David Grčki
dgrcki@uniri.hr
Faculty of the Humanities and Social Sciences in Rijeka
https://orcid.org/0000-0001-9068-5995

FROM BAD TO WORSE: WHY THE CONCEPT OF 'ANTIHERO' FAILS

Abstract: In this paper, I argue that the concept of antihero fails to adequately explain how and why we engage with fictional characters. In the first part, *The Structure of Sympathy*, I present the most prominent epistemic theory of character engagement – Murray Smith's theory of sympathy, which contains recognition, alignment, and allegiance. In the second part of the paper, *The Problem of 'Antihero'*, I explain the problem of 'antihero' as is presented in contemporary literature. I argue that the 'antihero' as a concept suffers from two problems: the unclear moral criteria and the unity of concept. I contend that these problems are unsurmountable and that we should reject the concept of 'antihero' when explaining character engagement. I conclude by arguing that the framework of tragedy and the rejection of villains explain how we engage with morally bad characters.
Keywords: fictional characters, character engagement, antihero, tragedy

The Structure of Sympathy

There are several ways to approach the concept of fictional characters. The first is from the point of view of metaphysics; what is the *nature* of fictional characters; do they exist and if they do; how do they exist (Thomasson, 1999; Bonomi, 2008; Motoarca, 2014)? The second is from the perspective of literary science, which encompasses literary criticism, comparative literature, and literary theory; what narrative function do fictional characters play in the story; is a character a protagonist, an antagonist, a 'side character', or a romantic interest (Tomashevsky, 1925; Greimas 1966; Barthes, 1981; Black, 2018)? The third is from the point of view of epistemology; how do we engage with fictional characters cognitively and emotionally while treating them as real people (Smith, 1995; Plantinga, 2010; Carroll, 2013; Felski et al., 2019; Rain & Mar, 2021)? Treating them as real people simply means attributing to them psychological and mental states, and ethical and moral characteristics that are comparably similar to our own (Grčki, 2023). My argument against the concept of 'antihero' is embedded in the third approach to the concept of fictional characters – epistemology.

The most prominent and influential epistemology of fictional characters is presented by Murray Smith in his book *Engaging Characters: Fiction, Emotion and the Cinema* (1995). Smith presents the epistemology of fictional characters in the following way.

> What is the nature (the logic, the rationality) of our responses to fictional characters? It what sense, if any, do we take characters in films to be real persons? Is such a (mis)perception a prerequisite for having an emotional response to a fictional character? If filmic characters aren't real people, what are they? What exactly do we mean when we say that we 'identify' with a particular character? What are the various senses of this term, and can they be developed into a systematic explanation of emotional response to fictional characters in cinema? (Smith, 1995: 2)

Firstly, by framing the discussion of fictional characters through audience's 'responses', 'emotional responses,' and 'identification', Smith is putting focus on audience's understanding and acknowledgment of fictional characters as real people. The field of philosophy that encompasses, broadly speaking, the concepts of 'knowledge' and 'understanding' is epistemology, and for that reason Smith is presenting the epistemology of fictional characters. Secondly, although Smith talks about 'filmic characters', and 'fictional characters in cinema', his work has been broadened to other narrative arts such as literature, comic books, and videogames (Carroll, 2004, 2013; Lankoski, 2011; Felski et al., 2019; Tobón, 2019; Brodie & Ingram, 2021). Third, and the most important, is Smith's 'systematic explanation of emotional response to fictional characters in cinema' which he terms „*the structure of sympathy* with distinct *levels of engagements"* (Smith, 1995: 5).

The structure of sympathy consists of three distinct levels of engagement – *recognition*, *alignment*, and *allegiance*. Generally, recognition is the way that spectators construct characters, alignment is the narrative or cinematic perspective that spectators have on the character, and allegiance is the spectators' moral evaluations of the characters. He states that "A schema is a 'mental set' or a conceptual framework which enables us to interpret experience, form expectations, and guide our intention" (Smith, 1995: 21). In this light, Smith's schema is a useful theoretical tool and a building block for an epistemology of fictional characters. Regarding alignment, Smith insists that this term is neither a 'identification' nor a 'point of view'. Instead, it is "… the process by which spectators are placed in relation to characters in terms of access to their actions, and to what they know and feel" (Smith, 1995: 83). Smith continues that alignment is akin to the literary notion of 'focalization' because it explains how much information is given to the readers/spectators in any given point in the story. Alignment is a sum of two interlocution functions*: spatio-temporal attachment* and *subjective access*. Spatio-temporal attachment is the way that the narration restricts itself to the actions of a single character or moves freely between the characters regardless of time and space. Characters may be on different continents (*Game of Thrones* 2011-2019) or in

