**The Paradox of Fiction: A Proposal for a Solution Based on the Information-Processing Approach**

Sam S. Rakover

Department of Psychology, Haifa University, Haifa, Israel 31905

Running head: The paradox of fiction

Rakover phone number: 972 4 8240924

Email: [rakover@psy.haifa.ac.il](mailto:rakover@psy.haifa.ac.il)

Correspondence should be addressed to Sam S. Rakover, Department of Psychology, Haifa University, Haifa, Israel 31905. (Email: [rakover@psy.haifa.ac.il](mailto:rakover@psy.haifa.ac.il))

**Abstract**

The paradox of fiction deals with the following question: how is it possible to react emotionally to a fictive image? After a discussion of two important solutions to the paradox, I present an outline of my solution. The “Real/Fictive Information-Processing” theory proposes that all kinds of stimuli (real or fictive) are undergoing information processing by the cognitive system. Each stimulus consists of bundle of a particular stimulus (e.g., a cat) and certain indicators that specify whether it is real or fictitious. Based on the result of this processing, one’s emotional response is done adequately, it fits reality.

**The Paradox of Fiction: A Proposal for a Solution Based on the Information-Processing Approach**

**Introduction**

How is it possible for the reader of Tolstoy’s novel *Anna Karenina* to feel sympathy and pity for a woman who never really existed? Anna is a fictional character, not a real person. How is it possible for a man to fall in love with a beautiful movie star like Marilyn Monroe or Elizabeth Taylor long after her death? Similarly, it seems nonsensical that a young person would identify with fictional characters like Tarzan, Superman, Sherlock Holmes and Hercule Poirot. For this reason, scholars who have explored this topic assert that people exposed to artistic creations, such as novels, theatrical productions and movies, who emotionally respond to the fictional characters, as well as the fictional situations, with identification, fear, anger, affection, pity, loathing, and hatred, are responding irrationally. In other words, scholars assert that emotional responses to characters featured in artistic creations are irrational (summary and discussion, Kroon & Voloini, 2019; Schneider, 2020). Since the days of Aristotle, philosophers have been addressing this problem in some way or other. Yet fruitful inquiry into this question really began when Radford (1975) tackled it in a modern way.

Radford presents the problem discussed above through the aid of three premises, each of which seems intuitively correct. Yet, when taken together, there appear to be inconsistencies between them:

1. **Belief in Existence:** In order to respond emotionally to an individual, we need to believe in his/her existence.
2. **Disbelief in Existence:** When we are introduced to fictional characters in an artistic work, belief in existence is not present.
3. **Fictional Influence:** Fictional characters lead us to respond emotionally.

If a growling grizzly bear entered your room, you would be overcome by fear; if you witnessed a quarrel between two people in which one stabbed the other with a knife, a desire to help the wounded individual and a fear that you be seriously wounded would make you horribly upset; if you were at a funeral and you heard a widow cry, you would be saddened and cry too. All of these simple examples support premise (A) **Belief in Existence**: real situations, such as the appearance of a predatory animal, a bloody feud, or a person’s funeral, naturally evoke emotional responses.

In almost every case when people read literature, or watch a theatrical work, a film or an opera, they know well that the characters they are encountering are not real people. Literature is not biography. Similarly, in the theater or in an opera, an actor who is murdered during the performance bows to the audience during the curtain call; after a film hero’s shocking on-screen death, one can see the actor who played him participating in a long television interview. These are just a few of the many potential examples that support premise (B) **Disbelief in Existence:** characters in artistic works are fictitious and not realistic. Yet, nonetheless, despite the fact that characters in artistic works are fictitious, we respond to them emotionally. That being the case, despite the fact that each of these premises seems to be true, it is clear that when the three of them are taken together a contradiction arises. It is impossible that if we accept premise (A) **Belief in Existence** and premise (B) **Disbelief in Existence** as true that we would accept that premise (C) **Fictional Influence** is true—that is just not reasonable. If we respond emotionally to figures that are genuine and real, it does not seem logical that we would respond emotionally to a character understood to be fictitious and nonexistent. For example, (A) David thinks that to pity somebody you need to believe that they exist; (B) David believes that Anna Karenina never existed; (C) David pities Anna Karenina. Statements (A) and (B) lead us to the following conclusion: *David does not pity Anna Karenina*. However, statement (C) leads us to the opposite conclusion: *David pities Anna Karenina*.

