

# RECLAIMING THE PARADOX OF FICTION

Mary Beth Willard  
*Weber State University*

1

It is a truism of modern parenting that entertainment aimed at children needs to be carefully screened to avoid violence, sexual situations, and coarse language. Fortunately, there is a wealth of programming that avoids all of this. What is desperately needed, however, when choosing entertainment for a four-year-old, is a ratings system aimed at minimising suspense. My son, for example, finds any kind of suspense, any kind of tension or fictional peril, to be nearly unbearable. He immediately wants to stop watching because he fears too much for the well-being of the characters.

In children's movies, dramatic tension usually resolves happily. Bad guys don't win. But my sensitive son resonates like a theremin to the tension of the narrative and wails: "I don't like this! Turn it off!"<sup>1</sup>

As a parent, my thoughts turn to how to calm him down so that he might successfully finish a film. As a philosopher, my thoughts turn to the paradox of fictional emotions, standardly formulated as:

1. We have genuine and rational emotional responses towards fictional characters and situations.

---

<sup>1</sup> Lest any reader be concerned that I am mocking my dear son for the sake of a lively philosophical introduction, I confess that he comes by his thereminic qualities honestly. At age five I ran out of a theatre because the villain was making Rainbow Brite scrub the floors.

2. We believe that fictional characters and situations do not exist in reality.
3. In order for us to have genuine and rational emotions towards a given character or situation, we must believe that they exist in reality.<sup>2</sup>

All three are plausible. Any two can be held consistently. Holding all three together results in a contradiction. Eureka! Philosophy shall be a consolation for my son! Thus, I whisper:

“Don’t worry. It’s not real.”

He shoots me a look of withering scorn. “I know *that*, Mum.”

So much, then, for philosophy.<sup>3</sup> He is young, and new to fiction; he does not know the conventions of story-telling. Maybe this story ends with a terrifying princess flambé. From his perspective, what’s happening *now* is distressing. It’s possible, of course, that as a small child, his response signifies nothing more than his inexperience at regulating his emotions. But I think it offers a clue about something that’s long bothered me about the paradox of fiction.

Attempts to resolve the paradox of fictional emotions focus on the non-existence of fictional characters, and star as examples in the larger debate between judgment theories of the emotions and embodied cognition theories of the emotions.<sup>4</sup> Judgment theories typically resolve the paradox by rejecting 1.: if emotions require cognitive judgments of danger, then whatever our physiological

---

<sup>2</sup> See Levinson 1997 for an overview of philosophical solutions to the paradox of fictional emotions; Gendler and Kovakovich 2005 delineate a particularly clear formulation of the paradox.

<sup>3</sup> In one sense, the theories are right. Pointing out that the characters aren’t real does *interrupt* his engagement with the fiction. My son has replaced fear with scorn. But arguably all this shows is that interrupting someone’s engagement with fiction makes it hard to engage with fiction.

<sup>4</sup> For example, Walton 1978, 1990, and 1997; Currie 1990; Stecker 2011; Tullman 2012; Tullman and Buckwalter 2014; Matravers 2014. But see Cova and Teroni 2016 for some pushback on this.

responses to fiction are, they're not genuine emotions. Embodied cognition theories, by contrast, explain that emotions result from affective appraisals of our circumstances, bodily changes that register what is significant to our well-being, without any intervening cognitive judgment.<sup>5</sup>

Cognitive appraisals of a situation can modulate our responses to the circumstances, but they do so by creating secondary affective appraisals; in other words, there is no change in emotions without first registering the change in the body. On Robinson's view, becoming emotionally involved in a movie requires nothing more than instinctively responding to the visual stimuli; bodies just aren't all that sophisticated. Robinson quips,

It does not matter to my emotion systems (fear, sadness etc.) whether I am responding to the real, the merely imagined, the possible or the impossible. [...] This might be irrational in some sense, but it happens all the time.<sup>6</sup>

Robinson shouldn't concede irrationality, or at least not on the grounds that the characters do not exist. As I shall argue, we should reject premise 3 in the standard paradox; existence turns out not to matter all that much to whether our emotions are rational. Instead, whether our emotional responses toward fiction are rational depends on our compliance with what I'll call the 'norms of aesthetic engagement': norms that govern the range of licensed emotional responses to the artwork.

