Fiction and Emotion

Stacie Friend

For *The Routledge Handbook of Philosophy of Imagination*, edited by Amy Kind
DRAFT – PLEASE DO NOT CITE OR CIRCULATE WITHOUT PERMISSION

Engagement with fiction often inspires emotional responses. We may pity Sethe while feeling ambivalent about her actions (in *Beloved*), fear for Ellen Ripley as she battles monstrous creatures (in *Alien*), get angry at Okonkwo for killing Ikemefuna (in *Things Fall Apart*), and hope that Kiyoaki and Satoko find love (in *Spring Snow*). Familiar as they are, these reactions are puzzling. Why do I respond emotionally if I do not believe that these individuals exist or that the events occurred? If I merely imagine that my best friend has betrayed me, I do not become angry with him; if I did, I would be considered irrational. Yet although beliefs seem to be necessary for emotions in other contexts, we respond to fiction without the relevant beliefs. These observations prompt two questions about our emotional responses to fiction (henceforth: *fictional emotions*). The first is *descriptive:* Should fictional emotions be classified as the same kind of emotions we experience in other contexts? The second is *normative:* Are fictional emotions irrational or otherwise inappropriate?

Both questions are typically framed in terms of the *Paradox of Fiction,* constituted by the following three statements:

1. We experience (genuine, ordinary) emotions toward fictional characters, situations, and events.
2. We do not experience (genuine, ordinary) emotions when we do not believe in the existence of the objects of emotion.
3. We do not believe in the existence of fictional characters, situations, and events.

Insofar as the three claims are individually intuitive but jointly incompatible, they cannot all be true. The debate over the descriptive question has focused on which claim to reject to resolve the paradox. The debate over the normative question has focused on whether all three claims must be accepted (though I question this approach below). I take each of these debates in turn.

The Descriptive Question

The standard approach to the descriptive question assumes that one claim of the Paradox of Fiction must be rejected. The majority of philosophers reject either (1) or (2). Although the rejection of (3) is a possibility—and some philosophers have argued that we have the relevant beliefs, perhaps subconsciously or partially (Hartz 1999; Suits 2006)—I will focus on (1) and (2) because the disagreement between advocates of each position centers on the role of imagination in emotion.

It was Kendall Walton's (1978) rejection of (1) that prompted interest in the descriptive question (see also Walton 1990, Ch. 7). Walton famously argued that moviegoer Charles, shrieking in terror of the onscreen blob coming straight at the camera, was not genuinely afraid. Instead, Charles's experience is better described as *make-believe* or *imagined* or *quasi* fear, and similarly for our pity of Sethe, anger at Okonkwo, and so forth. On the face of it, this conclusion is counterintuitive. Walton is not, however, making the patently implausible claim that in fictional contexts we fake or pretend to have emotions, or that our experiences must be less intense or more easily controlled than in other contexts. To the contrary, Walton agrees that we experience emotions that share features with pity, fear, and so on; what he denies is that these are genuine pity and fear (i.e., emotions of just the same kind as ordinary pity and fear). Walton and his supporters point out that the responses occur in the context of imagining rather than believing the content of the fiction and, as a result, they possess features that distinguish them from other instances of emotion.

The vast majority of philosophers and psychologists nonetheless disagree with Walton's conclusion. Some argue that we need believe only that certain claims are fictionally true, rather than that the objects are real, to have genuine emotions (Neill 1991; Livingston and Mele 1997). But by far the most popular position is that no kind of belief is necessary for emotion, and in particular that imagination without belief is sufficient. Proponents of this position conclude that there is no good reason to deny that our responses to fiction are ordinary, genuine emotions. They therefore reject (2).

Rejecting (2)

The debate over (2) is concerned with a particular form of *cognitivism* about the emotions. According to cognitivist theories, some kind of cognitive evaluation—often, but not necessarily, construed as a propositional attitude—constitutes an essential component of an emotion. In paradigm cases the attitude is a belief: in pity the belief that someone has suffered, in fear the belief that one is in danger, and so on. Many emotional states also have an affective component, involving phenomenological and physiological responses as well as associations with action. For example, fear usually involves a feeling of anxiety and an increase in heart rate and perspiration, along with a desire to flee. According to cognitivists, these aspects of an affective state or episode are insufficient by themselves to constitute an emotion. First, many emotions are associated with a similar physiological and phenomenological profile; the distinction between, say, fear and anger may simply be whether our feelings arise in evaluating a situation as dangerous or offensive. Second, these features do not distinguish emotions from moods—affective states lacking in content, as when I am anxious but not about anything in particular—or lower-level automatic responses, such as being startled by a loud noise. Emotions are open to rational assessment in ways that moods or reflex responses are not, and cognitivists explain this by pointing to the cognitive component.

Traditionally, the explanation assumed that the cognitive component was a belief (e.g., Lyons 1980; Oakley 1992), or even that emotions could be identified with judgments (e.g.,Solomon 2003; Nussbaum 2001), so that the rationality of emotions was determined by the rationality of the relevant beliefs or judgments. Anyone who accepts this traditional, *narrow* form of cognitivism accepts (2). Significantly, narrow cognitivism does not require that the relevant belief be *true* for emotion to be possible. I will experience genuine fear if I believe a burglar is in my house, even if it turns out that I've only heard the wind rattle the door.