different centuries (*Dark* 2017-2020) and still be a part of the same story. Subjective access simply refers to the degree of access spectators have to any particular character. These can be achieved in a variety of ways, such as cinematic voice-over narration, like in David Fincher's movie *Fight Club* (1999).

The last part of the structure of sympathy and the third level of engagement is allegiance. Allegiance refers to "… the moral evaluation of the characters by the spectator" (Smith, 1995: 84). This evaluation comes in two dimensions – affective and cognitive. In the affective dimension, we can be angry or upset at the protagonist when she does something that goes contrary to our moral norms or sentiments. For example, we can be angry at Iago for lying/manipulating[1] Othello (Shakespeare, 1992) and we can be filled with joy at the end of *A Christmas Carol* when the transformation of the protagonist Ebenezer Scrooge from a cruel and selfish man to a loving and caring uncle is complete (Dickens, 1943). In the cognitive dimension, we assess the moral attitudes and actions of a character. If those attitudes align with ours we sympathize with the character, and if those attitudes do not align with ours we disapprove of the character i.e., we have an antipathetic judgment towards the character. This explains why we sympathize with Shakespeare's Othello and not with Iago, and why we sympathize with Ebenezer Scrooge at the end of *A Christmas Carol* and not at the beginning.

These three levels of engagement form the structure of sympathy, the most influential contemporary epistemic account of how we engage with and understand fictional characters. Over the last decade the problem with Smith's theory has arisen, which can be formulated in the following manner; how do we epistemically engage with morally bad characters, and do we really have antipathetic judgments towards them? The most prominent case in contemporary literature of a morally bad character that we sympathize with is Tony Soprano from *The Sopranos* (1999-2007). He is, unquestionably, a ruthless criminal, adulterer, and murderer. But still, despite these facts, he garners sympathy from most of the audience. For some reason, we are on the side of a New Jersey crime boss who does clearly immoral things. This phenomenon is called the problem of the 'anti-hero' or the 'rough hero'.

The Problem of 'Antihero'

Antihero is a fictional character whose set of moral beliefs differs from or is contrary to a standard set of moral beliefs held by the majority of the audience (Grčki, 2023). Carroll formulates the problem of 'antiheroes' in the following way: "The problem is basically how a viewer can be sympathetic (care for, or have a pro-attitude) toward a fictional character whose real-world counterpart she would abhor totally? (…) Does this make any sense? How is it possible?" (Carroll, 2013: 235), and Smith says "How can we care about, or sympathize with, someone who would repel us in

---

[1] For an interesting discussion on the topic of whether Iago is lying to Othello or just manipulating him, see Blečić, 2019.

reality?" (Smith, 2011: 69). Both Carroll (2013) and Smith (2011) use the example of Tony Soprano to demonstrate the problem of being an 'antihero'.

There have been numerous attempts to solve the problem of 'antiheroes'. The first strategy is to argue that 'antiheroes' are not as morally bad as they seem. Carroll (2004) argues that morally bad characters are not actually morally bad, all things considered. They are bad in an absolute sense, but they are good in a relative sense, i.e., they are better relative to the other characters in the fictional world that they inhabit. For example, Tony Soprano is morally bad in an absolute sense, but he is comparatively better than the other corrupt police officers, criminals, and murderers in the fictional world of *The Sopranos*.

The problem for Carroll (2004) is that most morally bad characters are not good even in a 'relative to the other fictional characters' sense. The most obvious counterexamples are Alex DeLarge and Patrick Bateman, who are undoubtedly the evilest characters in their respective fictional world and possess no redeeming qualities (Grčki, 2023). But even if we take characters from serialized television, such as Walter White and Saul Goodman, Carroll's argument falls short in explaining how we engage with those characters. Walter White is a chemistry teacher at a high school, diagnosed with stage 3 terminal lung cancer, who, over the course of *Breaking Bad* becomes a ruthless leader of a drug empire he created. We sympathize with him at the beginning of the series because he is suffering from a terminal illness and is looked down upon by his family, but we despise him at the end because he is transformed into a narcissistic drug lord who cares only for himself.

