Resolving the Paradox of Fiction: A Defense of Irrationalism

Anna Brinkerhoff

Abstract: In this paper, I examine the Paradox of Fiction: (1) in order for us to have genuine and rational emotional responses to a character or situation, we must believe that the character or situation is not purely fictional, (2) we believe that fictional characters and situations are purely fictional, and (3) we have genuine and rational emotional responses to fictional characters and situations. After defending (1) and (2) against formidable objections and considering the plausibility of ~(3) in isolation of (1) and (2), I conclude that we should resolve the Paradox of Fiction by rejecting (3).

The so-called Paradox of Fiction is an inconsistent triad of propositions regarding our emotional responses to fiction: (1) in order for us to have genuine and rational emotional responses to a character or situation, we must believe that the character or situation is not purely fictional, (2) we believe that fictional characters and situations are purely fictional, and (3) we have genuine and rational emotional responses to fictional characters and situations. At first blush, all three propositions seem true. But, since they cannot all be true, resolving the paradox requires that we figure out which proposition to reject.

The thesis of this paper is that (3) is false in that our emotional responses to fictional characters and situations are not rational. In what follows, I will expound on the paradox itself, clarify what it means to have emotional responses, and lay out conditions for rational emotional responses. With this is mind, I will then consider what I take to be the most daunting objections to (1) and (2). Next, I will defend each proposition in light of its respective challenges. Finally, I will advocate the plausibility of ~(3) independently of the other two propositions that entail it.

The Paradox

Emotional engagement with fictional narrative, presented in novels, plays, or movies, is a familiar phenomenon: we pity the tragic hero upon his downfall, fear the malevolent monster as it lurks behind the bushes, and rejoice with the bride as she kisses her prince. The question arises: how is it rational to respond emotionally to characters or situations that we believe not to exist and to never have existed? Perhaps it is not rational; in other words, it may be that propositions (1) and (2) of the paradox are premises in an argument whose conclusion is the negation of (3):

> F1. If we have genuine and rational emotional responses to a character or situation, then we must believe that the character or situation is not purely fictional.
>
> F2. We believe that fictional characters and situations are purely fictional.
>
> F3. Therefore, we do not have genuine and rational emotional responses to fictional characters and situations.[1]

This is the argument I seek to defend. Before proceeding, it is important to note what, in my view, makes (3) false. Supporters of F3 can negate (3) by claiming that we do not have emotional responses to fiction; that we have emotional responses to fiction, but they are neither genuine nor rational; that we have genuine emotional responses to fiction, but they are irrational; or that we have rational emotional responses to fiction, but they are ingenuine. My claim is that (3) is false because we have genuine but irrational emotional responses to fiction. This view is called Irrationalism. More specifically, Irrationalism is the idea that it is irrational to have fictional characters and situations as the objects of our emotions.[2] The argument above in favor of

[1] F1 is (1), F2 is (2), and F3 is ~(3). For the remainder of the paper, F1 will be used interchangeably with (1), F2 with (2), and F3 with (~3).

[2] Throughout the remainder of the paper, when I refer to our emotional responses to fiction, I am referring to our emotional responses that have a fictional character or situation as their object.

Irrationalism is indisputably valid, but both premises are debatable. Before evaluating them, I will clarify pivotal terms in the section below.

Emotional Responses to Fiction and Conditions for Rational Responses

It is beyond the scope of this paper to offer or recommend an analysis of emotion; however, in order to eschew ambiguity, I will flesh out what it means to respond emotionally to fiction. For the purposes of this paper, I am using the strict sense of "emotion." This excludes moods, attitudes, and dispositions. A consequence of this is that the emotions to which I am referring must have an object. Thus, on this account, mental states such as regret, irritation, and fear count as emotions; but depression, apathy, and euphoria do not. With this in mind, suppose that Person S reads a fictional narrative in which one of the characters, Character A, is an innocent victim of domestic abuse. The narrative causes a mental state in S that seems to her like pity and which she readily identifies as pity. This emotional response of pity is genuine insofar as the way in which it seems to S is identical to the way in which the same emotion would seem to her were Character A a real person in the situation stipulated. Hereafter, I assume that the emotional responses we have to fictional narratives are genuine; this is a safe assumption given that the emotions we experience when consuming fiction tend to be phenomenologically indistinguishable from those when we hear about or witness real-life narratives.