2

Arguably, the paradox of fictional emotions has always, at its core, been a question of aesthetic appraisal, despite the direction the

---

<sup>5</sup> Robinson 2010, pp. 72–73.

<sup>6</sup> *Ibid.*, p. 86.

philosophical literature has taken. Recall one of Radford and Weston's original examples, in which a man weeps irrationally over imagining the counterfactual, in which his sister suffers from infertility. They conclude that the man is irrational because he is a "sort of Walter Mitty, a man whose imagination is so powerful and vivid that, for a moment anyway, what he imagines seems real, that his tears are made intelligible, though of course not excusable".<sup>7</sup> The clear implication is that if it is irrational to weep over purely imagined states of affairs, then it is also irrational to respond emotionally to fictional characters and situations, which after all are imagined states of affairs, the production of which has been outsourced to an artist. So, the historical discussion has proceeded by assuming that fictional characters and situations are on a par with merely imagined fictional characters and situations, and that the problem is one of reconciling rationality with non-existence.

But if we look at the norms that govern our emotional reactions to fiction, we see that the question of existence fades into the background in favour of the work's licensing attitudes, behaviours, and beliefs. The question of the appropriateness of emotions directed toward a fiction is not fully answered by whether the attitudes are toward existing objects. One might exhibit the *wrong* emotions, such as if someone were to laugh uproariously at 'Schindler's List'. One might suffer from a *deficit* of emotions, such as when someone is too cynical to find the casual violence of a gritty film noir shocking. One might suffer from an *excess* of emotions, as my son does when watching anything with dramatic tension.

In these cases, whether the characters exist settles nothing. We would not respond to the person who laughs at 'Schindler's List' by saying, "Don't laugh! It's only a fictional retelling of Oscar Schindler's heroism.". Nor would we point to the non-existence of the gritty noir world as a reason to have *more* shock at the violence. And even if my son did relax upon being reassured that the characters for

---

<sup>7</sup> Radford 1975, p. 74.

whom he fears do not exist, he wouldn't be engaging with the fiction properly if he felt nothing. Small children are *supposed* to be shocked and concerned when they see Lightning McQueen crash, but not to such an extent that they cannot bear to continue the film.

Failing to react emotionally to fiction signifies not a wise person who refuses to be taken in by an illusion, but someone who isn't engaging the fiction *properly*.<sup>8</sup> Our emotional responses to fiction are governed by the norms of aesthetic engagement. Aesthetic normativity, broadly construed, concerns what is good and bad in beauty and art. Norms of aesthetic engagement govern the responsibilities appreciators of art incur when they engage works of art. Roughly, they say, "to experience the work correctly, appreciators should respond in this way". These norms include appropriate emotional responses, and two norms are relevant for our purposes here:

1. Being open to engaging with fiction.
2. Exhibiting the appropriate emotional response.

First, consumers of works should be open to allowing the work to influence them emotionally. This norm is relatively easy to follow. As Robinson observes, our initial emotional response happens automatically. Robinson writes that "emotion is a *process* at the core of which is a set of bodily responses activated by an *affective appraisal* that is 'instinctive' and automatic".<sup>9</sup> The emotional systems bequeathed to humans by evolution simply aren't all that complicated, so they can be triggered by images, sounds, imagination, contemplation, and even mistakes in judgment. In the case of works of art, emotional responses may be occasioned, as in the case of literary fiction, by vividly imagining the characters, or listening to music with great focus, or by paying attention to the detail of a fine painting. On the embodied cognition model, we experience emotions

---

<sup>8</sup> See also Robinson 2004b.