The literature that addresses the paradox of fiction is fraught with various proposals for solving this paradox and debates that raise doubts about each one of these proposals. In what follows, I will briefly summarize two famous proposed solutions for the paradox and I will point to their primary flaws. The first of these proposals questions premise (C) **Fictional Influence** and the second questions premise (A) **Belief in Existence**. Finally, drawing on an approach of cognitive psychology that postulates that human cognitive processes function in a manner akin to computational processes in computers, I will propose my own solution.

**Two Famous Proposals for Solution of the Paradox of Fiction**

Although premise (A) **Belief in Existence** and premise (C) **Fictional Influence** both refer to instances where people respond emotionally to figures – real figures in the former and fictive ones in the latter – many scholars point out that people’s response to a real person is in many respects quite different from their response to a fictional character. For example, it is reasonable to assume that if a growling grizzly bear would come into your room, your behavioral responses would be completely different from the behavioral responses that you would have when watching a film in which a growling grizzly bear appears on screen. The difference is not just the intensity of responses like pulse and blood pressure. The types of responses are different too: In a real situation, you would scream as loud as you could and try to run away, or you would faint from fear. In a movie theater, you would get somewhat excited and your pulse and blood pressure would elevate slightly, but you would not let out a frightening scream and flee the theater or faint in your seat. On the contrary, you would probably continue eating popcorn and sipping the soft drink that you purchased at the kiosk.

Based on the differences between responses to actual figures and fictional ones, Walton (1978, 1990) proposed the following theory. The Make-Believe Theory proposes a solution to the paradox of fiction through denial of premise (C) **Fictional Influence**. The basic idea is that when we engage with aesthetic works (reading of literature, or viewing of theatrical, operatic, or cinematic creations), we enter into a state of play, pretend that we are present in a real situation, and respond to the fictional reality with seemingly real responses. Walton compares these fictional realities to games that children play with toys. In the world of children’s games, toys turn into living beings and children respond to them accordingly. A game in which a father plays a monster chasing after his young child constitutes an additional example. While the child screams in fear, he also enjoys it, because he is confident that nothing bad will happen to him. He understands full well that his father is not a monster and is just pretending. In other words, the small child does not really fear the monster. Rather, he is playing a game of pretend. In these examples Walton rejects premise (C) **Fictional Influence** and changes it: people do not respond to fictional characters in the same way that they respond to real figures. Instead, they start playing an imaginative game and they respond with imaginative responses, suited to the game, that constitute quasi-emotions.

One of the many critiques of Walton’s theory rejects his claim that responses to fictional characters are only quasi-responses or pretend responses. The critique points out that in many cases fictional characters and fictional situations provoke real emotional responses in the reader or the viewer, just like real people and situations do. For example, strong and sudden changes in stimuli introduced by either a theatrical performance or a film can provoke real startle responses or fears in the observer, because, as a matter of fact, every strong and sudden stimulus (a flash of light, a sharp increase in sound volume) arouse such responses. Furthermore, due to simple forms of learning like ‘fear conditioning,’ simple neutral stimuli are liable to arouse fear responses. For example, stimuli like light or sound that were previously linked with pain, will subsequently arouse fear responses in an individual. For that matter, stimuli that are similar to these stimuli will also arouse fear. A soldier who participated in gory battles during a prolonged and difficult war will likely respond with fear and panic when he subsequently hears sounds similar to those he heard during battle, such as strong sudden noises, sirens, or the sounds of airplanes overhead. There are works of art (literature, theatrical and operatic productions, and film) that from their inception are constructed to arouse strong emotional responses like mercy, sorrow, and empathy to the point of tears. (Films of this variety are referred to as a tearjerkers, since nobody can avoid shedding tears of pity and sorrow at the heroes’ and heroines’ bitter fates.) Who does not feel devastated, while reading *Bambi*, *a Life in the Woods* or watching its Disney adaptation, when the fawn sees its mother killed by hunters? And whose eyes do not mist up when hearing the wonderful arias of Verdi or Puccini? Even mafia dons shed tears. A final example: a friend told me that when he went to see a 3D film for the first time, a spear was thrown in his direction during one of the scenes and he felt a sharp and real pain in the middle of his forehead—the point towards which the spear had thrown. Based on these and other examples (like depictions of love and sexual intercourse), it is hard to say that the individual pretends and is just going through the motions of participating in a game.