What, then, does it mean for any genuine emotional response to be rational? I adopt (and slightly revise) the conditions that Derek Matravers, who endorses Irrationalism, proposes, where E is an instance of emotion: (a) S is justified in being in the cognitive state (e.g., believing that P, understanding that P, entertaining the thought that p, etc.) that elicits E, (b) E is a reasonable response given the cognitive state, and (c) E is of an appropriate intensity.[3] If an emotional response meets all three of these conditions, then it is rational. To clarify by example, if I peer through my window and see a swirling funnel cloud descend to the earth and swell with debris, I am justified in forming the belief that there is a tornado in the vicinity. Thus, I am justified in being in the cognitive state—believing—that elicits my emotional

[3] Derek Matravers, "The Challenge of Irrationalism, and How Not To Meet It," in *Contemporary Debates in Aesthetics and the Philosophy of Art*, ed. Matthew Kieran (Malden: Wiley-Blackwell, 2005), 254-64.

response—fear (condition a). Fear is a reasonable response given my belief that there is potentially devastating weather nearby (condition b). Gratefulness, however, is not. Also, depending on the particulars of the circumstance I am in, fear of a wide range of intensity is appropriate (condition c). So, in the situation postulated, my emotional response is rational.

Applying these conditions to S's emotional response to the narrative about Character A, S is justified in imagining that Character A is in unfortunate circumstances, as this is what the author prompts her to imagine. *Prima facie*, the misfortune and suffering that Character A undergoes render S reasonable in directing pity at A, and deep pity is of an appropriate intensity given the despondency of A's circumstances. Matravers argues, however, that responding emotionally to fiction is irrational because, in doing so, condition (b) is violated; in his words, pity of any intensity is not a reasonable response to "a proposition imagined rather than a proposition believed."[4] Before assenting to Matravers's view that emotional responses to propositions imagined are irrational, I want to examine other purported solutions to the paradox. It is by virtue of considering premises F1 and F2 that the irrationality of emotional responses to fictional characters and situations is illuminated. I will begin by addressing F1.

Objection to Premise F1

In an attempt to solve the Paradox of Fiction, some philosophers have denied premise F1. Their claim is that it is possible for S to have genuine, rational emotional responses to characters and situations that she believes do not exist and to never have existed. Noël Carroll defends this position. His view is that the mere thought of the vindictive masked killer is what inspires fear in person S when she watches a horror film or that the thought of Character A in vastly unfortunate situations is what motivates S to feel pity. Moreover, these thoughts are the objects of the emotion; borrowing his terminology, S fears and pities, respectively, the content of her thought. In articulating his view, Carroll writes that "with respect to fictions, the author of such works presents us with conceptions of things to think about . . . . And in entertaining and reflecting upon the contents of these representations, which supply us with the contents of our thoughts, we can be moved

[4] Ibid., 257.

to pity, grief, joy, indignation, and so on."[5] Contra Matravers, Carroll argues that actual pity, grief, joy, and indignation can be rational responses to a proposition entertained in thought, or imagined, rather than a proposition believed. Emotional responses to fiction are not irrational, he argues, because the thought contents are not based on psychotic or neurotic fantasies; the thought theory does not compel the consumer of fiction to embrace a contradiction; responding to fiction emotionally is normal and a natural component of our emotional and cognitive structure; and, when consumed as intended, such emotional responses do not interfere with practical pursuits.

I agree with Carroll that we can have emotional responses to things that we do not think exist. Furthermore, I am also inclined to accept his view that, when responding emotionally to fiction, the objects of our emotions are mental representations of content prompted by engaging with the fiction at hand. Yet, Carroll's reasons for thinking that these emotional responses are rational are inadequate. I will address this in the next section.

Defense of Premise F1

Carroll argues that it is possible to have genuine, rational emotional responses to characters and situations that we believe do not exist because, when consuming fiction, we entertain in thought these characters and situations. These thoughts are the objects of our emotions. Because it is rational to emotionally respond to the thought of something, it is rational to respond to the thought of fictional characters and situations. It is important to emphasize here that the question is not whether the mere thought of something can generate an emotional response; rather, it is whether such emotional responses are rational, specifically when fictional characters and situations generate them. Despite cataloging a plethora of ways in which responding emotionally to the thought of fictional characters and situations is not irrational, Carroll neglects the one that underlies his theory: Matravers's condition (b) of rationality. The question remains: is the emotion in question a reasonable response given the cognitive state at hand?

---

[5] Noël Carroll, *The Philosophy of Horror or Paradoxes of the Heart* (New York: Routledge, 1990), 88.