<sup>9</sup> Robinson 2010, p. 73.

in these cases only if our imagination, focus, and attention are registered physiologically.<sup>10</sup>

Second, consumers should exhibit the appropriate emotional response authorised by the fiction.<sup>11</sup> Robinson writes that “[t]his automatic appraisal gives way to *cognitive monitoring* of the situation, which reflects back on the instinctive appraisal and modifies expressive, motor, and autonomic activity accordingly, as well as actions and action tendencies”.<sup>12</sup> I suggest that cognitive monitoring of our emotional responses to fiction is informed by knowledge of aesthetic and artistic practices. Having the correct emotional and behavioural response requires at least tacit knowledge of the artistic and aesthetic conventions of the artwork.

For example, some artworks are more demanding than others, requiring greater knowledge of those consumers who would engage with them. A newcomer to opera might find some of the conventions risible rather than moving. Moreover, the genre of a work can characterise what emotional responses are appropriate when engaging a work. For example, familiarity with the conventions of horror films will lead the savvy appraiser to feel apprehensive when the beautiful young heroine leaves her quiet farmhouse on a still night to check out a noise she hears coming from the shed. The noise itself isn’t frightening, but the knowledge of how horror movies usually go induces tension in viewers. A satire is characterised by how the genre inverts attitudes that would be appropriate to the situation in the real world, so that to engage with a satire properly requires at least tacit knowledge that one is engaging with a satire.<sup>13</sup>

Appropriate engagement with ‘Die Hard’ requires allowing the character John McClane a certain amount of leeway as the hero of an action film, as heroes of action films are customarily permitted to

---

<sup>10</sup> Robinson 2004, p. 43.

<sup>11</sup> See Walton 1990.

<sup>12</sup> Robinson 2010, p. 73.

<sup>13</sup> Liao 2013, p. 274.

perform impossible feats of endurance without disrupting narrative engagement.<sup>14</sup>

Not every artwork will be successful in evoking the desired emotional responses in its consumers; in fact, we can often see the aesthetic and artistic conventions most clearly in cases where the work is not successful in evoking the correct emotional response. Consider the famous sequence in ‘Battleship Potemkin’ of the massacre on the Odessa steps. It is still a powerful scene, but one of its most iconic images, the orphaned baby in the carriage rolling helplessly down the stairs, has been copied by so many films that what was once no doubt a startling scene full of urgent pathos feels somewhat pedestrian. In other cases, we may recognise that a maudlin work is trying to make us feel sorrow, or that a comedy suffers from stale and flat jokes. In such cases, we recognise upon reflection that there is a norm that explains how we should feel, but one that we find impossible to fulfil.

Nor should we think of aesthetic affective engagement as a dramatic martinet, dictating the precise response of the audience. Rather, it is more like norms of social engagement. A norm that says that I should greet the barista politely before ordering my coffee, which admits of a wide range of social interactions: a friendly smile, a warm hello, a “good morning”, a joke about manners before caffeination, a conversation, or just a nod and a brief and efficient transaction. Similarly, an aesthetic norm that instructs us to respond emotionally to a fiction will allow for a broad but related range of responses.

I suggest that an appraiser’s emotional response to fiction is rational only if she satisfies the first two norms of aesthetic engagement. The physiological responses that we have once engaging with a fiction properly are nearly automatic, but whether they are *appropriate* requires a secondary cognitive judgment, which in the case of fiction derives from a knowledge of aesthetic norms. Here I

---

<sup>14</sup> Hazlett & Mag Uidhir 2011, p. 41.

propose that aesthetic norms parallel folk theories of the emotions. On Robinson's analysis, physiological responses are simple and automatic, but folk theories of the emotions are complex, culturally specific, and determine how the emotion is received. Cultural practices distinguish emotions by the context in which they are appropriate (e.g. anger vs. indignation). Some initial emotional responses are disregarded as inappropriate or inscrutable, so cognitive monitoring maps the initial affective appraisals onto a topography of what is salient or valued.