Those who reject (2) fall into two camps. In the first are those who agree with *broad cognitivism*, according to which emotions involve some form of cognitive evaluation, but not necessarily a belief (e.g., Stocker 1987; Greenspan 1988). From this perspective, pity involves the thought that someone suffers, fear the thought that one is in danger, and so on, but these thoughts can be entertained, supposed, imagined, or whatever. One motivation for this view is the existence of phobias, which appear to be cases of genuine fear in the absence of the belief that one is in danger. Another is the fact that the rationality of emotions departs in various ways from the rationality of beliefs (Greenspan 1980, 1988). Broad cognitivists who address the Paradox of Fiction, often called "thought theorists," argue that imagining in response to fiction is sufficient for genuine emotions (e.g., Lamarque 1981; Carroll 1990; Matravers 1998). In the second camp are *non-cognitivists,* who deny that cognition of any kind is necessary for emotion. They argue that emotions can be triggered by stimuli that directly affect the amygdala, prompting an affective response prior to any signal's being sent to cortical areas of the brain (LeDoux 1996). In this way emotions involve an "affective" or "embodied" appraisal of a situation, which can arise without any intervening cognition (Robinson 2005 and Prinz 2004, respectively). For example, seeing a coiled shape in the dark or hearing a growl might prompt fear even before I have had a chance to process the stimulus. Given that no cognition is necessary for an emotion, non-cognitivists such as Robinson and Prinz also reject (2).

Importantly, the disagreement between cognitivists and non-cognitivists concerns only the *trigger* of emotion: whereas non-cognitivists claim that emotions can be initiated without cognition, cognitivists only count a state or episode as an emotion once cognition is involved. But non-cognitivists do not deny a role for cognition in emotion. First, they acknowledge that cognition plays a part in the development of emotions even when they are initiated more directly (Robinson 2005: 75–79). For instance, if I realize that the coiled shape is just a rope, my fear dissipates. Second, non-cognitivism does not exclude the possibility of emotions triggered by cognitions; thinking of a sad incident can make me sad. Therefore non-cognitivists agree with broad cognitivists that genuine emotions can occur when we merely imagine, but do not believe, the content of a fiction. Emotions can be triggered by mere imaginings, as with our pity of Sethe, our hope for Kiyoaki and Satoko, and so on. And imaginings can play a role in the development of emotions triggered non-cognitively. Even if audiences of *Alien* experience fear as a gut reaction when the cat suddenly jumps out of its hiding place, this fear turns to relief once we recognize that the cat is not the alien. For this reason, the debate over (2) is not a debate between cognitivists and non-cognitivists, but between narrow cognitivism and other theories of emotion. According to narrow cognitivists, genuine emotions are not possible when we do not believe that the objects of emotion exist. According to other theorists, they are.

Now, it is fair to say that narrow cognitivism is a minority position. Nearly everyone—philosophers, psychologists, neuroscientists, ordinary folk—agree that we can experience genuine emotions in the absence of relevant beliefs, often precisely because this seems to be the case in fiction. Contrary to the assumption of most parties to the debate, however, rejecting (2) does not settle the issue of how to classify fictional emotions. In particular, rejecting (2) does not entail accepting (1), for there may also be *other* reasons to deny that fictional emotions are genuine emotions (Stecker 2011). I consider these in the next section.

Rejecting (1)

Among the reasons to reject (1) are the obvious dissimilarities between fictional emotions and emotions in other contexts. In ordinary circumstances there is a close relationship between our beliefs and central features of emotion that are missing in the fiction case. One is the connection between emotion and motivational force. Fear provokes a desire to flee, pity a desire to help, and so on. Ordinarily, if we believe that nothing endangers us, we do not desire to flee; if we believe that no one has suffered, we do not desire to help. Another is the way emotions are evaluated in light of our beliefs. If I do not believe that my friend has done anything wrong yet I am still angry at him, this response will be unjustified and perhaps even irrational. A third feature is the intentionality, or aboutness, of emotions. If you tell me about the terrible experiences of a friend I may pity her. If I then find out there is no such person, I can no longer experience pity *of your friend*, just as a child who knows there is no Santa Claus can no longer hope to meet *him*. (Notice that the problem is not merely that the object fails to exist. We may want to say that the Greeks worshipped Zeus even though there is no Zeus, given their false belief. Instead, the problem is the lack of any belief.) None of these features characterizes our emotions in response to fiction. Our pity of Sethe has no motivational force, we do not count ourselves unjustified in pitying someone we believe is unreal, and our feelings still seem directed toward Sethe despite her nonexistence. Fictional emotions thus depart in significant ways from emotions in ordinary circumstances.