In order to get a better grasp on condition (b), I need to explain what it means for an emotional response to be reasonable for S. In order for S to be reasonable in responding with pity to something or someone, that something or someone must warrant S's pity. If S responds with pity to something or someone that does not warrant it, S's response is unreasonable. For something or someone to warrant S's pity, he/she/it must undergo, have undergone, or will undergo misfortunate and suffer, have suffered, or will suffer accordingly. Similarly, in order for S to be reasonable in responding with anger to something or someone, he/she/it must have wronged or offended S. Likewise, in order for S to be reasonable in responding with fear to something or someone, he/she/it must threaten to harm S. There are parallel standards for regret, indignation, grief, joy, and a host of other emotions. Again, the issue is whether the emotion in question is a reasonable response given the cognitive state at hand. On Carroll's account, the relevant cognitive state is entertaining in thought, or imagining. Whereas the object of emotion when believing that P is the thing in the world that [S thinks] corresponds to the object of emotion when entertaining in thought that P is the thought itself; there is nothing in the world to which S thinks the content of entertaining in thought that P corresponds. Because the thought itself does not undergo misfortune, wrong or offend S, or threaten to harm S, it is not reasonable, and, by extension, irrational, for S to respond with pity, anger, or fear. In light of this, Carroll is wrong in claiming that it is possible to have rational emotional responses to situations or characters that we know not to exist. What is more, it seems that E is a reasonable response for S only if S's cognitive state is belief that P. When S believes that Pm only then is the object of her emotion something she takes to be in the world, and only some things in the world can endure suffering and misfortune, wrong, offend, or harm.

Objection to Premise F2

Another approach to resolving the paradox is to deny F2. The claim of philosophers who pursue this strategy is that S believes that the fictional characters and situations are real. As David Suits, a thoughtful supporter of this position, puts it, "[S] believes that the persons in the story are there, that [she is] in the places described in the story, and that the events of the story are occurring exactly as

described."[6] On Suits's account, S believes that Character A is real and that A is a victim of domestic abuse. Suits attends to the predictable objection: it is not the case that S believes that they are real because, if she did, she would react differently than she does upon encountering them. For example, if S thought that the malevolent monster on the big screen were real, she would flee. Since she does not flee, it is not the case that she thinks the monster is real.

Suits argues that those who make this inference are ignoring the context of belief. He claims that, when S is engrossed in the horror movie featuring the monster, she peripheralizes her physical situation. Before and after engaging with the narrative, S believes that the characters and situations are fictional; however, during her engagement, she forgets this, or perhaps intentionally suspends this belief. Suits's argument for why S does not flee even though she thinks the monster is real is this: in situations where S is actually confronted by a monster, it is not clear that she would flee because her reaction is contingent upon her other beliefs. In the scenario in which she is sitting at the cinema, S holds beliefs that counteract her impulse to flee.

Defense of Premise F2

This argument misrepresents what is actually going on. It seems right that, when caught up in a narrative, S is not attending to the fact that the characters and situations are purely fictional; however, this does not entail that she assents to the proposition that the situations and characters are real, as Suits implies. To be charitable, though, let us suppose with Suits that, when captivated by the narrative, S does believe the monster is real. Before going further, here it is helpful to introduce Matravers's idea of instrumental belief. According to Matravers, an instrumental belief allows us to act toward the object of our emotion.[7] Suppose that we pity the malnourished homeless man begging for money. The belief that gifting him a ten-dollar bill would ameliorate his suffering is an instrumental belief. Returning to Suits, his account is problematic because it seems that the following conditional is true: if S has the emotional response of fear and has an

[6] David Suits, "Really Believing in Fiction," *Pacific Philosophical Quarterly 87*, no. 3 (2006): 371.

[7] Matravers, "The Challenge of Irrationalism, and How Not To Meet It," 255.

instrumental belief that, if acted upon, she thinks will decrease her risk of harm, then, assuming that she is physically and psychologically able and that she has no overriding reason to do otherwise, she will act on the instrumental belief. When confronted with this conditional, Suits would claim that S does not flee because she has an overriding reason, namely that she believes that she is in a cinema and that the monster cannot breach the screen.