The solution that Lamarque (1981) proposes is built primarily around a rejection of premise (A) **Belief in Existence**, in other words, an unacceptance that the belief in a character’s actual existence is necessary for one to react emotionally. Just as the cognitive system’s representation of a real figure can evoke emotional responses, the content of the thoughts representing a fictional character (Anna Karenina, for example) can also evoke emotional responses. For this reason, this theory is referred to as Thought-Theory. It has also aroused a number of objections. I feel that the following objection is the most interesting one. If Thought-Theory proposes that thought in the mind (the thought of Anna Karenina committing suicide, for example) generates the emotional response, the fundamental problem of the paradox of fiction remains unresolved, because one still responds to a fiction rather than an actual figure! In other words, the claim that the fictional character is represented as thought does not convert it into a real tangible individual. It remains fictive. Therefore, the question at the heart of the paradox of fiction remains unresolved: How does it come about that we respond emotionally to a nonexistent stimulus?

**Outline for a Theory of “Real/Fictive Information Processing”**

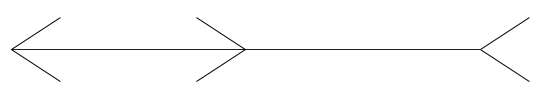
As just mentioned, the different attempts to solve the challenge posed by the paradox of fiction have spurred debates on the topic (see the surveys found in Kroon & Voloini, 2019; Schneider, 2020). To a certain degree, every proposed solution was met by serious critique. In the end, 36 years after Radford (1975) published his groundbreaking article, Stecker (2011) asked whether, after all the philosophic-literary tumult that followed it, there was really a need to address this paradox. Why? If most researchers do not accept premise (A) **Belief in Existence**, we can set aside this paradox, because, as a matter of fact, the paradox ceases to exist once this premise is rejected. Nevertheless, Stecker asserted that this paradox deserves continued discussion, because further research on it might help discover the mechanism through which the complex connection between emotional responses and the conditions responsible for their appearance, including fictions, is created.

I believe that Stecker has a good point. Indeed, in my proposal, I will follow his line of reasoning about the paradox of fiction. Given the cognitive psychology approach, I will propose an outline for solution of the paradox of fiction grounded in the belief that the human brain behaves in a manner comparable to the operation of a computer processing information. Therefore, I refer to this theory as the “Real/Fictional Information-Processing Theory.” Yet, before I sketch out this theory, I have to discuss two general observations that came to me while I read through the engaging material that has been written on this topic. To a certain extent, these observations serve as the foundation for the theory being proposed here.

1. **Rationality:** My impression is that the paradox is based upon the premise that human behavior, at its core, is rational. Why? Because without this premise, there is no paradox. Nobody would be surprised that the two premises (A) For David to feel pity for X, he needs to believe that X exists; and (B) David does not believe that Anna Karenina ever existed; do not contradict premise (C) David pities Anna Karenina. Rationality is what requires us to see a contradiction between premises (A) and (B) and premise (C). Yet, importantly, we know full well that human behavior frequently does not satisfy rational criteria and that common sense alone does not solve behavioral problems. If human behavior was grounded exclusively in logic, all that we would need to do to treat abnormal behavior would be to rationally explain what was wrong with such behavior and it would no longer be a problem. Yet, as stated, in a high number of cases human behavior is really not rational.
2. **The Extent of the Paradox of Fiction**: Rationality greatly expands the field of inquiry concerning the paradox under discussion. We are not only surprised that we pity a fictional character, but we also need to pay attention to the fact that we leave ourselves exposed to tragic and terrifying fictional situations. It is clear that no sane person would expose himself to a hungry grizzly bear, place his hand in a snake’s lair, ask to be fired from his job, or find pleasure when he or his family members are diagnosed with horrible diseases. A rational person would try to avoid these types of situations. Yet, nonetheless, we constantly expose ourselves to works of art that depict such horrible tragedies. It seems to me that these situations are no less paradoxical than the paradox of fiction. That being the case, I would proffer that the paradox of fiction is merely part of a much broader problem that I will refer to as “the problem of fiction.”