Defenders of (1) argue that these disanalogies do not justify treating fictional emotions as a distinct kind. They draw attention to other instances of genuine emotions where beliefs are lacking. If phobic Jack is afraid of the gentle Fido even though he does not believe that Fido poses any danger to him, no one would deny that his fear is genuine. And Jill might fear walking on a glass platform over a deep drop, without believing that there is a genuine risk of falling (Gendler and Kovakovic 2006, 250). If these are genuine emotions, it looks as if belief is not required for fictional emotions to be so classified. However, fictional emotions cannot be assimilated to phobias and automatic reflex responses. Jack and Jill are likely to recognize that they have no good reason for their fears. By contrast, readers of Morrison’s novel take themselves to have numerous reasons for pity. Moreover, if we take Jack and Jill as experiencing genuine emotions, this is only insofar as those emotions motivate action. If Jack did not avoid Fido or Jill was happy to step on the glass floor, we might doubt that they were afraid. Such motivations are lacking in the fiction case (Walton 1990: 201–202).

Proponents of (1) might then point to our emotional responses to some kinds of nonfiction representations: in these cases there is belief, but due to spatial or temporal distance no motivation to act. If something occurred a hundred years ago or a thousand miles away, we are as little capable of doing anything about it as we are in the case of fiction (Matravers 1998: 69–73; Gaut 2003). Moreover, they argue, our believing in the one case that the objects of emotion are (or were) real and in the other case that they are (or were) unreal has no impact on the experience. On the other hand, they must concede that by contrast with emotions directed at real individuals past or present, we cannot literally experience emotions *toward Sethe* if we believe there never was such a person (Lamarque 1981). The question is whether this is sufficient to show that the pity is of a distinct kind.

Emotions and Make-Believe

It seems as if we have reached an impasse. After all, there is little disagreement about the features of fictional emotions themselves, which are widely acknowledged to be both similar to and different from emotions in other contexts. Now, an empirical approach might identify other, less obvious features of fictional emotions than those so far discussed (cf. Weinberg and Meskin 2006). For example, one psychological study suggests that emotional reactions to fiction are less intense than emotional responses to nonfiction (Cova et al., ms). But it is unlikely that the matter can be settled empirically. First, most empirical studies of emotion simply presuppose that emotions are all of a kind, insofar as psychologists and neuroscientists draw their conclusions from responses to fictional stimuli such as pictures and stories. Second, although empirical results may be illuminating in various ways, further data by themselves would not settle the issue; the data have to be interpreted.

If closer consideration of the emotions themselves will not shed more light on their classification, we must look instead at their context: the *fictional* aspect of fictional emotions. For the most persuasive arguments against (1) rely, not on a general background assumption of narrow cognitivism, but rather on the idea that emotions in response to fiction arise in the context of imaginative engagement. Indeed Walton's own account of fictional emotions makes sense only in the context of his theory of make-believe.

Starting from examples of novels, plays, films, paintings, and sculptures, Walton (1990) aims to construct a theory of what he calls *representational art* or *fiction*. Walton defines a work of fiction as a work whose function it is to act as props in certain games of make-believe. The connection with children's games of make-believe, where the props may be dolls and toy trucks, is intentional; for Walton there is continuity between the two types of games. The idea is that works of fiction are designed to prescribe imaginings about their content, and imagining what is prescribed is participating in the game of make-believe authorized by the work. What a work prescribes imagining is what is "true in the fiction," or simply *fictional*. On Walton's theory our imaginings about the content of a work—for instance, my imagining that Sethe suffers because her children are born into slavery—are generated by more basic games of make-believe that involve imagining about ourselves. Specifically, we make believe that in engaging with the fiction we are learning about actual fact, and we respond to the fictional events from within the scope of this imagining. In reality I am reading a novel; in imagination, in my game of make-believe, I am reading a true report about actual people and events. And it is within this context that I have emotional experiences.

Suppose that I am caught up in *Beloved* and I find myself pitying Sethe. I am perfectly aware that there is no such person. But my imagining is sufficiently vivid that I have a sinking sensation, tears well up, I find it hard to turn the page. According to Walton, facts about my real feelings—specifically, the phenomenological and physiological state I am in due to my imagining, which Walton calls *quasi-pity*—make it fictionally the case that I pity Sethe. From a perspective outside the game of make-believe, the experience does not count as genuine pity. Just as it is the fact that I (really) read a novel that makes it fictionally the case that I am learning about actual fact, it is the fact that I (really) experience quasi-pity that makes it fictionally the case that I experience pity. And because it is *only* fictionally the case that I am reading a true report and learning that Sethe suffers because she is separated from her children—the imaginings that explain my quasi-pity—it is *only* fictionally the case that I pity her. In short, Walton's introduction of quasi-emotions is a natural consequence of his theory of make-believe.

Now, most of Walton's opponents agree that when I read *Beloved* I imagine learning about actual fact. They also agree that it is within this imaginative context that my emotional experiences occur. However they disagree with Walton's conclusion that the emotional experience is as much a part of the game as the events of the novel; for his opponents, the emotion falls outside the scope of the make-believe. To see what difference this makes, it is helpful to consider the distinction between imagining and believing.