The concept of 'antihero' suffers from two problems: the unclear moral criteria and the unity of concept. I claim that these problems are severe and that we should reject the concept of 'antihero' when building an epistemic theory for character engagement. The concept of 'antihero' lacks any clear moral criteria. Let us take a look at some of the definitions and descriptions of the 'antihero'. Carroll and Smith, respectively, offer no definitions of the concept of 'antihero'. Carroll says that an 'antihero' is "… a fictional character whose real-world counterpart she would abhor totally" (Carroll, 2013: 235), and Smith argues that it is "… someone who would repel us in reality" (Smith, 2011: 69). This description of an 'antihero' is quite vague, and of limited use when we discuss affective and cognitive engagement with fictional characters. Both Carroll and Smith point to Tony Soprano when describing a prototype 'antihero,' but based on their 'definitions,' Shakespeare's Macbeth and Homer's Agamemnon would be classified as 'antiheroes.'

Any moral or rational person would be 'abhorred' and 'repelled' by either of these characters. Bruun Vaage argues that the 'antihero' "… truly is immoral in the sense that he is continually violating moral principles" (Vaage, 2016: xi). This definition has two problems. First, it begs the question of what exactly the 'moral principles' are, which is still in dispute in moral philosophy. Second, even if we somehow agree on the moral principles, does that mean that all the antagonists, such as Iago and Shylock, are 'antiheroes'? García argues "… that the predominant traits of today's antihero is a mixture of hero and villain characterized by moral ambiguity; a certain Machiavellianism exists with regards to the achievement of certain ends, and a

contradiction between ideals (…) and actions" (García, 2016: 53-54). Mittell argues, in similar vein, that an 'antihero' is "… a character who is our primary point of ongoing narrative alignment, meaning that we closely follow their experiences and have some access to their knowledge or interior state, but whose behavior and beliefs provoke ambiguous, conflicted or negative moral allegiance" (Mittell, 2015: 75). Both García's and Mittell's definitions of 'antiheroes' are detailed, but vague and unhelpful. García implies that this notion is a recent occurrence brought about by film and television, but this is simply not the case. 'Moral ambiguity' and 'conflicted or negative moral allegiance', argued by Mittell, were not invented by film and television. We have the same engagement with Dostoyevsky's Raskolnikov from *Crime and Punishment* (1866), or with Homer's Achilles, and García and Mittell would probably not consider the latter two 'antiheroes'. Moral flaws and moral ambiguity are simply prerequisites for complex and multilayered characters. From this fact, it follows that our engagement and responses as an audience will provoke ambiguous, conflicted, or negative responses.

The last strategy to define 'antiheroes' is to argue that the concept of 'antiheroes' is reserved only for really morally horrendous characters. Eaton states that 'rough' heroes need to satisfy the following conditions:

> (a) flaws are, first, grievous: he is usually a sociopath, an outlaw, a murderer, a sex criminal, a sadist, or Satan incarnate
> (b) flaws are an integral part of his personality rather than peripheral failings or foibles
> (c) often fully intends to do bad and is remorseless about his crimes
> (d) the audience's forgiveness is not prescribed, nor are we offered reasons to dismiss his misdeeds as the result misfortune, weakness, folly, or ignorance
> (f) vices are not outweighed by some more redeeming virtues; although he is not entirely bereft of such virtues—more on this in a minute—they far from surpass his vices in gravity or importance. (Eaton, 2012: 284)

