

# Narrative Art, Knowledge and Ethics

---

## Edited book / Urednička knjiga

Publication status / Verzija rada: **Published version / Objavljena verzija rada (izdavačev PDF)**

Publication year / Godina izdavanja: **2019**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:186:633917>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-02-23**



Repository / Repozitorij:

[Repository of the University of Rijeka, Faculty of Humanities and Social Sciences - FHSSRI Repository](#)

# NARRATIVE ART, KNOWLEDGE AND ETHICS

Edited by  
IRIS VIDMAR JOVANOVIĆ



University of Rijeka, 2019

*Title*

Narrative Art, Knowledge and Ethics

*Editor*

Iris Vidmar Jovanović

*Publisher*

Faculty of Humanities and Social Sciences

University of Rijeka

Sveučilišna avenija 4, 51 000 Rijeka

[www.ffr.uniri.hr](http://www.ffr.uniri.hr)

[www.uniri.hr](http://www.uniri.hr)

*For the publisher*

Ines Srdoč-Konestra

*Language editing*

Iris Vidmar Jovanović and David Collins

*Reviewers*

Prof. dr. sc. Snježana Prijjić Samaržija

Prof. dr. sc. Nenad Mišćević

*Design, Print & Cover*

Grafika Helvetica d.o.o. Rijeka

[www.grafikahelvetica.com](http://www.grafikahelvetica.com)

*Publishing date*

December 2019

*Editor and contributors*

Iris Vidmar Jovanović, James O. Young, Ana Maskalan, David Collins,

Britt Harrison, James Hamilton

ISBN: 978-953-7975-92-0

The CIP record is accessible at the Computer catalogue at the University Library  
in Rijeka under the number 140817090.

This book is published with the support of the University of Rijeka (research  
project Literature as a domain of ethics, no. 17.05.2.2.05) and the Faculty of  
Humanities and Social Sciences Rijeka.

NARRATIVE ART,  
KNOWLEDGE AND ETHICS

Edited by  
IRIS VIDMAR JOVANOVIĆ

University of Rijeka, 2019



## Content

Notes on contributors . . . . .	7
Acknowledgments . . . . .	9
<i>Iris Vidmar Jovanović</i>	
Preface . . . . .	11
<i>Iris Vidmar Jovanović</i>	
Cognitive and Ethical Values and Dimensions Of Narrative Art . . .	17
<i>James O. Young</i>	
Literary Fiction and the Cultivation of Virtue. . . . .	87
<i>Ana Maskalan</i>	
Better Worlds and Mark Twain's Submarines: Utopian Literature as a Stimulus for Social Engagement . . . . .	109
<i>David Collins</i>	
(When) Is 'Engaged' Art Existentially Authentic? de Beauvoir contra Sartre . . . . .	123
<i>Britt Harrison</i>	
Introducing Cinematic Humanism: A Solution to the Problem of Cinematic Cognitivism . . . . .	175
<i>James Hamilton</i>	
Arts & Ents . . . . .	201
Index . . . . .	221



## Notes on contributors

**David Collins** is currently a doctoral candidate in the Department of Philosophy at McGill University in Montreal, where he is finishing a dissertation on R. G. Collingwood's aesthetics, exploring the links between Collingwood's thinking and process philosophy, American pragmatism, and existential phenomenology. He also has strong interests in moral philosophy (especially virtue ethics and existentialist ethics) as well as practical and theoretical issues concerning education and the psychological, ethical, and social costs of modern technologies. Coming to philosophy from an academic and practical background in film and theatre, he takes seriously Collingwood's rapprochement between theory and practice, and is concerned not just with ontological and epistemological questions about art, but with art's importance for human life at both the individual and social levels.

**James R. Hamilton**, Emeritus Professor of Philosophy at Kansas State University in Manhattan, Kansas, USA, works on aesthetic issues and issues in the philosophy of art. He is an empirically and formally oriented philosopher, informing his work using studies in decision theory, formal learning theories, neuroscience, cognitive psychology, and cognitive anthropology. He has published a book with Wiley-Blackwell (2007), articles in collections published by the University of Michigan Press, de Gruyter, Cambridge Scholars, and Roman & Littlefield, peer-reviewed articles in the *Journal of Dramatic Theory and Criticism*, *The Journal of Aesthetics and Art Criticism*, *British Journal of Aesthetics*, *Puppet Notebook*, *Journal of Aesthetic Education*, *Organon F*, *Revue internationale de philosophie*, *Philosophy Compass*, *Performance Research*, *Croatian Journal of Philosophy*, and others, in addition to entries in several encyclopedias and companions to aesthetics, theatre, Shakespeare, and performance

**Britt Harrison** is a screenwriter, script consultant and film-making teacher. Her first PhD, entitled *The Epistemology of Know-how*, challenged the propositional presumption of contem-



porary intellectualist views on know-how, drawing on resources from the Later Wittgenstein. She is currently at the University of York, UK, working on her second PhD, Cinematic Humanism, bringing together her interests in film and philosophy under the rubric of Philosophy of Film Without Theory.

**Ana Maskalan** is a Research Associate in the Centre for Youth and Gender Studies at the Institute for Social Research in Zagreb. She received her PhD from the Department of Philosophy at the Faculty of Humanities and Social Sciences, University of Zagreb. Areas of her research interest beside gender and utopian studies are theories of body identity, transhumanism, pop culture, cyberculture, culture of fear, and political and social issues of social justice and discrimination. She has participated in a dozen research projects. In addition to numerous research articles, she is the author of *Woman's Future: The Philosophical Discussion on Utopia and Feminism* (ISRZ; Plejada: Zagreb, 2015) and co-author of *Identity and Culture* (M. Labus, L. Veljak, A. Maskalan and M. Adamovic, ISRZ: Zagreb, 2014).

**James O. Young**, FRSC, is Professor of Philosophy at the University of Victoria. He is the author of several books and many articles on philosophy of art, metaphysics, and philosophy of language. He is currently working on a book on art and intellectual property and another on the history of philosophy of music.

**Iris Vidmar Jovanović** is an Assistant Professor in the Department of Philosophy, University of Rijeka. She works primarily in the philosophy of literature, but her interests include Kant's aesthetics, cognitive approaches to aesthetics, and issues in metaphilosophy. She is one of the organizers of the Philosophy of Art conference in Dubrovnik, Croatia, a secretary of Croatian Society for Analytic Philosophy, and a member of the Executive Committee of the European Society for Aesthetics.

## Acknowledgments

I would like to express my gratitude to all who supported this project: to the University of Rijeka and the Faculty of Humanities and Social Sciences for providing the means for publication; to all the authors of individual chapters, to the *Croatian Journal of Philosophy* – primarily to Nenad Miščević and Kruno Zakarija for their most generous permission to republish two essays originally appearing in the Journal; to *Popular Inquiry: The Journal of the Aesthetics of Kitsch, Camp and Mass Culture* and its editors Max Rynnänen and Jozef Kovalcik, including the generous assistance of Adam Andrzejewski, for allowing the publication of James Hamilton’s paper. In addition, in one way or another, the most valuable support came from my collaborators in the project *Literature as a Domain of Ethics*, from colleagues in the Department of philosophy in Rijeka, primarily Snježana Prijčić Samaržija and Elvio Baccarini, from the audiences at the conferences where the papers pertaining to the project were presented, particularly the Philosophy of Art conference held annually at the Inter-University Centre Dubrovnik, and from numerous students who took my class on Philosophy of Literature. The gratitude extends to the reviewers, whose generous comments helped improve certain arguments, and to Dora Kunović whose help with the technical aspects of preparing this book can hardly be overstated. This would not be possible without the generous support of David Collins – thank you for your faith in, and unyielding dedication to this project, not to mention your patience and kindness with proofreading the original drafts. Finally, thanks to my family and my husband for their love, support and encouragement.

This book is dedicated to all the teachers out there, who believe in the power of narrative art to make our lives more enjoyable, our future a bit brighter and our world a better place.

*Iris Vidmar Jovanović*  
Rijeka, October 2019



# Preface

*Iris Vidmar Jovanović*

---

The present collection sprang from a research grant awarded by the University of Rijeka, number 17.05.2.2.05. The research project, entitled *Literature as a Domain of Ethics*, was primarily dedicated to exploring the cognitive and ethical dimensions of literature. The aim was to show that, as a cultural practice, literature is saturated with cognitive potential and that it has the capacity to bring forward ethical concerns, thus providing its readers with an opportunity for developing moral sensibility. Overall, the hope was to reaffirm the educational role of literature and its social and cultural relevance. That same hope remains at the core of this collection, which expanded the domain of interest to narrative art, incorporating literature and film, with the prospects of extending to quality television shows and theatre plays.