Quarantined Emotions

The difference between imagining and believing is usually construed as a functional difference rather than a difference in content. So, for example, we can both believe and imagine that the sun is out; the contrast is not in the proposition believed or imagined, but rather in the function this content plays with respect to other cognitions, affective states, behavior, and so on. Two features of imagining are often noted, *mirroring* and *quarantining* (see Gendler 2003). Imagining *mirrors* belief to the extent that it is governed by similar inferential constraints. For example, if we imagine that the sun is out, we typically imagine that it is daytime. At the same time imagining is *quarantined* insofar as we treat its implications and effects as limited to a certain domain. For example, we do not take imagining that the sun is out to mean that we face an increased risk of skin cancer. The same applies to imagining in response to fiction. We draw conclusions about how Sethe feels based on what we know about people in general, but we do not expect to be able to her descendants or locate her grave. In short, at least some of what we imagine about Sethe remains compartmentalized from our beliefs about the world, so that it does not guide our ordinary actions. If Walton is correct that fictional emotions occur within the imaginative context, we should expect them similarly to be quarantined.

One way to test for the presence of quarantine is to see how we judge potential conflicts. In a pretend tea party, children playing with empty cups typically believe that the cups are empty while imagining that they are full. Because the imagining is quarantined from the belief, there is no conflict. Imagining in response to fiction works the same way. In reading *The Hitchhiker's Guide to the Galaxy* we imagine that the earth is an experiment designed by mice, while all the while believing that it is not. Quarantine prevents conflict; it also explains why scientists engage in astronomy and geology to explain the origins of the planet rather than trying to communicate with rodents.

There is some reason to think that the same applies to fictional emotions, as can be seen when we consider fictions about real individuals. Perhaps I believe that Richard III was a good king who did not murder his nephews in the Tower of London, but was instead the victim of Tudor propaganda. I admire him for his good deeds and pity him for the way he was killed. If I see a compelling production of Shakespeare's *Richard the Third*, I may well respond with very different emotions, condemning rather than admiring him, viewing his death as deserved rather than pitiable. Although some will disagree, I see no reason to deny that the Richard portrayed in Shakespeare's play is the real Richard; certainly the controversy about the portrayal, which Shakespeare borrows directly from Thomas More's history of Richard's reign, assumes that this is so (see also Friend 2003). In that case, if my fictional emotions do not conflict with my ordinary feelings about Richard, this is plausibly because they are quarantined. Whereas my admiration may prompt me to join the Richard III Society, it would make little sense for me to criticize the Society's views based on how I feel while watching the play. I might explain this difference by saying that I do not *really* condemn Richard; in reality I admire him.

There is, however, an objection to treating fictional emotions as quarantined within the imaginative context: that these experiences play a role in our ethical life that cannot be explained unless they are taken to be genuine emotions. If I read a description of a heinous murder in a novel I cannot be faulted for imagining an evil act, but it may seem that I can be faulted for responding with glee. The idea is that we are responsible for the emotions we experience in imagining, insofar as they reveal our characters in a way that the imaginings themselves do not (Moran 1994). In addition, we seem to be able to learn by responding emotionally to fiction, both by broadening the range of our emotional experiences and by acquiring ethical knowledge of the appropriate responses to various situations. If our experiences are not of genuine emotions, it is difficult to see how these kinds of learning are possible; in particular, non-genuine emotions would not seem capable of providing evidence for ethical conclusions (Gaut 2007: 206–207). So even if our imaginings in response to fiction are quarantined from our beliefs, the argument goes, our emotional responses to those imaginings cannot be.

These considerations are not entirely persuasive, however. First, not all responses in the fictional context reveal one's character, or else it would always be inappropriate to laugh at black comedies. It is true that in many cases fictional emotions mirror the emotions we would experience in response to similar, real-life situations; but mirroring of this sort is characteristic of imagining. Quarantine in the relevant sense does not mean total isolation. As for the possibility of learning, there is no doubt that we think of fiction as broadening our emotional experience, and that we sometimes transfer our emotional responses in the fictional context to real-life situations. The question, though, is whether or not these extrapolations are warranted. For whether or not the emotions are genuine they occur in response to imagined situations, and the move from the imagined to the real requires justification (Friend 2010). Ethical considerations therefore do not settle the answer to the descriptive question. They do, however, highlight the second issue concerning fictional emotions: the normative question.

The Normative Question

Just as Walton's counterintuitive conclusion in "Fearing Fictions" prompted the debate over classification, Colin Radford's (1975) argument for a different counterintuitive conclusion prompted the debate over rationality. In "How Can We Be Moved by the Fate of Anna Karenina?" Radford argued that emotions in response to characters we know do not exist and events we know did not occur are "incoherent" and "inconsistent." The implication is that in responding emotionally to fictions we are irrational; though Radford himself does not use this terminology, this is how he is usually understood.

The Standard Interpretation

According to the most popular interpretation of Radford's argument, the inconsistency he identifies is just our commitment to the three incompatible claims that constitute the Paradox of Fiction. Radford defends versions of the three claims in both his original paper and in subsequent replies to objections (e.g., Radford 1982, 1989, 1995, 2000). In accord with (1) he criticizes attempts to distinguish fictional emotions from ordinary emotions; even if they differ in some respects, they do not differ sufficiently to be subject to completely different considerations of consistency. In accord with (3) he argues that we do not, even temporarily, believe that the fiction is true; and if we did, we would have contradictory beliefs leading to a different form of incoherence. Finally he claims that "I can only be moved by someone's plight if I believe that something terrible has happened to him" (Radford 1975: 68). This is understood as a formulation of narrow cognitivism in defense of (2).