Eaton's idea is that the concept of 'antihero', or 'rough hero' as she calls it, should apply only to characters that have grievous flaws which are integral part of their personality, are remorseful, and are not outweighed by some more redeeming virtues. In other words, 'antiheroes' are villains, according to Eaton (2012). They are sociopaths, outlaws, murderers, sex criminals, sadists, or Satan incarnate. Her definition seems to be, *prima facie*, robust enough, until we look at her examples of what a 'rough hero' is. She offers more than twenty examples of 'rough heroes' in five subcategories: *The glorified criminal*, *The congenial murderer*, *The likeable sex criminal*, *The sympathetic sadist*, *The appealing mean-spirited person*. Her examples encompass literature, film, and serialized television. Michael Corleone (played by Al Pacino) from *The Godfather trilogy* (1972, 1974, 1990), Gus Fring (played by Giancarlo Esposito) from *Breaking Bad* (2008-2013), and the aforementioned Tony Soprano are placed in 'The glorified criminal' category, while Dr. Hannibal Lecter

and Dexter Morgan are placed in 'The congenial murderer' category, and Heathcliff from *Wuthering Heights* (1847) and The Underground Man from *Notes from Underground* (1864) appear in 'The appealing mean-spirited person' category. The problem with this is what I term the unity of concept problem.

Eaton's definition bundles together a specter of fictional characters that actually have little in common. Although they technically satisfy these conditions, we engage with them for completely different reasons and with utterly different emotions. What do Michael Corleone and Gus Fring have in common that is relevant for the explanation of our affective and cognitive engagement with them? I argue – very little. Michael Corleone is the youngest son of Vito Corleone (played by Marlon Brando) who, in the beginning, wants nothing to do with his family crime business, and later, after his father dies, becomes a ruthless don of the Corleone crime family. Gus Fring is a lieutenant of Vuente Cartel, who secretly plans the revenge on the said cartel for the death of his partner and lover Max Arciniega (played by James Martinez). We engage positively (if we do) with them both for completely different reasons. We sympathize with Michael because he is the protagonist and we want to see him achieve his goals and defeat his enemies who are worse or at least as bad as him. Gus is the main antagonist of *Breaking Bad*. We sympathize with him because he is in some sense, better than Walter, and reflects his hypocrisy and immoral behavior. In a similar manner, I could argue that Heathcliff and The Underground Man have very little in common and we engage with them for very different reasons.

Does this mean that there is no problem at all? Or is there still a question of how and why we sympathize with morally bad characters? There are two things that I would like to point out when discussing morally bad characters and our engagement with them: the rejection of villains and the framework of tragedy. These two ideas, in conjunction, explain the problem with morally bad characters. Firstly, we do not sympathize with characters that are clearly and unequivocally morally bad, such as Alex DeLarge, Patrick Bateman, or Cassie from *Promising Young Woman* (2020). We can ask a question: why do we engage with these works at all? One of the reason is because these works pose interesting and important moral questions. *A Clockwork Orange* (1971) tackles the question of an appropriate response to unapologetic delinquency and violence by people who simply enjoy it. *American Psycho* (2000) asks how much one needs to be psychopathic in order to effortlessly succeed in a late capitalist society. *Promising Young Woman* (2020) asks to what extent one has to go in order for a society to realize that systemic sexual violence against women exists and that it is morally wrong. Engaging with these works and understanding these themes does not mean sympathizing with characters through which these themes are cinematically delivered. I agree that the social response to delinquency and violence is a complex problem, but I do not sympathize with Alex when he, with his 'droogs', breaks into a disabled man's house and tortures him and his wife. The two can perfectly be separated. But what about characters that are not completely evil, but still morally bad, such as Tony Soprano, Walter White? Our engagement with these types of characters can be explained with the concept of tragedy. The framework for this is actually established by Smith himself.

Setting *The Sopranos* within the framework of tragedy is instructive in terms of understanding how it is that we may sympathize with Soprano. In his commentary on tragedy in the *Hamburg Dramaturgy*, G.E. Lessing argues that it is essential that the protagonist of a tragedy is 'one of ourselves'—that is, recognizable as a human being, a mix of virtue and vice, rather than 'an incarnate devil.'" Lessing's remarks are prompted by the character of Richard III in Christian Felix Weiss's play Richard III (completed in 1765), whom Lessing called "the most loathsome monster that ever trod the stage. In Lessing's view, this ensured that the play could not succeed as a tragedy. According to Lessing, the emotions of pity and fear, prompted by tragedy, are interrelated in a particular fashion: our pity for the tragic protagonist hinges on seeing him as precisely not monstrous, enabling us in turn to fear such a fate for ourselves. Soprano is rather more like the protagonist of a tragedy—"neither a wholly virtuous nor a wholly vicious man" —than he is like Weiss' monstrous Richard III.12 Soprano is sufficiently ordinary that we may, in Lessing's terms, recognize him as 'one of ourselves'. (Smith, 2011: 74)