The outline of the Real/Fictional Information-Processing Theory will attempt to address the problem of fiction and it will take into account a number of ideas from the two proposed solutions to the paradox of fiction described above. I will begin by attempting to address the limitation of the Thought Theory – the claim that it does not respond to the following question: How is it possible that we respond emotionally to thoughts about fictional characters, when thoughts about the fictional character Anna Karenina, for example, are themselves nothing more than fictions?

The present theory is grounded in the premise, based upon numerous experimental observations, that individual response is produced after the cognitive system processes the stimuli absorbed by the perceptual system. The out-put of this processing will be referred to as “cognitive result.” Visual perception research points out that perceptual illusions testify to how individuals respond to stimuli after they have been processed, in other words to their cognitive result, because the information processing system is liable to fail in many cases and to produce a distorted perception of reality. To illustrate this, let us look at the famous Müller-Lyer illusion. As it is possible to see, the length of the horizontal line on the right side of the sketch appears longer than the length of the line on the left side, even though the length of both lines is equal! (Please take a ruler and measure them.) That being the case, this cognitive result does not accord with reality. If perception directly reflected the physical stimuli itself, it would be very difficult to explain this illusion. In order to explain it, one must assume that the cognitive information-processing mechanism is liable to err in certain situations.



The Real/Fictional Information-Processing Theory accepts the fundamental idea behind the Thought-Theory of Lamarque, according to which the individual’s emotional reaction is a response to the mental representation that appears in individual mind. It just expands upon the idea to present it in a more thorough manner. The present theory proposes that information processing almost always takes place at high speed without the individual even being aware of it. Furthermore, the theory proposes that all stimuli, whether they are real or fictive, go through the appropriate information processing that concludes with a cognitive result to which the individual responds (Here I should note that I assume that such a mechanism for information processing exists and functions without trying to describe it.) The foundation of the present theory rests on two fundamental questions:

1. **Compatibility with Reality:** How does the cognitive result accord with reality in most cases?
2. **Distinguishing Real and Fictive Responses:** How is the distinction between responses to cognitive results of real stimuli and responses to cognitive results of fictive stimuli created? In other words, how can the difference between real and fictive responses be explained?

The answer that I will propose here responds to both of the questions simultaneously. The basic idea is that the topic of the present discussion is multidimensional: The stimulus is multidimensional and the response is too. When we are exposed to a real figure like a cat, for instance, one can propose that the pattern of the cat’s visual stimuli (the cat’s shape and size) is connected to a collection of stimuli, which I will refer to as “indicators,” that signal that a real figure, rather than a fictional one, is present before us. For example, the cat’s visual form is three-dimensional, it moves in a specific way, it makes sounds, it has a specific smell, and we are able to feel it with our hands. The connection between the visuality of the cat (or any other creature) and the indicators that signal that it is a real creature is something that we learn through daily experience from birth.

With the aid of two examples, I will concisely demonstrate the importance of learning about reality. The first example is connected to my friend whose sudden feeling of sharp pain in his forehead from a fictive spear thrown in a 3D film was discussed above. This was the first and last time that my friend felt such pain. He very quickly learned that events that take place in a 3D film are just as fictional as those that occur in a regular film. The second example is connected to what is referred to as the perception of ‘size constancy’. Through the learning process that begins at a very young age, people learn that the size of a creature moving towards or away from them remains constant despite the fact that changes in distance change the size of the creature’s projection onto the retina: when one becomes farther away, the projection on the retina is small and vice versa.

By contrast, fictive stimuli lack these reality indicators. When I read about Anna Karenina, there is no real three-dimensional shape present before me to observe, no real movement through space, no voice to hear, and I do not have a real interaction, like a conversation in a café, with her. I only read words that successfully create a specific impression of Anna Karenina in my mind. It is not just that the fictive character lacks reality indicators. The act of reading is accompanied by indicators that point out that one is clearly reading a fictional novel written by an author named Tolstoy; one knows that nobody considers it a biography written about a real figure whose life is depicted by an historian who undertook careful historical research about somebody who lived in Russia named Anna Karenina.