The usual reply to Radford's argument is to deny that we are, in fact, committed to all three claims in the Paradox of Fiction. Since the incoherence is supposed to stem from our commitment to these three inconsistent claims, this would avoid the charge of irrationality. From this perspective, the descriptive and normative questions are addressed in exactly the same way. Denying that fictional emotions are genuine, or arguing that emotions do not require beliefs, would on this interpretation remove the threat of maintaining inconsistent beliefs.

However, this interpretation fails to capture Radford's puzzle. Radford's ardent defense of (1) shows that he cannot be committed to narrow cognitivism. Rather, his claim in (2) is that *in ordinary contexts* we are not moved when we lack the relevant beliefs (cf. Teroni and Cova, ms). This claim is consistent with both (1) and (3) and therefore produces no paradox. Radford argues for this claim with a series of examples. Most famously, he points out that if he hears a harrowing story about someone's sister he will be harrowed, but this feeling disappears—to be replaced perhaps by indignation—if he then learns that the sister does not exist. The surprise is not the key factor. If we know in advance that our friend is merely acting out the experience of pain, we will not pity her. In such cases, pity in the absence of relevant beliefs would not just be unwarranted, it would be incomprehensible. Yet in the fictional context we are systematically moved even though we do not believe anyone has suffered. That is what Radford finds puzzling.

Radford's Normative Argument

If it is not the Paradox of Fiction that explains our inconsistency and incoherence, what is it? Radford's focus is the experience of fictional emotions themselves. Regardless of our theoretical position on the three claims that constitute the Paradox—even if, for instance, we follow Walton in rejecting (1)—Radford claims that we would still be incoherent in responding emotionally to fiction. The explanation must be that in so responding we violate a *norm* that governs emotions (cf. Gendler and Kovakovich 2006). Radford's examples indicate that in ordinary contexts, we violate the norm when we do not believe that there exists an appropriate object for the emotion. So the relevant norm might be that an experience of an emotion presupposes that such an object, with the relevant properties, exists—or at least presupposes the belief that it exists—and is otherwise defective. If this is a genuine normative constraint on emotions, there is something deeply incoherent about having an emotion while believing that there is no appropriate object. It would be like intending to perform an action while believing oneself incapable of performing it.

It is widely accepted that emotions are subject to normative constraints. For example, to be appropriate an emotion must fit its formal object, the property ascribed to the target in normal occurrences of the emotion: fear for danger, pity for suffering, and so on. Emotional responses that fail to meet this condition, such as elation at the death of a close friend or terror of a teddy bear, are the ones most likely to be called irrational. An additional dimension of fit is proportionate intensity; mild annoyance may fit a situation that boiling anger does not. The fictional emotions that Radford discusses arguably meet this constraint. The problem with pity of Anna Karenina is not that it is directed at, say, a happy situation or one that invites only mild concern; the formal object is intense suffering, even if no actual object suffers.

A different condition for the appropriateness of emotions is correctness (also called "fittingness"; see D'Arms and Jacobson 2000), which is associated with objective rationality. An emotion is correct so long as the evaluation associated with the emotion is accurate. I may have excellent reason to think that my friend Andrew betrayed me, but anger at him is incorrect if he is in fact innocent. Even if fictional emotions fail to meet this condition—since no appropriate object exists—this cannot be the source of Radford's claim that they are incoherent. My anger at Andrew coheres with my beliefs regardless of whether it turns out to be undeserved. Similarly, a child's feelings about Santa Claus may be incorrect since there is no Santa, but they make sense in light of her beliefs.

Radford seems rather to be concerned that fictional emotions fail to meet a condition of *justification* (sometimes also called "warrant"), associated with subjective rationality. An emotion is justified so long as the evaluation associated with it is adequately responsive to one's evidence about the situation. If I have good reasons to think I am in danger, or am sufficiently sensitive to danger cues, fear is justified; if I have every reason to think I am safe, fear is unjustified. Emotions may be more or less justified insofar as evidence is stronger or weaker. But they will seem irrational if the evidence is all on the opposite side. Once I know that Andrew is innocent, continuing anger would be entirely unjustified and probably irrational. Similarly, anyone who knows Santa does not exist but still harbors hopes of his largesse can be described as incoherent. Since this seems to be our position with fictional emotions—we are fully aware that no appropriate object exists—Radford takes them to violate this normative constraint.

Objections to Radford

There are objections to accepting such a normative constraint on emotions, however. Coherence and consistency are clear requirements for rational belief, but the same does not seem to apply to emotion. For example, we often experience "mixed feelings," as when we are happy for a friend who wins a prize but unhappy because we would have liked to win. By contrast with beliefs, it may be more rational to maintain both emotions rather than trying to resolve the conflict, because doing so strikes the right balance between our interests in ourselves and in others (Greenspan 1980, 1988). In this way emotions in response to fiction might be "emotionally rational" insofar as they are socially adaptive, getting us to identify and sympathize with other people (Robinson 2005, 146–148). Their being rational in one sense is, however, consistent with their being irrational in a different sense, as Robinson herself points out. As we have seen, in ordinary contexts emotions without the relevant beliefs are unjustified. If fictional emotions are not subject to the same normative constraints as such emotions, this cannot be because of general features of emotional rationality.