The essays gathered here are held together by jointly shared intuitions: works of narrative art provide not only aesthetic delight, but opportunities for learning, advancement of cognitive skills and refinement of moral sensibility. They also stand united in recognizing deep connections that obtain between narrative art, our reality and experience, connections which quite often give narrative artworks a dimension of social significance and engagement. Another shared thread is the experience of turning to art for its capacity to give us means – cognitive, emotional, ethical – to make sense of our experience, ourselves and other people. Thus, taken together, the essays present a defence of the view that the arts should be awarded an important place among our cultural and educational values.

Against such shared common ground, each essay develops a unique line of inquiry into a more narrowly defined area. The first part of the book is an extended essay intended as an outline of the

view that narrative art is epistemically valuable and that it relates to our ethical concerns. It is not a polemical or argumentative essay, though it is developed as an epistemological defence of aesthetic cognitivism – the view that art is cognitively valuable. Iris Vidmar Jovanović presents the main claims behind this view, bolstered by examples from literary tradition and the practice of literary criticism. She then introduces anti-cognitivists' challenges and provides epistemological pointers specifying how to address these challenges. The resources she employs from contemporary epistemology enable her to offer a more coherent theory on the cognitive value of art, one which does not succumb to challenges traditionally issued against art's capacity to convey truth. Readers sympathetic to such views can look for details of her account in her doctoral thesis, *The Nature of Fictional Testimony and its Role in Reaching, Fulfilling and Promoting Our Epistemic Aims and Values* (2013).

In the second part, five philosophers defend their particular views on matters pertaining to narrative arts, knowledge and ethics, discussing art's capacity to benefit individuals and to stir social change. In doing so, they start from epistemology and ethics rather than art or aesthetics. In other words, questions of whether, or how, art is cognitively valuable and how it relates to our moral and ethical concerns are, for the most part, kept apart from a slightly different question, the one asking whether the cognitive/ethical dimension of arts promote its overall value. Though the latter question is of grave importance, particularly for philosophical aesthetics and philosophy of art, the scope of the book is not wide enough to include it. Rather, the focus remains on exploring narrative art's particular and distinctive manners of being an epistemic instrument in our attempts to understand our world, and its notable ways of engaging our cognitive capacities and moral agency.

James O. Young's essay, originally published in the *Croatian Journal of Philosophy* (XIX, 56, 2019), develops the philosophical claim regarding the capacity of literary fiction to make people more virtuous. Young begins his essay by defending the view that this claim is empirical. He goes on to review the most recent empirical literature and finds that it supports the claims philosophers have

made. He identifies three mechanisms whereby reading literary fiction makes people more virtuous: empathy is increased when readers enter imaginatively into the lives of fictional characters; reading literary fiction promotes self-reflection; and readers mimic the pro-social behavior of fictional characters. The paper concludes with a caution: there is a danger that readers could mimic antisocial behavior displayed in literary fiction. If they do, reading some literary fiction could make readers less virtuous. The particular value of this essay is in bringing together conceptual work done by literary and aesthetic cognitivists and contemporary psychological research. While the findings of such collaboration might seem to jeopardize the overall project of advocating for the educational value of art – if, after all, moral corruption is just as possible as moral development, we might be better off if we follow Plato’s strict paternalism or abandon art altogether – Young’s essay provides plenty of reasons to stick to the main premises of aesthetic cognitivism and continue further with the incorporation of psychological evidence into philosophical research.

Ana Maskalan’s essay introduces the topic of art’s social engagement. In her contribution, she discusses features of literary utopias that make them socially engaging. Understanding social engagement as a form of transcending the existent and creating a better life, a better society and a better world, in the first part of the paper Maskalan argues that the contemporary social rejection of utopias is not primarily a consequence of their naivety or inconclusiveness, but of their promise and incitement of social change. As she sees it, there are four attributes that can explain the engaging capacities of literary utopias – feasibility, criticality, democracy of authorship and seductiveness. Each of these is described and contextualized in the second part of her essay. The seductiveness of literary utopias lies in their ability to immerse readers in their imaginary worlds, describing life as it could and should be. Utopian feasibility reminds utopian readers of their responsibility for creating a better world and future. Utopian criticality represents an articulation of awareness and attitude towards the real world. Utopian democracy of authorship refers to the creation of a space for

different voices, different experiences and different visions. Mas-kalan concludes by pointing out the dangers of giving up utopia, not just for the fate of utopian literature and utopian thinking in general, but for the fate of social change as well.

Continuing the topic of art's social engagement, David Collins turns to Jean-Paul Sartre's and Simone de Beauvoir's views on literature. Collins begins by exploring Sartre's view, expressed in *What is Literature?*, where Sartre outlines a theory of politically committed or 'engaged' literature and argues that authors have an obligation to write engaged literature, so understood. While Sartre makes a number of remarks that are not always obviously compatible, the dominant account of engaged literature that emerges presents it as a matter of writing about the social issues of one's day in order to critique political problems and advocate for progressive solutions that promote social freedom, with the implication that a work is better or worse qua literature to the extent that it does, or does not, do this. As Collins argues, not only is this criterion of literary value implausible but Sartre's account faces problems, including apparent internal inconsistencies and the likelihood that it ends up in what Sartre calls 'bad faith'. While not directly responding to Sartre, two essays of Simone de Beauvoir's set out an alternative account of literature (which is compatible with several 'outlier' remarks of Sartre's that don't obviously fit with his apparent position) that gives another way to understand how literature – and by extension, other artforms – can be politically engaged and promote human freedom, not by being about freedom at the level of 'content' but by embodying it at the level of 'form' and in the relation that is established with the reader. After explicating both positions and outlining Sartre's existentialism as the background against which they can be understood, Collins develops and argues in favour of de Beauvoir's account and discusses how it speaks more plausibly to the cognitive, moral, and social or political values of literature and other arts.

The final two essays take a more skeptical turn towards the main claim of this collection, though not entirely a negative one. Britt Harrison sets out to defend cinematic humanism while simul-

taneously harshly criticizing the theoretical machinery of aesthetic and literary cognitivism. A Cinematic Humanist approach to film is committed *inter alia* to the following tenet: some fiction films illuminate the human condition thereby enriching our understanding of ourselves, each other and our world. As such, Cinematic Humanism might reasonably be regarded as an example of what one might call ‘Cinematic Cognitivism’. This assumption would, however, be mistaken, Harrison claims. For Cinematic Humanism is an alternative, indeed a corrective, to Cinematic Cognitivism. Motivating the need for such a corrective is a genuine scepticism about the very notion of the cognitive. Using historical reconstruction, Harrison reveals how ‘cognitive’ has become a multiply ambiguous, theory-laden term in the wake of, indeed as a consequence of, Noam Chomsky’s original stipulative definition. This generates a constitutive problem for cognitivism as both a research programme and a set of claims, and as such poses a trilemma for philosophers of film, art and beyond. Harrison proposes a Cinematic Humanist solution to the problematic commitments of cognitive film theorising and, in so doing, gestures towards a methodology she is calling ‘philosophy of film without theory’. Harrison’s paper was originally published in the *Croatian Journal of Philosophy* (XIX, 56, 2019).

James Hamilton’s contribution points to some of the worries that a cognitivist might face in her development of a theory supporting the main suppositions of aesthetic cognitivism. Hamilton presents three problems, and constructs a fourth, that arise concerning arts and entertainments. The first problem with the distinction between arts and entertainments relates to the fact that there may be no successful way to draw the distinction. Thus, Hamilton’s essay engages with the hard question of differentiating what is sometimes construed as good vs. bad art, i.e. art imbued with cognitive and ethical value and art devoid of any but purely entertaining features. Hamilton further considers whether arts or entertainments can do philosophy in the same way that philosophy does it. The problem is, how, precisely, do we learn normative lessons from works of art and entertainments when they express a deep philosophical insight? It has been argued that works of popular en-



tertainment are thought experiments or, more narrowly, intuition pumps. However, Hamilton argues, overlooked in the discussion is that some works expressing such insights seem to act as if they were parables. He describes thought experiments, intuition pumps, and parables, and notes some ways in which thought experiments and intuition pumps differ from parables. His third problem concerns what Hamilton calls the “two-audiences” problem. This problem arises initially from the simple fact that there are cases of popular entertainment that are attended to both by people who normally attend only to traditional instances of art *and* by people who normally only attend to popular entertainments. When this simple fact is conjoined to a plausible back-story about how we determine what the audience is for a particular art form or bit of entertainment, we get the problem: are we the same auditors when we “like” Bach’s unaccompanied cello suites and also “like” Leon Redbone’s blues – or any other pair of sufficiently disparate objects? Well, yes, of course, we are. But it is also true that we have some work to do in order to maintain a coherent sense of our beliefs and attitudes. The problem he constructs is an uncomfortable implication of an obvious way of conjoining (i) the facts that some works of popular culture enable us to learn the lessons embedded in popular entertainments in the same manner as parables do and (ii) the two-audiences problem. In a nutshell: if we engage in telling the parable expressed in a work of art or of popular entertainment and if we endorse its normative stance, then it becomes unclear how the parable can be uttered and to whom it can be delivered. Although this is a problem that besets the practices of only a particular group of scholars, mostly some of those who work in what is now called “cultural studies,” Hamilton suggests that the obviousness of the conjunction means it is a quite general problem for a very common way of thinking about, learning from, and teaching the lessons embedded in popular entertainments. Hamilton’s paper was originally published in *Popular Inquiry: The Journal of the Aesthetics of Kitsch, Camp and Mass Culture*, 2/2018.