For example, *liget*, an emotion peculiar to the culture of the Ilongot in the Philippines, described as "a powerful energy running through the body" or the feeling of "want[ing] to take a man's head and throw it", can be explained only with regard to the circumstances of that particular tribe.<sup>15</sup> As initial emotional responses are coarse-grained physiological responses, a Westerner can experience the initial physiological analogue of *liget* but won't be embedded in a culture where describing the feeling as 'liget' makes sense. Nor will others in her culture be inclined to blame or excuse her behaviour based on her experience of *liget*, nor teach their children to expect feelings of *liget* to emerge after a profound personal loss. Similarly, for those who are not members of the relevant culture, the Japanese emotion of *amae*, a propensity to 'depend or presume upon another's love' or the Western emotion of *schadenfreude*, the feeling of taking joy in someone else's misfortune, simply will not be conceptualised in the same way.<sup>16</sup>

I suggest that just as the concept of *schadenfreude* rationalises the mixture of joy and fascination that a Westerner feels in response to the misfortune of a rival, the norms of aesthetic engagement rationalise the emotions we have in response to an artwork. Crying for no reason at all may be a sign of mental instability; but being moved at the fate of Anna Karenina is understood culturally as an appropriate

---

<sup>15</sup> Spiegel 2017.

<sup>16</sup> Robinson 2004, p. 39.

response to a great work of art. The initial physiological response happens as we engage and imagine the world depicted by Tolstoy, but the norms of aesthetic engagement help us identify this emotion as rational. Thus, whether the ostensible objects of one's emotions exist turns out to be a misleading question. Radford's wool-gathering man is irrational when he weeps because we do not have a common cultural practice of working ourselves into tears over isolated, imagined situations. Were he to share his imaginings as a fiction about a woman who dearly wanted children but couldn't have any, he might find that the other would applaud his weeping, and weep with him.

Moreover, thinking of the emotions as modulated by the norms of aesthetic engagement helps explain the very subtle character of the emotions we might feel towards fictional characters, in ways that are not easily explained if the only question we have to answer is whether the character exists or not. We might welcome the death of a character, even as we feel sad, knowing that the death will lead the hero to a satisfying revenge. We might tremble with dread at a happy scene, recognising from foreshadowing that disaster will soon befall the characters. My son's emotional overreaction to fictional peril can be tempered by reassuring him, "Don't worry. The good guys will win in the end."

So, when Robinson quips that feeling emotions toward fiction might be in some sense irrational, but that it happens all the time, she concedes too much even by her own lights. "It happens all the time" because our physiological systems are not terribly sensitive, initially, to the source of the emotional stimulus. If the emotion that results from an encounter with a stimulus is believed to be rational, the culture must have a way of conceptualising it.<sup>17</sup> It is plausible that some of those norms concern our interaction with works of art.

So, we may revise the third premise of the paradox of fiction: In order for us to have genuine and rational emotions towards a given character or situation, we must believe either that they exist, or our

---

<sup>17</sup> Note that the claim is not that emotions are relative to cultures, but that *folk theories* of the emotions are relative to cultures.

reflective emotional response must be sanctioned by the norms of aesthetic engagement.

Note that I'm proposing, in accordance with the traditional formulation of the paradox, a necessary condition, not a sufficient condition. It is possible that a reflective emotional response could be sanctioned by aesthetic practices and nevertheless be irrational for other reasons.<sup>18</sup> A subculture that treated 'Schindler's List' as a farce would be irrational even if they had a well-developed aesthetic practice, because their practice would be monstrous!

More worryingly, we might object that aesthetic engagement is not necessary, because novices to artwork can exhibit appropriate emotions without any knowledge of the norms of aesthetic engagement. Imagine a person recently escaped from a hyper-Platonic enclave, in which all fiction had been banned. Upon her escape, the said person encounters a copy of 'Anna Karenina', reads it, and is profoundly moved although she has no theory of fiction or knowledge of aesthetic norms. But surely, she isn't irrational to respond with sorrow and grief; on the contrary, her perfectly appropriate reaction might be thought to demonstrate the timelessness and universality of the artwork.<sup>19</sup>

First, we should distinguish between the initial emotional response and the emotional response that develops as the result of cognitive monitoring, which is where knowledge of the norms of aesthetic engagement refine and conceptualise our emotional response. Our Platonic escapee will experience the story of 'Anna Karenina' as sad, on the assumption that despite her Platonic training, she is otherwise a typically functioning human being. In this restricted sense, her emotional response is rational.