To understand the glitches in this rejoinder, it is beneficial to specify the supposed process of S's switching between believing that the monster is real and believing that it is fictional that Suits's account demands. Engrossed in the movie, S believes that the monster is real. This belief occasions fear and, naturally, she poises to flee. Preparing to flee distracts her from the film, and so she does not flee because the fact that the monster is fictional is remembered. As it is unreasonable to assume that S would conscientiously hold two propositions that are blatantly inconsistent with each other simultaneously, it can be concluded that S no longer believes that the monster is real. If this pattern holds, then, when S believes that the monster is real, she fears that the monster will harm her. In response to that fear, she prepares to flee, but, on the brink of fleeing, she is tempered by, and thereby reminded of, the fact that the monster is not real. By virtue of this, she is distracted from the film. No longer captivated, she recalls that it is purely fictional and so stays seated. Granting that it is the case that S believes at moments that the monster is real, this belief is so stunted by fear and urge to flee that it is insignificant. Rather than assenting to Suit's analysis, it is simpler and more plausible to say that, although she experiences genuine fear when watching the horror movie, she does not flee because she knows throughout that it is fictional. Either way, F2 is not undermined. Now that F1 and F2 have been defended, I will consider Irrationalism on its own.

The Plausibility and Palatability of Irrationalism

The potency of the paradox is derived from the fact that, when considered in isolation, (1), (2), and (3) all seem plausible. Even though I have argued that (1) and (2) are true, and that implies that (3) is false, it will be helpful to consider the plausibility of Irrationalism in isolation. Irrationalism is the view that S's emotional responses to fictional characters and situations are irrational; specifically, I have argued that such emotional responses are irrational insofar as the

object of the emotion is not something S takes to be in the world. It is not the view that we do not have genuine emotional responses to fiction, nor is it that no good can arise from engaging emotionally with fictional narratives. Perhaps it is by dint of this irrational element of our cognitive and emotional structure that we gain experiential knowledge or expand our understanding of the human condition.

Furthermore, the Irrationalism I endorse does not condemn all emotional responses prompted by fiction as irrational. To see this, consider the Counterpart Theory that Gregory Currie recommends, which accounts for what happens when S emotionally engages with fiction. He claims that "we experience genuine emotions when we encounter fiction, but their relation to the story is causal rather than intentional; the story provokes thoughts about real people and situations, and these are the intentional objects of our emotions."[8] On Currie's view, when S consumes fiction, the fictional characters and situations cause her pity, but what she really pities are situations and people she believes exist. Admittedly, Currie's interpretation of what happens when we read fiction is not true of most cases. When I read *Pride and Prejudice*, I am not happy for my cousin who defied societal norms to marry her beloved; nor do I consider Mr. Darcy and Elizabeth tokens of a type of which there are real-life tokens that are the objects of my happiness. Rather, Mr. Darcy and Elizabeth and their unlikely love (or, more precisely, the mental representations of these prompted by the text) are the objects of my happiness, even though I know they are fictional. From what I can tell, most consumers respond in the same way I do, with fictional characters and situations as the object of their happiness. Perhaps, though, Currie's interpretation is true of some cases; indeed, it is possible for fictional characters and situations to cause an emotion that has a real-life counterpart as its object. When this does happen, emotional responses caused by fiction are rational. It is important to emphasize that this does not undermine Irrationalism. When the object of our emotions is a counterpart of a fictional character or situation, the emotion is merely triggered by the fiction, but the fictional character or situation is not the object of our emotion. In other words, in these rare cases, our emotional response is rational because it was caused by but not to fiction. My thesis remains formidable.

---

[8] Gregory Currie, *The Nature of Fiction*, (Cambridge: Cambridge University Press, 1990), 188.

The last thing I want to note about Irrationalism's plausibility is that it is not necessary to prove F1 and F2 to maintain Irrationalism; all one has to do is show Matravers's conditions of rationality are sensible to accept and that emotional responses to fictional characters and situations do not meet all three. In light of all this, it seems that the stripe of Irrationalism that I support is plausible in isolation and, additionally, is quite palatable.

Conclusion

The most plausible and palatable way to resolve the Paradox of Fiction is to embrace Irrationalism. Propositions (1) and (2) emerge unscathed from formidable objections; thus, the two premises of the argument that entail the negation of (3), F1 and F2, respectively, seem true. Moreover, when evaluated in isolation of the other two propositions, the negation of (3) seems plausible and, ultimately, a relatively soft bullet to bite to resolve the Paradox of Fiction.[9]

[9] I would like to dedicate this paper to my dad (1957 - 2013): my biggest supporter, fellow wonderer, and funder of my library. I love you and miss you lots!