# Cognitive and Ethical Values and Dimensions of Narrative Art

*Iris Vidmar Jovanović*

---

... *Madame Bovary* addresses major problems in human relations and ethics. (...) Flaubert's sharp criticism of the excesses of capitalist societies – greed, exploration, and consumerism – apply to our own day. His keen analyses of people's difficulties in communicating with each other in dysfunctional relationships illuminates much of our own experience, regardless of our gender, ethnicity, or background. Despite the humble, obscure social setting in which it plays out, Emma's narcissistic personality disorder typifies the tragic hero throughout literature. The fate of this charming, intelligent, but impulsive and deeply selfish person raises many feminist issues in the context of a society where women were not allowed to vote, move, open a bank account, hold a passport, or start a business without their husbands' permission until World War II (Porter & Gray 2002, xiv).

Underneath this powerful commentary regarding one of the world's most famous literary masterpieces is the idea that literature is most deeply connected to our worldly experience, so much so that it can be seen as its mirror. To claim that Flaubert *addresses* 'major problems in human relations and ethics' suggests his capacity to carefully observe human relations and identify such problems, that is, to truly understand difficulties, challenges, and issues contained within them, and to write about them in a way which enables the audience to gain a similar kind of understanding. And so does the claim that, through his fiction, he *criticises, analyses* and *illuminates* our social reality as well as our most personal, subjective domain. Finally, his capacity to *raise issues* demonstrates literature's capacity to be provoking and challenging and to make us think, reflect and reconsider things we take for granted as well as those that never crossed our minds. Most significantly, the commentary reveals why it would not be an exaggeration to say that Flaubert's

novel is, in addition to being aesthetically and culturally praiseworthy, a valuable source of knowledge. How can it be that it – and others like it – can be all of these, is the question I pursue here.

The main premise in this essay is that some works of narrative arts, including literature, films, theatre plays and TV series, are not only artistically, aesthetically and culturally valuable, but also epistemically and ethically valuable. To claim that narrative art is cognitively valuable means, in the most ‘epistemic’ sense of the word, that we can learn from it and obtain significant cognitive benefits. To claim that it is ethically valuable means that a lot of what we read about in works of narrative art relates to the fundamental ethical challenges that we face as human beings, which potentially enables us to learn something about those very challenges. In the first part I present theories which explain where such value is, and those that negate it. I then move on to say a bit more about the connection between narrative art and human experience, and to explicate ways in which engaging with narrative art is cognitively beneficial. Towards the end, I explore the intersection of narrative art and ethics, before turning to explaining cognitive mechanisms which enable narrative art to influence us along the cognitive dimension.

## **1. Narrative Art and Cognitive Value: Aesthetic Cognitivism**

The task of explicating the cognitive value of narrative art consists in showing and explaining what cognitive benefits are available to those who engage with it, and how they are available. This task has mostly been carried out by aestheticians, not by epistemologists, which is why our debates today exhibit a considerable lack of epistemological insights and are underlined with a conceptual framework of aesthetic thoughts on the matter. For the most part, the cognitive dimension of art and its influence on the cognitive functioning of the audience has been accounted for via aesthetic resources, primarily by arguing that the cognitive value of a work relates to its overall artistic and aesthetic value. Thus, the gist of aesthetic cognitivism – an umbrella term referring to multiple the-

ories which acknowledge art's cognitive dimension<sup>1</sup> – is “that the cognitive (or, equivalently, epistemic) merits of works of art are, under certain conditions, aesthetic merits in those works or condition their aesthetic merits.” (Gaut, 2007 p. 136). In this essay, I rely upon such a cognitivist framework, but I approach it from the standpoint of epistemology: I seek to explain how narrative art can be an epistemically reliable and justified source of knowledge and other cognitively valuable states.

The claim that narrative art is cognitively valuable implies that it offers cognitive benefits or payoffs to its readers, i.e. that it has effects on readers which can best be described as cognitive (rather than emotional or aesthetic). The theory does not argue that by definition, each reader will pick up or make use of such benefits: what is needed is an active reader, willing to reflectively engage with the work. Such a reader can greatly profit, not only in terms of gaining aesthetic and artistic pleasure or entertainment, but in terms of enriching her conceptual scheme, developing her imaginative capacities, sharpening her moral sense and learning new things. Not all works of narrative art are in this sense valuable, but some are, and the aim in this essay is to explore how the audience can make use of those which are.

The most straightforward way to motivate aesthetic cognitivism is to say that readers pick up true propositions from a work, that is, that narrative art is a source of truth. However, there are two senses in which we can speak of narrative art and truth.<sup>2</sup> One relates to the truth that is internal to the work, when we ask *what is true in a given work*. Whether or not Othello killed Desdemona depends on how things are in the fictional world of *Othello*, that is, on what Shakespeare wrote. It is the play, not the real world, that determines such truth. An integral part of engaging with any work

1 See for example: Carroll 2002, 2007, Davies 2007a, 2007b, 2010a, Elgin 1993, 1996, 2000, 2007, Friend 2007, Gaut 2005, 2006, 2007, 2016; Gibson 2003, 2007, 2009, 2011; Graham 1997, 2000; Green 2010; Kieran 1996, 2005; Kivy 2006; Mikkonen 2013, 2015; Novitz 1984; Row 2010; Zamir 2002, 2006; Walsh 1969; Young 2001; Vidmar 2010, 2012b, 2013, 2014b, 2017a, 2019.

2 See Davies 2016.

consists in the audience's interpretive attempts to come up with a coherent account of what happened and why. This is not always straightforward, as various instances may remain ambiguous and thus subject to individual interpretations. The question of Hamlet's motivation is one such example, as evidenced by decades and decades of multiple interpretations of the famous play. For all that Shakespeare put in it, it remains unclear whether our tormented prince suffered from an Oedipus complex or was simply too contemplative to act.

In order to understand any given work, the readers or spectators bring in and rely on their knowledge of how the world is – they know that Gertrude will die after she drinks the poison, even though Shakespeare does not say that poison is lethal. They are able to make sense of what they read and to supplement the content that is not explicitly given because the fictional world and real world resemble one another. This is the idea behind the so called *principle of verisimilitude*. As explained, “fictional states of affairs (objects, event, personages) can be assumed to be like ordinary states of affairs (objects, events, personages) failing indications to the contrary” (Lamarque and Olsen, 1994, p. 95). This is why, as some cognitivists claim, we can learn about the real world, that is, why occasionally we pick up truth about the real world from the work. It is this sense – *the truth proper*, i.e. the truth about the real world, as revealed through work – that matters to us here. The fact that Jim was a runaway slave is true in Mark Twain's novel, but it is also a real world fact that there were people like Jim. Slavery was not made true by Twain's novels but by the state of affairs in the world, by political, juridical and social institutions in America. In writing *The Adventures of Huckleberry Finn*, Twain incorporated the worldly truth into his fictional world, thus opening up the possibility to the audience to learn something about the real world from his novel.

Not everyone agrees that an epistemological account of narrative art is appropriate. Most famously, Peter Lamarque and Stein Haugom Olsen object to emphasizing truth and other cognitive merits in discussing literature:

*Of course* readers can pick up information about people, places, and events from works of fiction; *of course* readers can learn practical skills, historical facts, points of etiquette, insights into Regency England, etc., from literary works; *of course* writers of literary fiction often offer generalizations about human nature, historical events, political ideologies, and so forth, in their works; *of course* what readers take to be true (in the world) will affect how they respond to literary works, including how they understand the works; *of course* readers often need to have background knowledge of a cultural, psychological, or historical kind, even moral or philosophical preconceptions, to understand some literary works.

The theoretical interest lies not in defending these commonplace observations, but in integrating them into a satisfactory account of literature, literary value and fictionality (Lamarque and Olsen 1994, pp. 4-5. *Italics original*).

According to their theory, no such integration is needed, as the cognitive dimension of a work does not relate to its artistic value. What they object to is the view of

those who want a ‘stronger’ sense of ‘truth’ and ‘falsity’ applied to literature; i.e. those who see the aim of literature as conveying or teaching or embodying universal truths about nature, the human condition, and so on, in a sense at least analogous to that in which scientific, or psychological, or historical hypothesis can express general truths (Lamarque and Olsen, 1994, p. 6).