Smith's idea is that what explains our engagement with 'antiheroes' is the framework of tragedy. Tony Soprano is, as a protagonist of most tragedies, neither a wholly virtuous nor a wholly vicious man. This allows us, as audience members, to sympathize with him because he is one of ourselves. The same can be said for Walter White and Saul Goodman, but not for Alex DeLarge, Patrick Bateman. The former operates in a context of tragedy, while the latter does not. I would build on Smith's idea and claim that the framework is not limited to the protagonist. Gus Fring is an antagonist but operates in the framework of tragedy, and this explains most of the sympathy we have for him. Jamie Lannister (played by Nikolaj Coster-Waldau) from *Game of Thrones* (2011-2019), a side character, operates in the framework of tragedy, and this explains most of the sympathy we have for him, despite the horrible things he has done. In any explanatory framework, there will be hard cases, in which we are not sure where they belong. One such example is Dexter Morgan. He is either a villain and we should reject engagement with him or our engagement with him is explained in a framework of tragedy. My point is simply that there is nothing problematic or mysterious with how we engage with morally bad characters and that the concept of 'antiheroes' offers little to no explanatory value.[2]

References

A Clockwork Orange (1971). Directed by Stanley Kubrick. US: Warner Bros. Pictures.

American Psycho (2000). Directed by Mary Harron. US: Lions Gate Films.

Barthes, R. (1981). *Introduction to the Structural Analysis of the Narrative*. Birmingham:

Centre for Contemporary Cultural Studies, University of Birmingham.

Black, N. (2018). *Virginia Woolf as a feminist*. Cornell University Press.

Blečić, M. (2019). Shakespeareov Jago kao protuprimjer tradicionalnoj definiciji laži,

Filozofska istraživanja 39 (04/156), 827-851.

Bonomi, A. (2008). Fictional Contexts. In Paolo Bouquet, Luciano Serafini, & Richmond H. Thomason (eds.), *Perspectives on Contexts*, Stanford, CA: CSLI Publications, 215–249.

Breaking Bad (2008-2013). Created by Vincent Gilligan. US: AMC.

Brodie, Z. P., & Ingram, J. (2021). The dark triad of personality and hero/villain status as predictors of parasocial relationships with comic book characters. *Psychology of Popular Media*, 10(2), 230–242. https://doi.org/10.1037/ppm0000323

Bronte, E. (2012) *Wuthering Heights*. London, England: Penguin Classics.

Carroll, N. (2004). Sympathy for the Devil. In Richard Greene; Peter Vernezze (Eds.),

The Sopranos and Philosophy. I Kill Therefore I Am (121-136). Illinois: Open Court.

Carroll, N. (2013). *Minerva's Night Out: Philosophy, Pop Culture, and Moving Pictures*

Oxford: Wiley Blackwell.

Dark (2017-2020). Created by Baran bo Odar & Jantje Friese. Germany: Wiedemann & Berg Television.

Dickens, C. (1943). *A Christmas Carol*. London: Chapman & Hall.

Dostoyevsky, F. (2006). *Crime and Punishment*. Translated by Sidney Monas, Signet Classic.

Dostoevsky, F. (2011). Notes from Underground. (Zinovieff K. & Hughes J. Trans.) London: Alma Classics Ltd. (1861)

Eaton, A.W. (2012). Robust Immoralism. *The Journal of Aesthetics and Art Criticism* 70: 281-2.

Felski, R., Anderson, A., Moi, T. (2019). *Character: Three Inquiries in Literary Studies*. Chicago: University of Chicago Press.

Fight Club (1999). Directed by David Fincher. US: 20th Century Fox.

Game of Thrones (2011–2019). Premiered on April 17, 2011, and concluded on May 19, 2019. US: HBO.

García, A. N. (2016). Moral Emotions, Antiheroes and the Limits of Allegiance, in *Emotions in Contemporary TV Series*, García, Alberto N. (eds), Palgrave McMillan, Basingstoke, 2016, pp. 52-70.