One might object further, however, that my proposal implies that *knowledge* of aesthetic norms is necessary for genuine rational emotions. If our initial affective appraisal is automatic, and we refine

---

<sup>18</sup> Thanks to an anonymous reviewer for encouraging me to clarify this point.

<sup>19</sup> Thanks again to the exceedingly helpful, anonymous reviewer for this objection.

it only with cognitive monitoring informed by the norms of aesthetic engagement, we must know what those norms are. Thus, we should conclude that whether the Platonic escapee is rational or not depends on her knowledge of aesthetic norms. Aesthetic norms vary, and by hypothesis, she is from a culture that has no norm that concerns fiction. Doesn't this imply that *rationality* is dangerously relativistic?

This conclusion can be resisted. In this specific case, I find that my intuitions regarding her rationality depend on how we flesh out what our escapee thinks she's doing when she reads 'Anna Karenina'. Does she believe herself (à la 'Galaxy Quest') to be reading a historical document of a doomed woman? Then it seems that she is rational, though mistaken, to the extent that it is rational to mourn the believed bad end of any stranger. Has she heard only that within this new culture she's encountered, they regularly read fictional tales to induce strong emotions, and that one may feel sad even if one knows it is no more than a story? Then it seems to me that what this example shows is that only a very minimal knowledge of aesthetic conventions is necessary in order to conceptualise and refine one's emotional response to fiction. At least with respect to the very basic emotions, this seems right; one does not need a lot of exposure to fiction to understand the basic norms of aesthetic engagement. Does she fail, because of her hyper-Platonic heritage, to cognitively monitor her initial affective appraisal because she has no assumptions at all about 'Anna Karenina'? Then it's very hard to say how she would be able to reflect on what she's read; if that's all irrationality means in her case, it does not seem to pose a problem for my proposal.

More broadly, it's plausible to suppose that the norms of aesthetic engagement do not develop in isolation from other moral, social, and ethical norms. If that's the case, then while we might find that the norms of aesthetic engagement vary from culture to culture, the variance is no greater than we might find in folk theories of the emotions or moral theories, and if we can avoid a pernicious commitment to relativism in these cases, it is reasonable to suppose

that we can similarly avoid a commitment to relativism with respect to aesthetic engagement.

### 3

I have sketched two features of aesthetic engagement that allow us to treat the paradox of fictional emotions as a puzzle of aesthetics. We experience emotions by being open to engaging with fiction, and once engaged, we experience physiological changes that arise automatically. Our knowledge of aesthetic conventions and norms can revise our emotional responses to a work of fiction. Consequently, only emotions that are sanctioned by the norms of aesthetic engagement are appropriate to have toward fictional characters and situations. On the assumption that the embodied cognition model is correct, we should vigorously deny that the existence or non-existence of the characters has anything to do with the rationality of our emotions when it comes to fiction.

Much more needs to be said about aesthetic normativity, of course, especially as it related to ethics and practical reasoning. What grounds an aesthetic norm? How do we learn aesthetic norms? What about cases where we invert aesthetic norms, as when we enjoy a work, but only ironically? But I hope here to have shifted the conversation about the paradox of fictional emotions. Philosophers have ably responded to the paradox by developing theories of the emotions, but at its heart, the paradox is one about how we respond to the arts. Aesthetics should take this paradox back.<sup>20</sup>

[marybethwillard@weber.edu](mailto:marybethwillard@weber.edu)

---

<sup>20</sup> Thanks to Thi Nguyen for helpful comments on an earlier draft of this paper.

**ABOUT THE AUTHOR** Mary Beth Willard is associate professor of philosophy at Weber State University. She earned her Ph.D. from Yale University in 2009, and works primarily on issues at the intersection of metaphysics and aesthetics. The topics of her recent publications range widely, from a criticism of simplicity as a criterion of theory choice in metaphysics to the ethical dimensions of street art. Her current project concerns art, public space, and cyberspace.