On the view defended here, it is not *the aim* of literature (and narrative art generally) to convey or teach or embody truths – though, as the history of literary periods and genres reveals, sometimes that is precisely the aim – but the fact is that (some) literature does so (and much more in terms of providing cognitive benefits), which gives us a reason to subject literature to an epistemological analysis, without thereby robbing it of its aesthetic and artistic value.<sup>3</sup> To do so, we will start by looking at those theories which embrace the intuition we want to defend regarding the cognitive value of some works. One such is the view of Berys Gaut:

Aesthetic cognitivism, then, is best thought of as a conjunction of two claims: first, that art can give us (non-trivial) knowledge, and second, that the capacity of art to give us (non-trivial) knowledge (partly) de-

---

3 See Zamir 2002 for an elaborate version of this argument.

termines its value qua art, i.e. its aesthetic value. Aesthetic anti-cognitivism is a denial of one or both of these conjuncts (Gaut, 2005, pp. 436 – 437).

Our epistemological perspective will focus on the first of these claims – art can give us non-trivial knowledge – although, as we will see shortly, we will expand the cognitive impact of art beyond knowledge. One reason to take an epistemic perspective on narrative art has to do with the fact that it has traditionally been attributed some cognitive values (such as understanding) that have only recently been given an epistemological grounding, with the so called value turn in epistemology.<sup>4</sup> Art is saturated with different cognitive values which can be neglected if the challenge is cast in terms of the pro-truth vs. no-truth debate, i.e. if we only care about defending (or negating) the view that readers pick up true statements from the work (a view sometimes called the propositional theory of cognitive value). Most aesthetic cognitivists argue that truth is neither the most important nor the only available cognitive benefit that art affords. In fact,

most cognitivists correctly hold that there is a wide variety of different kinds of knowledge that art can impart to its appreciators: propositional knowledge, know-how (skills), phenomenal knowledge (knowledge of what it is like to experience something), conceptual knowledge, knowledge of values and of significance, for example (Gaut, 2006, p. 115).

If indeed we can show that narrative art can afford us benefits along these lines, it would seem we have strong reasons to engage with art in our everyday life, and to give it a more firm space in our culture and education. Let us then start by exploring some of the reasons why we might think that art does not afford us cognitive benefits.

## 2. A Sceptical Take on Narrative Art's Cognitive Value: Aesthetic Anti-Cognitivism

Debates on the cognitive value of art and literature originate

---

<sup>4</sup> See Elgin 1993, 1996, 2000; Kvanvig 2003; Prijic-Samaržija and Bojanić 2012; Prijic-Samaržija 2019 (part 1); Riggs 2003; Vidmar 2013.

historically with Plato and what is known as the ‘ancient quarrel’.<sup>5</sup> Plato’s criticism of art is layered and complex, and it presupposes a particular view of art (i.e. a mimetic theory of art, which he refers to as imitative poetry) as well as his ontological views on the distinction between the world of appearances and the world of ideas. Most importantly however, his beef with the artists is primarily caused by the educational role that art, primarily epic poetry, had in his days.<sup>6</sup> Poetry was considered revelatory of the truth about the origin of the world, and imbued with morally relevant lessons that the young were supposed to learn in order to become ethically and politically virtuous. However, Plato set out to refute this standing, though without negating its aesthetic and artistic appeal.

Plato’s epistemic criticism of art is scattered throughout his dialogues *Republic*, *Ion* and *Apologia*. The most pressing epistemological challenge presented in *Republic* is in book X, where he argues against the epistemic authority of literature. The gist of his argument is the claim that imitative poetry is thrice removed from the ‘the real thing’ and is only presenting a thing under some appearance (not as the thing is in itself). An artist ‘makes things’ by holding a mirror against nature (*Republic*, X, 596e). However, by creating things in such a way, one “couldn’t make the things themselves as they truly are” (596e), only appearances of them. Plato thus concludes that poets are not presenting things (moral order, human nature, god’s nature etc) as they are, but only under some appearance. Consequently, art does not give us true but only partial images of things.

As we will see below, when we explore James Young’s (2001) and Matthew Kieran’s (2005) theories, contemporary cognitivists have turned this argument into an advantage, claiming that, because art presents things under some appearance (i.e. only one aspect), it is particularly good at focusing the audience’s attention to that particular aspect of what is presented. In everyday life, the argument goes, we are often too busy to consider things in their

5 See Barfield 2011; Kaufmann 1992, on the ancient quarrel.

6 See Delija Treščec 2005.



entirety. Consequently, there are many nuances of how things are that we miss. The value of art is in throwing light on such nuances which, quite often, go unnoticed.

Another layer of epistemic criticism holds that ‘imitation can produce everything’ independently of the actual thing. Plato says:

For example, we say that a painter can paint a cobbler, a carpenter, or any other craftsman, even though he knows nothing about these crafts. Nevertheless, if he is a good painter and displays his painting of a carpenter at a distance, he can deceive children and foolish people into thinking that it is truly a carpenter (*Republic*, 10, 598,c).

The claim that imitation ‘can produce everything’ implies that artistic and literary creation is not bound with ‘what is out there, in the real world’. David Davies incorporates this into his theory of fiction, when he claims that, unlike in the case of non-fictional narrative, fictional narratives do not have to follow *the fidelity constraint*, i.e. be true to the facts.<sup>7</sup> Consequently, while no one will conflate a painting of a carpenter with a real carpenter, someone might take a fictional description or statement for a factual one. Thus, one might come to believe that Jane Austen is right when she claims that “It is truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife.” (Austen,

---

7 In Davies’ words: To read a narrative as *non-fiction* is to assume that the selection and temporal ordering of *all* the narrated events was constrained by a commitment, on the narrator’s part, to be faithful to the manner in which she takes actual events to have transpired. I term this constraint on narrative construction the ‘fidelity constraint’. To read a narrative as *fiction*, on the other hand, is to assume that the choices made in generating the narrative were not governed in the first instance by this constraint, but by some more general purpose in story-telling, such as entertaining the reader or illuminating some more general feature of the world, as in scientific or philosophical thought experiments (Davies, 2015, pp. 39-40; see also his 2007a). John Gibson explains the problem in this way: “What an author of a work of fiction writes is constrained only by her ‘creative’ imagination, as it is sometimes put; it is not in any obvious way guided by an attempt to depict how things stand in reality. (...) But if an attempt to transcribe how things stand in the world does not guide an author’s pen, then how could the product of that pen’s activity cast light on that world? Put differently, the very idea of writing a work of imaginative literature – of waving fictional narratives – seems incompatible with the notion of fidelity to the world, of bearing witness to it.” (Gibson, 2007b, p. 3).

2005, p.11). An even more pressing problem is the claim that an artist does not have knowledge of the things he writes about, which makes his work epistemically unreliable as a source of knowledge. In *Apology* Plato argues:

I soon realized that poets do not compose their poems with knowledge, but by some inborn talent and by inspiration, like seers and prophets who also say many fine things without any understanding of what they say. The poets seemed to me to have had a similar experience. At the same time I saw that, because of their poetry, they thought themselves very wise men in other respects, which they were not (*Apology*, 22,b,c).

As Plato sees it, a poet can deceive the audience into believing his descriptions, in light of the alluring poetic phrasing:

And in the same way, I suppose we'll say that a poetic imitator uses words and phrases to paint coloured pictures of each of the crafts. He himself knows nothing about them, but he imitates them in such a way that others, as ignorant as he is, who judge by words, will think he speaks extremely well about cobblery or generalship or anything else whatever, provided – so great is the natural charm of these things – that he speaks with meter, rhythm, and harmony, for if you strip a poet's works of their musical colourings, and take them by themselves, I think you know what they look like (*Republic*, 10, 600e, 601b).

On Plato's view, the artist and the audience are both ignorant: the artist is not an 'expert' with respect to the thing he is imitating, while the audience has no expertise regarding the subject of imitation so as to judge and evaluate the represented qualities of the object. However, the advantage that an artist has over the audience is his *techné*: he has the skills to use artistic means (words, phrases, meter, rhythm) so as to create an illusion that he is knowledgeable. This is precisely why imitative poetry is so dangerous: poets only give us the appearance of wisdom because they know how to employ artistic means in order to create an imitation, which in fact doesn't reveal anything about the true nature of the thing imitated. In *Ion*, Plato adds a new charge to this argument, when he points out that there are people who are more knowledgeable than Homer on matters Homer writes about. Why then should we trust Homer, rather than those other people?