A different objection to the normative constraint focuses on the role played by emotional responses to the merely imagined in reasoning. For example, successful planning typically involves imagining different possibilities and choosing among them, something we accomplish in part by responding affectively to the possible outcomes (Currie and Ravenscroft 2002: 98). Gendler and Kovakovich (2006) appeal to experiments by Antonio Damasio and his colleagues to argue that emotions in response to the merely imagined are essential to rational decision-making. The experiments compare normal individuals to patients with damage to the ventromedial prefontal cortex in a gambling task. Normal participants' performance improves once they begin to respond affectively in anticipation of high-risk choices as measured by skin conductance, even before they can explicitly identify which options are high-risk. The performance of the patients with damage, who fail to experience negative affect, never improves even when they are consciously aware that they are making high-risk choices. Gendler and Kovakovich conclude that because anticipatory affect plays a key role in decision-making, it contributes to rationality rather than undermining it.

These cases do not undermine Radford's claim that fictional emotions are unjustified, however. First, the low-level "gut reactions" described by Damasio are closer to reflex responses than to genuine emotions. Second, the results show only that certain affective responses are a precondition of making good decisions, not that they contribute to the goodness of those decisions. In other cases our gut reactions may lead us wildly astray. A similar point can be made about the use of heuristics in reasoning. Deploying simple heuristics is adaptive insofar as it results in quick judgments under conditions of uncertainty, but in at least some cases the heuristics generate wrong judgments, most notoriously with respect to probability (Tversky and Kahneman 1974). Finally, even if affective responses to imagined scenarios contributed positively to planning and decision-making in every case, this would indicate only that they were rational in the sense of being instrumentally valuable (cf. Kim 2010). Radford (1989) himself takes the experience of fictional emotions to constitute one of the great pleasures in life. As we have seen, it is consistent with Radford's argument that fictional emotions be rational in one sense but not in another.

A more significant concern is that we do not criticize ourselves or each other for responding emotionally to fiction (though overreactions might be subject to criticism). By itself this observation does not undermine Radford's argument, for we could be subject to norms we do not recognize. My beliefs should be consistent whether or not I realize this and my reasoning is constrained by the laws of probability whether I am aware of them or not. However, in the case of faulty reasoning or inconsistent beliefs, I can in principle be brought to see where I have gone wrong. In the case of fictional emotions, we remain convinced that our responses are unproblematic even upon reflection, when we are fully aware that the objects of emotion do not exist and thoroughly familiar with Radford's arguments. In other words, the norms governing fictional emotions seem to differ from those governing emotions in ordinary contexts. Saying this does not resolve the problem, though, without an explanation of why. In the absence of such an explanation we are simply distinguishing fictional emotions from other emotions on the grounds that only the latter require beliefs in existence; this looks like an ad hoc maneuver to avoid Radford's conclusion that fictional emotions are incoherent without such beliefs.

One explanation of why only certain emotional responses seem to be subject to a norm of consistency with beliefs might be the frequently noted connection between emotion and the motivation to act. In contexts where the demand for action is pressing or immediate, as when we are confronted directly with a situation, our emotional responses presuppose the existence of the appropriate objects. In fictions about purely invented characters and events, where action is impossible, we are not subject to the same constraint.

If this is right, rather than a dichotomy between emotions in the fictional context and emotions in other contexts we should find a range of cases between the two extremes. For instance, although direct action is typically ruled out in engagement with nonfiction representations, indirect actions are possible and subject to a justification condition. In response to a sympathetic history of Richard III (or even a historically researched fiction) I might decide to join the Richard III Society and work to exonerate him; this only makes sense if I believe the work to be accurate. Similarly, even though planning involves the imagination of potential outcomes, insofar as action is the goal our affective responses must be constrained by beliefs about what is possible. In planning a dinner party I might be elated picturing myself creating wonderful dishes with a *Bewitched*-style crinkle of the nose, but it would be irrational to take this feeling as a guide to action. On this picture, if Radford is wrong it is not because no emotions are constrained by beliefs, but rather because fictional emotions typically lack the connections to action that characterize emotions in other contexts.

Norms for Fiction

Although fictional emotions are not subject to the same norms of consistency with beliefs as other emotions, they are subject to other norms (cf. Gilmore 2011). We have already seen that such emotions should fit their formal objects. Pitying Ivan Denisovich is appropriate, whereas envying his life in the gulag is not. In addition, we can specify versions of the conditions of justification and correctness for the fictional context (Friend 2010). An emotion counts as fictionally justified so long as it is reasonable in light of one's fictional information; justification reflects competence at deploying the "principles of generation" (Walton 1990) that generate fictional truths. An emotion is fictionally correct so long as the associated evaluation accurately represents the fictional situation, even if one is unaware of this. Being worried that Mr. Hyde will harm Dr. Jekyll is fictionally justified up to a certain point in Stevenson's novella, but given that they are the same person it is fictionally incorrect. By contrast detesting Sethe for killing all her children (rather than just Beloved) is neither correct nor justified.