Thus, according to the Real/Fictional Information-Processing Theory, one can differentiate between two types of stimuli that are represented in the cognitive system, stimuli to whose cognitive result the individual responds in different ways: One type of stimuli is a “bundle of real stimuli,” in other words, a figure and reality indicators, and a second type is a “bundle of fictive stimuli,” that is a figure and fictive indicators. These stimuli bundles have indicators that signal reality or fictitiousness. Hence, it turns out that the way one responds to a real grizzly bear will be different from the way that one responds to a fictitious bear (in a story or a film). One can further propose that the emotional reactions to certain bundle of real stimuli (such as fear, anger, a feeling of being threatened, joy, pleasure, and nausea) are mostly universal reactions based on the evolutionary process (for an overview and discussion, see Barret, 2006; Ekman, 1992, 2016; Rakover & Cahlon, 2001). In contrast, the emotional responses to a bundle of fictive stimuli are weakened or inhibited by the lack of reality indicators and the presence of fictive indicators. (It is possible to suggest that this is where the basis for Walton’s make-believe theory lies.) In view of this, the present theory explains with great simplicity why we respond to a fictional literary character (the response is always to a cognitive result, whether the bundle of stimuli is real or fictive) and why the response to the bundle of fictive stimuli is different from the response to the bundle of real stimuli (the real cognitive result includes real indicators while the fictive cognitive result contains indicators of fictive stimuli).

The present theory also addresses the problem of fiction, according to which, as mentioned above, it is unclear from a rational point of view why we expose ourselves to artistic works that depict horrible tragedies. The answer is connected to the fact that the fictive bundle of stimuli makes this possible: Since it is clear to us that we will not be harmed at all, we can enjoy the failures and misfortunes that befall another, especially since we know that the harm that befalls the protagonists of literature, theater and film is mere fiction. Nevertheless, I need to add here that in certain cases fictive stimuli that resemble real stimuli are liable to arouse responses in individuals that are similar to responses to real stimuli, as in the example of fear-response or the example of my friend the first time he went to see a 3D film and a fictive spear was thrown in his direction (see the section on the critique of Walton above).

As a conclusion to the present discussion, I will address the extent to which the Real/Fictional Information-Processing Theory night be able to help solve general literary theoretical questions about the difference between fiction and non-fiction, the role of author and reader’s imaginations in the reading process and similar questions aroused by works of art (see the discussion in Kroon & Voltolini, 2019). The framework of the current article does not allow for a full response to these questions. All that I can say is the following: I believe that if discussion of the two basic concepts, “bundle of real stimuli” and “bundle of fictive stimuli,” underlying the present cognitive-psychological approach is expanded, this theory will offer interesting answers to theoretical questions that arise in discussion of artistic works. Let us take the following imaginative possibility as an example. In an archaeological dig, researchers find very ancient scrolls that tell about a highly powerful king who built a huge city at a specific location, but, when his great haughtiness led him to try to conquer the world, his sons poisoned him. Do the scrolls narrate actual events or are they telling a fictional tale? The current proposal would be to search for more indicators. If, for example, the scrolls tell a story containing deeds that were beyond human capabilities at the time of the scrolls’ composition, we would tend to say that the story is fictional. Yet, if archaeological digs that take place in the location specified in the scrolls reveal strong evidence of a huge and ancient city that existed there, the pendulum would swing towards an historic description.

**Bibliography**

Barrett, L. F. (2006). Are Emotions Natural Kinds? *Perspectives on Psychological*

*Science*, 1, 28-58.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, 6,

169-200.

Ekman, P. (2016). What Scientists Study Emotion Agree About? *Perspectives on*

*Psychological Science*, 11, 31-34.

Kroon, F. & Voltolini, A. (2019). Fiction. In E. N. Zalta (Ed.). *The Stanford*

*Encyclopedia of Philosophy.*

Radford, C. (1975). How Can We Be Moved by the Fate of Anna Karenina?

*Proceedings of the Aristotelian Society*, *Supplementary*, 49, Pp. 67-80.

Rakover, S. S. & Cahlon, B. (2001). *Face Recognition: Cognitive and*

*Computational Processes*. Amsterdam: John Benjamins Publishing Company.

Schneider, S. (2020). The Paradox of Fiction. In J. Fieser & B. Dowden (Eds.)

*Internet Encyclopedia of Philosophy.*

Stecker, R. (2011). Should We Still Care about the Paradox of Fiction?

*The British Journal of Aesthetics*, 51, 295-308.

Walton, K. L. (1978). Fearing Fictions. *The Journal of Philosophy*, 75, 5-27.

Walton, K. L. (1990). *Mimesis as Make-believe: On the Foundations of the*

*Representational Arts.* Cambridge, MA.: Harvard University Press.