In contemporary discussions, this line of reasoning is sum-

moned under the ‘no expertise’ argument, here explicated by Noel Carroll: “Artists as such, it is charged, have no special expertise in any branch of knowledge other than that pertaining to their art form and its medium” (Carroll, 2007, p. 28). The no expertise argument is convincing, but it is not a conclusive argument against the cognitive value of art. For one thing, knowledge of artistry does not preclude knowledge in other domains; in fact, literary creation and artistic practice are compatible with (and best explained by) the assumption that authors themselves have knowledge of the things they write about. Otherwise, it would be hard to explain our sense of recognizing the truthfulness of their descriptions, the sense that they speak informatively about reality. It is not uncommon for authors to conduct all sorts of enquiries on the subject they write about – certainly Homer could have obtained knowledge about the crafts he writes about, and authors generally can consult relevant sources. A telling example comes from Flaubert, who claimed to have read about hundreds of books in preparation for writing *Salambo*, and although this claim has been challenged, the fact remains that most of what we read in *Salammbô* can be traced back to Polybius’ *Histories*.<sup>8</sup>

A different approach to the no expertise argument is to point to the fact that most of what literature puts to view does not fall under any particular expertise, as it relates to human experience in the widest sense of the term. For that reason, it is not uncommon to claim, at least within literary and art criticism, that authors are observers of the society to which they belong to, and thus have a firsthand experience of. For example, Henry James presents to his readers “the world he knew, that of rich and cultivated upper class of his time in America and Europe” (Scofield, 2006, p.79), and one of his most obvious interests was in “the changing nature of the modern (usually American) woman and the social attitudes that

---

8 Krailsheimer 1977, *Introduction to Salammbô* (Penguin Classics) claims that „factual accuracy was not optional extra for Flaubert and in his attempted recreation of the past he demands to be judged by the most exacting standards of scholarship, as well as art“.

surrounded her” (Scofield, 2006, p. 80). Thus, James – and, we may add, a whole bunch of other writers such as Edith Wharton, Theodore Dreiser, Jane Austen, Charles Dickens, D. H. Lawrence, Leo Tolstoy, Dostoyevsky – have “‘snapshotted’ one small but vivid corner of the social scene” (Scofield, 2006, p. 82). Even outside of realism – where the epistemic norms of accurately reporting what one observes collides with aesthetic norms of presenting the reality as truthfully as possible – writers can be said to provide a ‘snapshot’ of reality. Writing about Chaucer’s *Canterbury Tales*, Ronald Carter and John McRae claim that the work

absorbs literary, historical, religious, social, and moral concerns, and transcends them all. It gives a wide ranging view of the late fourteen century world and its people. The specific people and places described become emblems of their period and the text becomes an image of its time. (...)

Literature, with Chaucer, has taken on a new role: as well as affirming a developing language, it is a mirror of its times – but a mirror which teases as it reveals, which questions while it narrates, and which opens up a range of issues and questions, instead of providing simple, easy answers (Carter and McRae, 2001p. 33 - 35).

This take on the no expertise argument can help us explain why the artistic *techne* is so important: what is needed to make someone recognize particularly intriguing aspects of shared human experience and social reality is precisely the technique to present common experience in ways in which we can recognize it as relevant to our own lives.<sup>9</sup> Certainly, men and women of Flaubert’s time participated in the same set of material circumstances as he did, but it was only Flaubert who managed to write about those circumstances with enough precision to enable his audience to come to grasp the underlying conditions of those circumstances. It is for that reason that we walk away from *Madam Bovary* with the sense of realizing what the struggle for an accomplished, happy life can mean, and how external circumstances determine the boundaries of one’s actions.<sup>10</sup> Absolutely then is artistic expertise relevant, for

9 Consider Attridge 2015 for one way in which this idea has been developed.

10 See Porter and Gray 2002 for an exploration of social circumstances in *Madam Bovary*.

without such expertise, we would be left with a dull and trivial imitation of reality that would not reveal its significant layers.<sup>11</sup> Naturally, this isn't to suggest that in matters when a certain expertise is needed – in matters of whaling, as is the case with *Moby Dick* or with the show business industry as evidenced by *Sister Carrie* – the authors do not conduct research which enables them to write knowledgeably on these matters.

Another way to diminish the no expertise argument is to point to the fact that readers have discriminatory capacities which enable them to track reliable descriptions. One set of such capacities relates to genre familiarity. As Stacie Friend argues, “readers familiar with genre conventions or the techniques of certain authors – even if they cannot articulate the specific conventions or techniques – are more likely to track true and false information accurately” (Friend, 2007, p. 48). Familiarity with genre extends to the capacity to discriminate differences with respect to different authors, not only in terms of artistic/aesthetic aims they are pursuing, but with respect to their reliability.<sup>12</sup> A sensitive reader is aware of a difference in presenting psychological, sociological, political and economic facts in Theodore Dreiser and in Hedwig Courths-Mahler. Both novelists create the background to their stories by depicting social and economical circumstances surrounding the development of factories. However, while in Dreiser this dimension is described in depth and provides a psychological grounding for the motivation of the main character, in Courths-Mahler the setting is notoriously banal, unconvincing and it obviously presents an oversimplified view on industrial development. The portrayal of human psychology in Dreiser and Courths-Mahler differs radically precisely because Dreiser is epistemically more reliable, more ‘true to life’, more informative on what he is describing.

---

11 Consider Kieran's take on this: “If the artistic means utilized are poor, clumsy or impoverished, then a work has failed to realize the affective understanding we value in much great art. In such cases we are unlikely to care about or take much interest in whatever insight is implicit in our experience of the work” (Kieran 2005, p. 120).

12 I develop this argument in Vidmar 2012b.

Finally, readers rely on their own experience and knowledge in evaluating a given work in terms of what that work presents, and can thus be alert to potential discrepancies between the state of affairs in the world and in the fictional world. However, it is not always easy to balance the fine line between fictional and factual elements. Plato would consider this a conclusive argument against ever taking narrative art as reliable: since we are never safe in forming beliefs on the basis of what we read in narrative art, we always lack justification for accepting such beliefs. Therefore, we should not believe anything we pick up from this source. That however seems like an exaggeration. On the one hand, quite often, narrative art is a reliable source of knowledge, even if it operates within slightly different norms of epistemic reliability. On the other, non-artistic sources of knowledge can occasionally be misleading. Art, in other words, is not the only context in which we are not epistemically safe.

Plato's no expertise charge is further supported in *Republic* by his observation that neither the poets (book 10 599b) nor the audience (book 10 600c-d) treat poets as the knowledgeable ones. While in itself this is at best an empirical claim which does not amount to a philosophically solid argument against the cognitive reliability of literature, Plato's further challenge might be:

But about the most important and most beautiful things of which Homer undertakes to speak – warfare, generalship, city government, and people's education – about these it *is* fair to question him, asking him this: "Homer, if you are not third from the truth about virtue, the sort of craftsman of images that we defined an imitator to be, but if you're even second and capable of knowing what ways of life make people better in private or in public, then tell us which cities are better governed because of you, as Sparta is because of Lycurgus, and many others – big and small – are because of many other men? What city gives you credit for being a good lawgiver who benefited it, as Italy and Sicily do to Charondas, and as we do to Solon? Who gives such credit to you?" Will he be able to name one?

None.

Or, as befits a wise man, are many inventions and useful devices in the crafts or science attributed to Homer, as they are to Thales of Miletus and Anacharsis the Scythian?

There's nothing of that kind at all (*Republic*, 10, 599d, 600).

The poets are now asked to prove that the knowledge they express in their poems can be put to ‘real use’ and can benefit those who rely upon it, which Plato thinks they cannot do. In contemporary discussions, Peter Lamarque issues a similar request to the audience:

Would we expect that those immersed in the great works of literature understand people and the world better than those who are not so well read? Yet there seems no evidence that such readers are especially knowledgeable about human traits, as are psychologists, or social scientists, or even philosophers. Literary critics are not sought out as experts or advisers on human affairs (Lamarque, 2007, p. 21).

Indeed, it can be argued that literature is useless, as it does not create anything new or practically useful, as the sciences do. However, there are other ways in which a certain practice can be useful. Literature reflects scientific discoveries by incorporating them into stories, and in that sense, contributes to dissemination of human knowledge. As a way of example, think of the psychological theories central to the Victorian novel of 19<sup>th</sup> century. Even if these theories are abandoned today, they testify to the progress of human knowledge.<sup>13</sup>

One aspect of literature that Plato ignores and that Lamarque does not consider is the fact that literature, in a manner similar to philosophy, is often concerned with questioning, criticizing, challenging or promoting certain values and beliefs, as well as political, religious, philosophical and scientific systems. One negative example of where literature did this comes from the so called America can frontier romances which struggled to come to terms with the concept of national identity in the face of white European settlers. Although various portrayals of Indians and their way of life can be found in literary fiction of the time, the fact remains that these novels advocate the need to eradicate them, or at least to find means to tame them and force them to accept the settlers’ way of life, in order for progress and civilization to settle in. Here’s how a literary theorist explains it:

---

13 For an analysis of this argument, see Vidmar 2014b, 2019.

Without exception, as far as I am aware of, all of the nineteenth-century frontier romances assume the inevitability of the conflict between whites and Indians as well as the eventual dominance of white Americans, but they approach this conflict and its result from somehow different perspectives. Some see it as tragic if inevitable (e.g. Cooper and Simms). Others see it as a subject of some degree of national disgrace (e.g. Child, Sedgwick, and Jackson). And still others view the first two positions as absurd, even dangerous, given the threat posed by this savage race to the forces of civilization (Bird, Edward Ellis, and Edward Wheeler) (Crane, 2007, p. 45).