Grčki, D. (2023). Learning through Stories: Epistemic Understanding as a Cognitive Value of Narrative Arts. *The Journal of Aesthetic Education* 57(3), 49-68. https://www.muse.jhu.edu/article/905818.

Greimas, A. J. (1983). *Structural Semantics: An Attempt at a Method*. (McDowell, D. et al. Trans.) Lincoln, Nebraska: University of Nebraska Press. (1966).

The Godfather (1972). Directed by Francis Ford Coppola. US: Paramount Pictures.

The Godfather Part II (1974). Directed by Francis Ford Coppola. US: Paramount Pictures.

The Godfather Part III (1990). Directed by Francis Ford Coppola. US: Paramount Pictures.

Halliday, F. E. (1964). A Shakespeare Companion 1564–1964. Baltimore, Penguin.

Rain M, Mar RA. Adult attachment and engagement with fictional characters. J Soc Pers Relat. (2021). Sep;38(9):2792-2813. doi: 10.1177/02654075211018513. Epub 2021 Jun 9. PMID: 34504387; PMCID: PMC8419286.

Lankoski, P. (2011). Player Character Engagement in Computer Games. *Games and Culture*, 6(4), 291-311. https://doi.org/10.1177/1555412010391088

Mittell, J. (2015). Complex TV: The Poetics of Contemporary Television Storytelling. NYU Press. JSTOR, http://www.jstor.org/stable/j.ctt15r3zwk. Accessed 12 June 2024.

Motoarca, I. (2014). Fictional Surrogates, *Philosophia*, 42(4): 1033–1053. doi:10.1007/s11406-014-9522-1

Rain, M., & Mar, R. A. (2021). Adult attachment and engagement with fictional characters. *Journal of Social and Personal Relationships*, 38(9), 2792-2813. https://doi.org/10.1177/02654075211018513

Plantinga, C. (2018). *Screen Stories: Emotion and the Ethics of Engagement*. Oxford: Oxford University Press.

Promising Young Woman (2020). Directed by Emerald Fennell. US: FilmNation Entertainment & LuckyChap Entertainment.

Sartre, J.P. (1949). *Nausea*. (Baldick R. Trans.) London: Penguin Books. (1938)

Shakespeare, W. (1992). *Othello* (C. Watts, Ed.). Wordsworth Editions.

Smith, M. (1995). Engaging Characters: Fiction, Emotion, and the Cinema. Oxford: Oxford University Press.

Smith, M. (2011) Just What Is It That Makes Tony Soprano Such an Appealing, Attractive Murderer?. In Ward E. Jones, Samantha Vice (Eds.) *Ethics at the Cinema* (66-90). New York: Oxford University Press.

The Silence of The Lambs (1991). Directed by Jonathan Demme. US: Orion Pictures

The Sopranos (1999-2007). Created by David Chase. US: HBO.

Thomasson, Amie L. (1999). *Fiction and Metaphysics*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511527463

Tomashevsky, B. (1965). *Thematics*. (Trans. Lemon, L. T., Reis, M. J.) Lincoln: University of Nebraska Press. (1925)

Vaage, M. B. (2022) 'Should we be against empathy? Engagement with antiheroes in fiction and the theoretical implications for empathy's role in morality', in Mezzenzana, F. and Peluso, D. M. (eds) *Conversations on Empathy:*

Interdisciplinary Perspectives on Imagination and Radical Othering. Routledge, pp. 116-134. doi: 10.4324/9781003189978-9.

Vaage, M. B. (2021). 'Five theses on the difficulty of ending Quality TV series', in Nannicelli, T. and Pérez, H. J. (eds) *Cognition, Emotion, and Aesthetics in Contemporary Serial Television*. United Kingdom: Routledge. Available at: https://www.routledge.com/Cognition-Emotion-and-Aesthetics-in-Contemporary-Serial-Television/Nannicelli-Perez/p/book/9781032037158.

Vaage, M. B. (2015) 'On the Repulsive Rapist and the Difference between Morality in Fiction and Real Life', in Zunshine, L. (ed.) *The Oxford Handbook of Cognitive Literary Studies*. Oxford University Press. doi: 10.1093/oxfordhb/9780199978069.013.0021.