These normative constraints reflect the idea that our emotional responses to fiction typically follow our emotional responses to real life. But there are further conditions that fictional emotions may need to meet that do not apply to emotions in other contexts (Currie 1990: 213–215; Livingston and Mele 1997). For example, genre conventions play a role. Feeling depressed because of the death and destruction wrought by the Killer Rabbit in Monty Python's *The Holy Grail* is inappropriate, even if such carnage would warrant bleak despair and terror in reality. Black comedies typically invite responses such as amusement that are neither justified nor correct in light of the fictional truth (Currie 1990: 213). Similarly, artistic skill is relevant. The justified and correct response to a scene of someone dying of a terminal illness might be sadness, but if the film is mere sentimental rubbish this response will not be appropriate. Some philosophers also take moral considerations to be relevant. If a story invites us to sympathize with an evil character, we might treat this response as "unmerited" insofar as it is unethical (Gaut 1998). Others argue, though, that entering into immoral points of view, and experiencing the attendant emotions, is a potentially valuable feature of engaging with fiction (Kieran 2003; Eaton 2012). If this is right, it is another way in which the normative constraints on fictional emotions depart from the constraints on emotions in other contexts.

Conclusion

Emotional responses to fiction raise two distinct questions: the descriptive question, which concerns whether these responses constitute genuine emotions of the same sort as emotions in other contexts; and the normative question, which concerns the rationality or appropriateness of these emotions. According to the standard interpretation, both questions can be answered by rejecting one of the three claims that constitutes the Paradox of Fiction. I have argued, however, that although resolving the paradox is relevant to the descriptive question, it is not the right strategy for addressing the normative question. Nonetheless imagination plays a key role in how we approach both questions, by providing a context within which fictional emotions may be quarantined and by generating distinctive norms for evaluating those emotions.

Acknowledgements: I would like to thank Amy Kind and Shen-yi Liao for helpful comments on a previous draft. I would also like to thank audiences at conferences in Paris, Leeds, and Ann Arbor, Michigan, for discussion of related material.

References

Carroll, Noel. 1990. *The Philosophy of Horror: Or, Paradoxes of the Heart*. New York: Routledge.

Cova, Florian, Amanda García, and Vanessa Sennwald. ms. "Is What I'm Feeling Really Genuine? Fiction vs. Reality."

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

Currie, Gregory, and Ian Ravenscroft. 2002. *Recreative Minds: Imagination in Philosophy and Psychology*. Oxford: Oxford University Press.

D'Arms, Justin, and Daniel Jacobson. 2000. "The Moralistic Fallacy: On the 'Appropriateness' of Emotions." *Philosophy and Phenomenological Research* 61: 65–90.

Eaton, Anne W. 2012. "Robust Immoralism." *Journal of Aesthetics and Art Criticism* 70: 281–92.

Friend, Stacie. 2003. "How I Really Feel about JFK." In *Imagination, Philosophy and the Arts*, edited by Matthew Kieran and Dominic Lopes, 35–53. London: Routledge.

———. 2010. "Getting Carried Away." *Midwest Studies in Philosophy* 34 (Film and the Emotions): 77–105.

Gaut, Berys. 1998. "The Ethical Criticism of Art." In *Aesthetics and Ethics: Essays at the Intersection*, edited by Jerrold Levinson, 182–203. Cambridge: Cambridge University Press.

———. 2003. "Reasons, Emotions, and Fictions." In *Imagination, Philosophy and the Arts*, edited by Matthew Kieran and Dominic Lopes, 14–34. New York: Routledge.

———. 2007. *Art, Emotion and Ethics*. Oxford: Oxford University Press.

Gendler, Tamar Szabó. 2003. "On the Relation Between Pretense and Belief." In *Imagination, Philosophy, and the Arts*, edited by Matthew Kieran and Dominic McIver Lopes, 125–41. London: Routledge.

Gendler, Tamar Szabó, and Karson Kovakovich. 2006. "Genuine Rational Fictional Emotions." In *Contemporary Debates in Aesthetics and the Philosophy of Art*, edited by Matthew Kieran, 241–53. Oxford: Blackwell.

Gilmore, Jonathan. 2011. "Aptness of Emotions for Fictions and Imaginings." *Pacific Philosophical Quarterly* 92: 468–89.

Greenspan, Patricia. 1980. "A Case of Mixed Feelings: Ambivalence and the Logic of Emotion." In *Explaining Emotions*, edited by A. O. Rorty, 223–50. Berkeley and Los Angeles: University of California Press.

Greenspan, Patricia S. 1988. *Emotions and Reasons: An Inquiry Into Emotional Justification*. New York: Routledge.

Hartz, Glenn A. 1999. "How We Can Be Moved by Anna Karenina, Green Slime, and a Red Pony." *Philosophy* 74: 557–78.

Jacobson, Daniel. 1997. "In Praise of Immoral Art." *Philosophical Topics* 25: 155–99.

Kieran, Matthew. 2003. “Forbidden Knowledge: The Challenge of Immoralism.” In *Art and Morality*, edited by Jose Luis Bermudez and Sebastian Gardner, 56–73. London: Routledge.

Kim, Seahwa. 2010. “The Rationality of Emotion toward Fiction.” *Midwest Studies In Philosophy* 34 (Film and the Emotions): 106–19.

Lamarque, Peter. 1981. “How Can We Fear and Pity Fictions?” *British Journal of Aesthetics* 21: 291–304.