Commentaries like this imply that narrative (and other art, see Kieran 2005) art often strives to make social impact and change. Sometimes, as David Collins analyzes in his contribution to this collection, philosophers tie literary value itself to the capacity to bring about social change.

Another telling example of how literature and art can make a change to how we understand various aspects of our world is science fiction, a genre ridden by a certain tension that springs from two of its seemingly opposite aspects. On the one hand, no other genre relies so extensively on scientific laws and pays so much attention to staying attuned to what the sciences have to say about such things as cybernetics, robotics, time and space travelling, artificial intelligence, etc. On the other hand, no other genre is as free as science fiction is to break the rules of mimesis and to abandon the principle of verisimilitude. Hence, no factual truths are available from such works. However, cognitive benefits latent in science fiction are multiple. Science fiction, it is said, “treads in the unknown, distances us from what we are familiar with, and puts in front of us a ‘novum’ – a discrete piece of information recognizable as not-true, but also as not-unlikely-true, not-flatly (in the current state of knowledge)-impossible” (Shippey pp. 13-4), thus creating ‘cognitive estrangement’. The crucial aspect of science fiction scenarios, the way that the novum ‘works’ and creates the estrangement, consists in focusing readers’ attention on some particular aspect of our world, humans, our scientific endeavours, the consequences of our scientific achievements or implementation of our designs, our political, sexual, psychological, ethical aspect, etc. By doing so, a



work puts in front of us, directly, that which it aims to explore, and it triggers our intuitions about it. The value is in literature's manner of forcing us to reflectively engage with things we might not consider without such incentives. In some cases, as Ana Maskalan describes in her contribution, such literature can have significant social impact.

One of the most optimistic replies to Plato's (and Lamarque's) challenge regarding literature's alleged uselessness comes from philosophers who claim that literature – or, more precisely, the insight we gain from literature into our human manners – helps us navigate better our social world and understand those around us. The underlying claim here is that literature, focused on imitating human relations and characters (in the psychological sense), fosters our understanding of how different types of people function. Literature manages to show us the intricate nexus of characters-emotions-actions, thus enabling us to grasp the causal mechanisms which underlie human behaviour. By carefully monitoring different characters' actions, we might realize how these follow from their beliefs, desires, attitudes and convictions. In some cases, there are valuable lessons in this, as we can sharpen our perception of human interaction in the real world. Noel Carroll traces this line of thinking to Aristotle. He writes:

Tragedy, on Aristotle's view, teaches us certain scenarios – certain regularities or tendencies in the course of human affairs – that are apt to occur when people with this or that set of dispositions or character traits are placed in various situations. In these cases, the constitution of their character will explain the ensuing pathway of events ... But since these tendencies are universal, given the character-structures that are in place in the pertinent situations, they will be operative not only inside the theatre but outside as well (Carroll, 2016, p. 86).

The assumption here is that the process of reading or watching a film or a play incorporates the audience's interpretation of what characters are doing and why, and in doing so, they rely on the knowledge of folk human psychology they employ in everyday life. Thus, encounters with narrative art are an exercise in explaining, understanding and predicting human behaviour, with the additional advantage in giving us opportunities to observe characters

and situations that might not be available to us in reality.<sup>14</sup>

Empirically based arguments in line with Carroll's suggestive response about the usefulness of narrative art have recently been developed by the growing amount of research in cognitive science and psychology. On the one hand, it seems that people generally use the same mental processes when they read fiction as they use when they read non-fiction.<sup>15</sup> This suggests that cognitive lessons associated with indirect humanism that are available from non-fiction, might be available through fiction – or, to put it differently, that the mode of presentation might not deter cognitive gain. On this basis, some believe that narrative art “is a set of simulations of selves and their interactions in a range of social circumstances” (Oatley, 2017, p. 261; see also Hogan's 2003, 2013 ) which is why “people who read a lot of fiction might have a better understanding of the social world” (Oatley, 2017, p. 269). The empirical evidence Oatley summons to support his claims are various. Research conducted by psychologists suggest that there are all sorts of valuable consequences that reading provides (a summary of which is given by James Young in his contribution to this volume), such as increased empathy. Philosophers and psychologists still debate how permanent such changes might be, but if empathy is conducive to our morality, then certainly Plato's dismissal of narrative art on the grounds of its usefulness was premature.

In contemporary discussions, the loudest voice against the cognitive value of literature is that of Jerome Stolnitz. Due to his commitment to formalism, a view according to which a work of art has no reference beyond itself and should be judged only by its formal properties, Stolnitz argues that the proper evaluation of a work of art should only concern the question of how artistic means are put together. His arguments against cognitive value mostly depend on the comparison between arts and sciences. He concludes:

there is no method of arriving at [the truth] in art and no confirmation

---

<sup>14</sup> See also Carroll 2013, Newman 2009, Vidmar and Pektor 2019.

<sup>15</sup> For details, see Metravers 2014 and the collection of papers in Troscienko and Burke (eds.) 2017.

or possibility of confirmation in art. Artistic truths, like the works of art that give rise to them, are discretely unrelated and therefore form no corpus either of belief or knowledge. Hence formal contradictions are tolerated effortlessly, if they are ever remarked. Only rarely does an artistic truth point to a genuine advance in knowledge. Artistic truths are, preponderantly, distinctly banal. Compared to science, above all, but also to history, religion, and garden variety knowing, artistic truth is a sport, stunted, hardly to be compared (Stolnitz, 1992, 2004, p. 342).

Stolnitz's arguments are only effective if literature is seen as separate from, yet rival to, the sciences. But that is not the only way in which we can discuss literature's connection to truth and knowledge – it is in fact a view that significantly misrepresents literature's cognitive dimension.<sup>16</sup> Arts are dedicated to human experience and human experience, unlike the physical world, cannot be described by a set number of concrete propositions or put into exact theories. As Rita Felski suggested (2008), literature is not a depository of timeless truths, as natural sciences seek to be, and is distinctive from the sciences such as history – an example often employed by Stolnitz – because it pays attention to particularities that sciences miss and neglect. To use her example, the particular subculture of New York's aristocracy depicted by Edith Wharton is not represented in historical books. Without her insight, contemporary readers might never get the chance to experience the “long-vanished America”, as she herself put it in her autobiography. Stolnitz's ultimate error is thus in not appreciating the kind of insights literature makes available as he misrepresented the domain that literature seeks to expose. In the next section we turn to this domain.

### 3. Narrative Art and Human Experience

Philosophers who defend art's cognitive value do so, roughly, by claiming that art and literature have their own domain of knowledge, different and independent from, but equally valid as the sciences. The sorts of insights that art delivers relate to *human concerns* and are thus independent from things that the sciences

---

16 See Vidmar 2019 for a developed response to Stolnitz.

generally care about, such as physical and biological aspects of the natural world. Many cognitivists, implicitly or not, presuppose that insights available through literature and art cannot be entirely embraced by the notion of knowledge, defined as justified true belief; rather, they call for recognizing other cognitive states, processes and mechanisms – such as understanding, change of perspective, realization, imagination, etc., as epistemically important, and they argue that literature and art are primarily efficient in generating such states, contributing to such processes or bringing about such changes.

The historical roots of such an approach to art are found in Aristotle's *Poetics*. His solution to the 'ancient quarrel' was to put art back on the pedestal, by claiming that

The poet's function is to describe, not the thing that happened, but the kind of thing that might happen, i.e. what is possible as being probable or necessary. The distinction between historian and poet is not in the one writing prose and the other writing verse – you might put the work of Herodotus into verse and it would still be a species of history; it consists really in this, that the one describes the thing that has been, and the other a kind of thing that might be. Hence poetry is more philosophical and of graver import than history, since its statements are of the nature of universals, whereas those of history are singulars. By a universal statement I mean one as to what such a *kind of man* will probably or necessarily say or do (*Poetics*, 2322-3).

Aristotle's claim that art deals with universals, with what humans can generally recognize as applicable to their experience, helps us circumvent the problem of finding the referents for the truth we extract from art. One version of this worry was formulated by Stolnitz, who asked whether "the statements of psychological truth refer to all or most or few of the flesh-and-blood beings they designate?" (Stolnitz, 1992, p. 339). Aristotle here suggests that we do not need to worry about that; rather, art is sufficiently universal to provide us with insights that can be applied to human experience generally.