## References

Cova, Florian & Teroni, Fabrice. (2016). "Is the Paradox of Fiction Soluble in Psychology?". *Philosophical Psychology* 29(6), pp. 930–942.

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

Gendler, Tamar, & Kovakovich, Karson. (2006). "Genuine Rational Fictional Emotions". In: *Contemporary Debates in Aesthetics and the Philosophy of Art*. Matthew Kieran (Ed.). Oxford: Blackwell Publishing, pp. 241-253.

Hazlett, Allan, & Christy Mag Uidhir. (2011). "Unrealistic fictions". *American Philosophical Quarterly* 48(1), pp. 33–46.

Lamarque, Peter. (1981). "How can we fear and pity fictions?". *British Journal of Aesthetics* 21(4), pp. 291–304.

Liao, Shen-yi. (2013). "Moral Persuasion and the Diversity of Fictions". *Pacific Philosophical Quarterly* 94(3), pp. 269–289.

Matravers, Derek. (2014). "Fiction and Narrative". New York: Oxford University Press.

Prinz, Jesse. (2004). "Which Emotions are Basic?". In: *Emotion, Evolution, and Rationality*, Dylan Evans & Pierre Cruses (Eds.). New York: Oxford University Press, pp. 69-88.

Radford, Colin. (1975). "How can we be Moved by the Fate of Anna Karenina?". *Proceedings of the Aristotelian Society, Supplementary Volumes* 49, pp. 67–93.

Radford, Colin. (1977). "Tears and Fiction." *Philosophy* 52(200), pp. 208–213.

Robinson, Jenefer. (1995). "Startle". *Journal of Philosophy* 92(2), pp. 53–74.

—. (2004a). "Emotion: Biological Fact or Social Construction". In: *Thinking About Feeling: Contemporary Philosophers on Emotions*, Robert C. Solomon (Ed.). New York: Oxford University Press.

—. (2004b). "The Art of Distancing: How Formal Devices Manage our Emotional Responses to Literature". *Journal of Aesthetics and Art Criticism* 62(2), pp. 153–162.

—. (2010). "Emotion and the Understanding of Narrative". In: *A Companion to the Philosophy of Literature*. Garry L. Hagberg & Walter Jost (Eds.). Malden: Blackwell, pp. 71-92.

Spiegel, Alix. "Invisibilia: A Man Finds an Explosive Emotion Locked in a Word". NPR, June 1, 2017. <https://www.npr.org/sections/health-shots/2017/06/01/529876861/an-anthropologist-discovers-the-terrible-emotion-locked-in-a-word>. Retrieved 11/15/2017.

Stecker, Robert. (2011). "Should We Still Care about the Paradox of Fiction?". *British Journal of Aesthetics* 51(3), pp. 295–308.

Tullman, Katherine. (2012). "HOT Emotions: Dissolving the Paradox of Fiction". *Contemporary Aesthetics*. [https://digitalcommons.risd.edu/liberalarts\\_contempaesthetics/vol10/iss1/9/](https://digitalcommons.risd.edu/liberalarts_contempaesthetics/vol10/iss1/9/). Retrieved 5/7/2018.

Tullmann, Katherine & Buckwalter, Wesley. (2014). "Does the Paradox of Fiction Exist?". *Erkenntnis* 79(4), pp. 779–796.

Van Leeuwen, Neil. (2016). "The Imaginative Agent". In: *Knowledge through Imagination*, Amy Kind & Peter Kung (Eds.). New York: Oxford University Press, pp. 85–109.

Walton, Kendall L. (1978). "Fearing fictions". *Journal of Philosophy* 75(1), pp. 5–27.

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

—. (1997). "Spelunking, Simulation, and Slime: On Being Moved by Fiction". In: *Emotion and the Arts*, Mette Hjort & Sue Laver (Eds.). New York: Oxford University Press, pp. 37–49.