LeDoux, Joseph. 1996. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon & Schuster.

Livingston, Paisley, and Alfred L. Mele. 1997. “Evaluating Emotional Responses to Fiction.” In *Emotion and the Arts*, edited by Mette Hjort and Sue Laver, 157–76. New York: Oxford University Press.

Lyons, William E. 1980. *Emotion*. Cambridge: Cambridge University Press.

Matravers, Derek. 1998. *Art and Emotion*. Oxford: Oxford University Press.

Moran, Richard. 1994. “The Expression of Feeling in Imagination.” *The Philosophical Review* 103: 75–106.

Neill, Alex. 1991. “Fear, Fiction and Make-Believe.” *The Journal of Aesthetics and Art Criticism* 49: 47–56.

Nussbaum, Martha C. 2001. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge University Press.

Oakley, Justin. 1992. *Morality and the Emotions*. London: Routledge.

Prinz, Jesse J. 2004. *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.

Radford, Colin. 1975. “How Can We Be Moved by the Fate of Anna Karenina?” *Proceedings of the Aristotelian Society* 49 (Suppl.): 67–80.

———. 1982. “Stuffed Tigers: A Reply to H. O. Mounce.” *Philosophy* 57: 529–32.

———. 1989. “Replies to Three Critics.” *Philosophy* 64: 93–97.

———. 1995. “Fiction, Pity, Fear, and Jealousy.” *Journal of Aesthetics and Art Criticism* 53: 71–75.

———. 2000. “Neuroscience and Anna; a Reply to Glenn Hartz.” *Philosophy* 75: 437–40.

Robinson, Jenefer. 2005. *Deeper than Reason: Emotion and Its Role in Literature, Music, and Art*. Oxford: Oxford University Press.

Solomon, Robert C. 2003. *Not Passion’s Slave: Emotions and Choice*. New York: Oxford University Press.

Stecker, R. 2011. “Should We Still Care About the Paradox of Fiction?” *British Journal of Aesthetics* 51: 295–308.

Stocker, Michael. 1987. “Emotional Thoughts.” *American Philosophical Quarterly* 24: 59–69.

Suits, David B. 2006. “Really Believing in Fiction.” *Pacific Philosophical Quarterly* 87: 369–86.

Teroni, Fabrice, and Florian Cova. ms. “Why There Still Is a Paradox of Fiction.”

Tversky, Amos, and Daniel Kahneman. 1974. “Judgment under Uncertainty: Heuristics and Biases.” *Science* 185: 1124–31.

Walton, Kendall L. 1978. “Fearing Fictions.” *Journal of Philosophy* 75: 5–27.

———. 1990. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press.

Weinberg, Jonathan M., and Aaron Meskin. 2006. “Puzzling Over the Imagination: Philosophical Problems, Architectural Solutions.” In *The Architecture of the*

Imagination: New Essays on Pretence, Possibility, and Fiction, edited by Shaun Nichols, 175–202. Oxford: Oxford University Press.

Further Reading

Boruah, Bijoy H. 1988. *Fiction and Emotion: A Study in Aesthetics and the Philosophy of Mind*. Oxford: Oxford University Press.

Dadlez, Eva M. 1997. *What's Hecuba to Him?: Fictional Events and Actual Emotions*. University Park: Pennsylvania State University Press.

Davies, Stephen. 2009. "Responding Emotionally to Fictions." *The Journal of Aesthetics and Art Criticism* 67: 269–84.

Feagin, Susan L. 1988. "Imagining Emotions and Appreciating Fiction." *Canadian Journal of Philosophy* 18: 485–500.

Joyce, R. 2000. "Rational Fear of Monsters." *British Journal of Aesthetics* 40: 209–24.

Matravers, Derek. 1997. "The Paradox of Fiction: The Report versus the Perceptual Model." In *Emotion and the Arts*, edited by Mette Hjort and Sue Laver, 78–92. New York: Oxford University Press.

———. 2005. "The Challenge of Irrationalism and How Not to Meet It." In *Contemporary Debates in Aesthetics and the Philosophy of Art*, edited by Matthew Kieran, 254–64. Oxford: Blackwell.

Meskin, Aaron, and Jonathan M. Weinberg. 2003. "Emotions, Fiction, and Cognitive Architecture." *The British Journal of Aesthetics* 43: 18–34.

Schaper, Eva. 1978. "Fiction and the Suspension of Disbelief." *British Journal of Aesthetics* 18: 31–44.

Schneider, Steven. 2002. "The Paradox of Fiction." *Internet Encyclopedia of Philosophy*. Accessed 6 September 2014. http://www.iep.utm.edu/fict-par/.

Tullmann, Katherine, and Wesley Buckwalter. 2014. "Does the Paradox of Fiction Exist?" *Erkenntnis* 79: 779–96.

Walton, Kendall. 1997. "Spelunking, Simulation, and Slime: On Being Moved by Fiction." In *Emotion and the Arts*, edited by Mette Hjort and Sue Laver, 37–49. New York: Oxford University Press.

Yanal, Robert J. 1999. *Paradoxes of Emotion and Fiction*. University Park: Pennsylvania State University Press.