Turning to contemporary discussions, one influential theory was developed by Dorothy Walsh. She claims:

we can recognize that literature is not adequately thought of as direct

report about something and that if there *is* a meaning of truth that is relevant to literature it must be different from the meaning of 'truth' associated with specifically formulated knowledge claims about this or that (Walsh, 1969, pp. 113-4, emphasis original).

On Walsh's view, art provides a distinctive kind of knowledge, one that is unavailable via scientific discourses. She first rejects the claim that literary works of art offer knowledge in propositional form: neither do readers extract such propositions from the work, nor do they test them against the real world to see if they are reliable.<sup>17</sup> The view of literary art delivering propositional knowledge is further complicated by the fact that different works generate different, sometimes contradictory propositions, all of which, by necessity, lack justification. If literary art were assessed on these grounds, it would be very hard to satisfyingly defend their cognitive value. Therefore, Walsh develops a different approach to defending such value, by arguing, first, that literature is a representation, that is, a mimetic image of the real world, which does not *refer* to the real world, and thus does not tell the truth about it, but resembles it. In light of such resemblance – captured by the notion of verisimilitude – we recognize literary works as true of the real world. Such resemblance – often experienced via the 'Ah yes effect!' that readers have – ties in with 'literary impressiveness', i.e. the manner in which a work exploits artistic resources available to literature. The cognitive power of a work is tied to its expressive, i.e. artistic and aesthetic power, which is how literature's way of being cognitively valuable differs from the scientific.

On Walsh's view, treating literary works as representational (or mimetic) images enables us to recognize their revelatory powers: literature makes the features of the world luminously displayed. The relevant features are those pertaining to human experience: "the general subject matter of all literary art is human experience" (Walsh, 1969, p. 80). For this reason we turn to literature for a better understanding of the world, which is why we are interested in numerous works. The experience that a literary work provides has

---

17 For an alternative view, see Novitz 1984 and Putnam 1976.

“permanent presence” and unlike real world experience, it is available for “full realization”. Therefore, “the kind of knowledge and the mode of knowing afforded by successful literary art seems to have an intimacy and an immediacy not characteristic of knowledge as knowledge about this or that” (Walsh, 1969, p. 138).

Walsh’ argument is premised upon her particular account of knowledge, and the relation between knowledge and experience. Influenced by John Dewey’s theory of experience, Walsh argued that, in addition to propositional knowledge, that is, *knowing that* (something is the case) and practical knowledge, i.e. *knowing how* (to do something), there is a third kind of knowledge, namely *knowing what is like*. As she sees it, “the kind of knowledge literature can afford is understanding as realization in the particular sense of the realization of what something might come to as a form of lived experience” (Walsh, 1969, p. 113). Such knowledge is characterized primarily by authenticity: authors present experiences for the readers, experiences that might otherwise go unnoticed.<sup>18</sup> If they are good at what they do, then these experiences are authentic and provide readers with the knowledge of how it feels to undergo such experiences.

More recently, a theory of the cognitive value of art and literature which sees them as complementary to the cognitive value of science was defended by James O. Young (2001). The main difference between art and science is that science employs semantic representations and rational demonstrations, while art, including literature, provides illustrative representation and illustrative demonstration and delivers knowledge not through the use of arguments and testimony like science, but by providing the audience with perspectives on the depicted matter. On Young’s view, the primary mode of representation in literature is illustrative representation, which is effective when the audience realizes the type or class of people or

---

<sup>18</sup> Here is Walsh: “Were it not for the testimony of literary arts, those of us who are not literary artists would never suppose, dream or imagine that the kind of complicated material that finds expression in literature was ever available for realization” (Walsh, 1969 p. 107).

events represented by the particular description.

The cognitive potential of literary works of art, claims Young, should not be explicated in terms of propositional theories: literary works do not express statements that are liable to assessment in terms of their truth. Rather, they provide illustrative demonstrations, i.e. open perspectives on objects so that the audience may achieve fuller understanding of them. In other words, while an artwork does not provide arguments to support its claims, it can afford an illustrative demonstration by enabling the audience to come to realize the correctness of a certain perspective. Various techniques are employed by the artists in order to enable such realization: the choice of objects and their various aspects (i.e. selection) guides the audience's attention, thus enabling them to recognize instances of the representation not available in other ways or simply ignored in the course of real life. An artist can modify certain aspects of the representations by enhancing them (amplification), by diminishing them (simplification), or by idealizing them, thus emphasizing some particular features of the represented thing. Novels, argues Young, often simultaneously juxtapose characters to make their different features more salient, or correlate them so as to underlie their relations. The purpose of all these techniques is to draw the audience's attention to the perspective employed by the work. Sometimes, the work does this via affective illustrations: those which evoke affective responses from the viewers. The particular value of such representations is in revealing what something feels like.

Part of Young's argument is the claim that art and science do not deliver the same kind of knowledge and do not conduct inquiry in the same way. Although science and art both rely on observation and interpretation, they use different methods. Scientists, who conduct researches in laboratories and in monitored conditions, rely on models and theories. Artists, who rely on observing people in their natural environment, rely on creating and providing perspectives, whose value is primarily in "giving us the capacity to discriminate features of complex phenomena and to navigate the problems posed by daily life." Therefore, Young claims, science is

better equipped to give us knowledge of the laws and natural phenomena such as global warming, while art is better equipped to teach us about emotional experience. For some aspects of our world – Young’s example is the French revolution – both art and science can be a valuable source of knowledge. However, unlike science, art “provides insight into complex, diverse subjects whose general laws are elusive or non-existent” such as human emotions, relations and predicament in the world. These domains of our experience cannot be ‘known’ in the scientific sense of the word, but can only be grasped if the appropriate perspective is taken on them.

A similar approach to the cognitive value of art is defended by Matthew Kieran (2005), who claims that “much art is aimed at prescribing and promoting, through the artistically manipulated conventions, particular ways of seeing the world” (Kieran, 2005, p. 102). Commenting works by Millet, Rodin, Degas, Constable etc, he claims:

The techniques, viewpoints, and aspects manipulated in all these artworks aim to promote imaginative understanding. (...) Their point is to evoke a particular imaginative understanding in relation to the subject portrayed and thus to deepen our imaginative understanding of our own world. (...) The artwork directs us toward the way certain things are to be seen and imaginatively understood, as opposed to merely stating that “they are or might be.” The work, its manipulation of conventions, style, and associations, prescribes particular imaginative experiences and, possibly, the reordering of our expectations (Kieran 1996, p. 343).

Emphasizing the capacity of a work to prescribe a certain perspective points to another capacity of literature overlooked by those who, like Stolnitz, ignore the multiple and distinctive cognitive ways of literature. Works of narrative art can function cognitively not by delivering true statements about the world, but by presenting some aspect of it under a particular perspective. Historically, the most famous example of a literary author who brought about a significant social change via the perspective she prescribed is Harriet Becher Stowe, whose *Uncle Tom Cabin* is still credited as one of the most powerful elements in the social processes that brought about the abolishment of slavery. Precisely how an artwork can function in this way is the topic of our next section.



#### 4. Literature, Humanism and Cognitivism

The claim that arts and literature have their own distinct area of concern not available to scientific inquiries is related to *literary humanism*, a view which states that literature deals with what is *humanly important*. Lamarque and Olsen relate such literary engagements to its mimetic aspect: “The interest which literature has for human beings, it has because it possesses a humanly interesting content, because what literature presents or says concerns readers as human beings” (Lamarque and Olsen, 1994, p. 265). On this conception, literature is distinct from the sciences because sciences are not concerned with human values and perspectives and do not raise questions regarding the meaning of our lives, principles we hold dear, relations we find worthy and attitudes we have toward things.

A defence of literature’s cognitive value against the background of literary humanism was offered by John Gibson. Literature, he claims, “is the textual form to which we turn when we want to read the story of our shared form of life: our moral and emotional, social and sexual—and so on for whatever aspects of life we think literature brings to view—*ways of being human*” (Gibson, 2007, p. 1). This ‘human’ aspect of what literature deals with is important to cognitivism because it explains why we turn to literature with the expectations of gaining understanding about reality and why we feel that “in literary experience we often come to know ourselves and our world better (Gibson, 2007, p.1).<sup>19</sup> Humanism “marks in literary aesthetics a very modest proposal: that there is an important link between literature and life, and that this link, whatever it may precisely consist in, accounts for one of the central reasons we value literature” (Gibson, 2007, pp. 15-16). This link between literature

---

19 As Gibson further claims, “... literature offers us a window on our world. We might call this the humanist intuition and characterize it as the thought – or hope – that literature presents the reader with an intimate and intellectually significant engagement with social and cultural reality” (Gibson, 2007a, p. 2). See also his “Introduction” to *A Sense of the World* (Gibson, Huemer and Pocci eds. 2007).

and life means that

we have grounds for claiming that part of the project of many (though certainly not all) literary works is to articulate an insight into some specific region of human experience and circumstances (...) The humanist wants to assert that through works of literature the significance of very real human experiences, practices, and institutions can be revealed when they were once mysterious or obscure; that a grasp of reality can be gained from close reading (hence literary *humanism*, for the claim is that literature speaks to human reality). The humanist means nothing metaphysical, implies nothing foundationalist, when he speaks of 'reality'. He gestures only toward the everyday world we inhabit – that is, the world of actual human experience and action (Gibson, 2007, p. 16).

