/spr2014/entries/emotion/>
- Deigh, J. (1994). Cognitivism in the Theory of Emotions. *Ethics*, 104(4), 824-854.
- Eldridge, R. (2003). *An Introduction to the Philosophy of Art*: Cambridge University Press.
- Feagin, S. L. (1983). The Pleasures of Tragedy. *American Philosophical Quarterly*, 20(1), 95-104.
- Gendler, T. S. (2000). The puzzle of imaginative resistance. *Journal of Philosophy*, 97(2), 55-81. doi:10.2307/2678446
- Goldie, P. (2000). *The Emotions: A Philosophical Exploration*: Clarendon Press.
- Goldie, P. (2009). *The Oxford Handbook of Philosophy of Emotion*: OUP Oxford.
- Goldie, P. (2012). *The Mess Inside: Narrative, Emotion, and the Mind*: OUP Oxford.
- Hartz, G. A. (1999). How We Can Be Moved by Anna Karenina, Green Slime, and a Red Pony. *Philosophy*, 74(290), 557-578.
- Hume, D. (1757). *Four Dissertations: I. The Natural History of Religion. II. Of the Passions. III. Of Tragedy. IV. Of the Standard of Taste*: A. Millar.
- James, W., Haupt, I. A., & Lange, C. G. (1922). *The Emotions*. Baltimore,: Williams & Wilkins Company.
- Joyce, R. (2000). Rational fear of monsters. *British Journal of Aesthetics*, 40(2), 209-224. doi:10.1093/bjaesthetics/40.2.209
- Lamarque, P. (1981). How Can We Fear and Pity Fictions? *The British journal of aesthetics*, ISSN 0007-0904: vol. 21 (pp. 291. \END).
- Lamarque, P. (1991). Review: Mimesis as Make-Believe: On the Foundations of the Representational Arts by Kendall L. Walton. [Mimesis as Make-Believe: On the Foundations of the Representational Arts, Kendall L. Walton]. *The Journal of Aesthetics and Art Criticism*, 49(2), 161-166. doi:10.2307/431705
- Lamarque, P. (2003). Fiction. In J. Levinson (Ed.), *The Oxford Handbook of Aesthetics* (pp. 377-391). New York: Oxford University Press Inc.
- Lamarque, P. (2005). 21. Fiction. In J. Levinson (Ed.), *The Oxford Handbook of Aesthetics*: OUP Oxford.
- Levinson, J. (1996). *The Pleasures of Aesthetics: Philosophical Essays*: Cornell University Press.
- Neill, A. (1991). Fear, Fiction and Make-Believe. *The Journal of Aesthetics and Art Criticism*, 49(1), 47-56. doi:10.2307/431648
- Neill, A., & Ridley, A. (2013). *Arguing About Art: Contemporary Philosophical Debates*: Taylor & Francis.

- Radford, C. (1977). Tears and Fiction. *Philosophy*, 52(200), 208-213.
- Radford, C., & Weston, M. (1975). How Can We Be Moved by the Fate of Anna Karenina? *Proceedings of the Aristotelian Society, Supplementary Volumes*, 49, 67-93.
- Robinson, J. (2004). The Emotions in Art. In P. Kivy (Ed.), *The Blackwell Guide to Aesthetics*: Wiley.
- Rorty, A. (1988). *Mind in Action: Essays in the Philosophy of Mind*: Beacon Press.
- Routley, R., & Sylvan, R. (1980). *Exploring Meinong's jungle and beyond: an investigation of noneism and the theory of items*: Australian National University.
- Saatela, S. (1994). Fiction, Make-Believe and Quasi Emotions. *British Journal of Aesthetics*, 34(1), 25-34. doi:DOI 10.1093/bjaesthetics/34.1.25
- Snyder, B. (2005). *Save the Cat!: The Last Book on Screenwriting You'll Ever Need*: Michael Wiese Productions.
- Stocker, M., & Hegeman, E. (1996). *Valuing emotions*. Cambridge England ; New York: Cambridge University Press.
- Walton, K. L. (1978). Fearing Fictions. *Journal of Philosophy*, 75(1), 5-27. doi:Doi 10.2307/2025831
- Walton, K. L. (1990). *Mimesis as Make-believe: On the Foundations of the Representational Arts*: Harvard University Press.
- Walton, K. L. (2014). *In Other Shoes: Music, Metaphor, Empathy, Existence*: Oxford University Press.