Some scholars see literary humanism as primarily concerned with the question 'how to live', or 'what forms of life are valuable'. These scholars see the value of literature in its capacity to answer this question, usually by showing different ways in which life can be worth living. Martha Nussbaum and Philip Kitcher have both written extensively on this. Nussbaum argues that the question 'how should one live' is an empirical question that literature can answer because it provides us with "the patterns of possibility – of choice, and circumstance, and the interaction between choice and circumstance – that turn up in human lives with such a persistence that they must be regarded as our possibilities" (Nussbaum, 2010, p. 244). Given that literature presents these options to us, it can foster a deeper understanding of ourselves and our social circumstances and thus show us the possibilities that lie ahead for us.

#### **4.1. Direct and Indirect Humanism and Literary Cognitivism**

Given the variety of cognitive benefits available through engagements with narrative art, it is not unusual for philosophers to make a distinction between so called direct and indirect humanism. As Gibson explains, a direct humanist needs to show that the link between literature and reality is immediate, in the sense that reality is (directly) contained *within the words in a literary work*. Reality (represented in a work) is not supposed to be something external to the work, something that the work refers to, but is "something we come into contact with when we explore the interior of the

work” (Gibson, 2009, p. 472).

The proper conception of (direct) humanism Gibson defends, and the one he sees as the only plausible account of literary cognitivism, can be spelled out in terms of the following four theses:

- i) Part of the project of at least some literary works is to articulate an insight into some specific region of human experience and circumstance
- ii) Through literary works the significance of very real human experiences, practices and institutions can be revealed
- iii) A grasp of reality can be gained from close reading: literature speaks to human reality, to the world of actual human experience and action
- iv) Literary experience may be a direct appreciation of and engagement with the real world

(i) to (iv) share the assumption that it is the literary works themselves, not the readers, that ‘do’ the cognitive work. For example, it is often said that Faulkner’s novels provide a great picture of the American South, including being particularly revealing of the connections between racism, slavery and identity.<sup>20</sup> Consider an array of critical commentary on William Faulkner’s works:

These extraordinary stories embody the essence – the people, the atmospheres, heat and fermenting tension – of America’s Deep South. Themes of the hunt, violence, friction between black and white, the primal corrupting influence of women, the past living in the present, surface again and again, handled with Faulkner’s insider’s genius for

---

20 The claim is not that we need to read Faulkner in order to get the information that the American South was for a part of its history saturated with racial issues, as this would amount to an instrumentalization of literature. The claim is rather that certain benefits are available to those who read. As Gibson explains in commenting on Faulkner’s *Intruder in the Dust*: „The structure of this story of racial injustice is hardly original: a black man is wrongly accused of murdering a white woman, and a lawyer who reluctantly comes to believe in his innocence fights against a community set on lynching him. We’ve heard this story (or at least seen the movie) before, in one form or another. Faulkner’s accomplishment was not to construct a terribly original story but to tell a story in a particular way, a way that rendered intelligible how certain features of Southern culture give rise to these familiar, intractable problems of race” (Gibson, 2011).

re-creating an intensely human world... (Malcolm Cowley, *in* The Penguin Collected stories of William Faulkner).

To some readers, it [*Absalom, Absalom!*] perfectly illustrates southern history. To some, it explains American race relations (Towner, 2008, p. 40).

The first three chapters of *Go Down Moses* very carefully trace (...) patterns of racial behaviour and their relationship to individual identity (Towner, 2008, p. 58).

These critical comments stress precisely those aspects of literature that direct humanism insists on: the real world as contained within the covers of a novel. Notice how each comment relates to the inherent bond between art and experience, a bond which Walsh and Young claim is the main focus of literature. These comments shed further light on Young's claim about literary works representing types of people. Of course, the epistemologically interesting question concerns the justification of those representations of the world contained in the text itself. In some cases, as with the realist novel, the problem of reliability is surmounted once we recognize that the artistic norms of realism coincide with epistemic norms of reliability. On this conception, the literary writer serves as an informer on how society is and what it does.<sup>21</sup> Consider a critical commentary of William Dean Howells's novel *A Hazard of New Fortunes*:

Howells paints a panoramic portrait of urban life. His novel abound in richly detailed descriptions of people representing the socio-economic spectrum, including recent immigrants, transplanted Southerners, old money and the newly rich, artists and writers. The points of view expressed by these characters include a property-is-theft socialism, a conservative Gospel of Wealth capitalism, and a remnant of the Old South's feudal aristocratic perspective. The crisis of Howells's novel, a bloody riot, reflects the harsh inequities of capitalism in the late nineteenth century and the class conflict simmering just below the surface of New York society... (Crane, 2007, p. 161).

Against the background of this conception of realism, it is not an exaggeration to claim that realist novels can be the source of factual knowledge. Consider, as an example, Dreiser's meticulous description of the development of Chicago in *Sister Carrie*:

---

<sup>21</sup> See Carroll 2007 for a defense of the cognitive value of realism.

In 1889 Chicago had the peculiar qualifications of growth which made such adventuresome pilgrimages even on the part of young girls plausible. Its many and growing commercial opportunities gave it widespread fame, which made of it a giant magnet, drawing to itself, from all quarters, the hopeful and the hopeless--those who had their fortune yet to make and those whose fortunes and affairs had reached a disastrous climax elsewhere. It was a city of over 500,000, with the ambition, the daring, the activity of a metropolis of a million. Its streets and houses were already scattered over an area of seventy-five square miles. Its population was not so much thriving upon established commerce as upon the industries which prepared for the arrival of others (Dreiser, 1981, p. 15).

Works of narrative art often introduce fictional elements and distort the descriptions of real world, usually for artistic and aesthetic purposes, or purposes of dramatization and other affective effects. Such distortions are part of literary practice, which is why we generally make a distinction between fictional and factual. However, we might wonder if such distortions are damaging to our intuition that we can learn something about the world and humans from art. For all the praise attributed to Faulkner's insight into the racist culture, his novels take place in Yoknapatawpha County, which, although modeled on and inspired by a Mississippi county, is not, nor it was back in Faulkner's days, real. Can we still claim that his novels reveal something of great significance about racism? Young's analysis of descriptive representations can help us answer this question, as can Walsh's notion of resemblance and the principle of verisimilitude. Given that readers can easily recognize the real world reflected in the novels, they can form new beliefs and come to understand something about those aspects that the novel puts to view. This is because

the literary perspective (and the artistic perspective more generally) is the definitive human perspective: the standpoint from which we are best able to bring to light the range of values, desires, frustrations, experiences, and practices that define the human situation. On this view, works of literature, at least when they live up to their promise, represent cognitive achievements: they embody ways of knowing the world (Gibson, 2009, p. 467).

A more relevant reason for maintaining that literary works are resources of cognitive values is the fact that there is much more

than propositional statements and factual truths that readers pick up from works. As most literary cognitivists argue, because of the way reality is presented in the work, an active and reflective engagement with it not only reveals something factual about the world but also modifies the audience's cognitive, emotional and imaginative economy and it enhances her conceptual framework. These are important cognitive achievements, because they make one, on the whole, a better, more sensitive, more discriminatory, more reflective cognizer.<sup>22</sup> Reading about the racial relations in Faulkner, the argument goes, can have various sorts of impact on the reader. It can help one to better understand causes and consequences of such relations, it can make one more attentive to various kinds of social and political actions giving rise to it, more sensitive (on a positive reading) to one's own treatment of others, and thus more appreciative of the nuances of the human predicament. Because a reader can imaginatively experience the sufferings of victims of racism, she may develop a higher sensitivity towards people of different races in her everyday conduct and become less prone to act as a racist.<sup>23</sup> Recognizing the struggling of people like Joe Christmas, one might be induced to reconsider one's own origin and to question the extent to which each of us is predetermined to certain kinds of behaviour. In such cases, the relevant processes one undergoes, the attitudes one thus develops or the perspectives one comes to recognize as correct or flawed, are not contained within the work, but are triggered by what the work presents. Thus, they take place in the reader, as she goes through the work and in the course of her engagement with it comes to think of the world in a more informed way.

As Gibson explains, in such engagements with the work an attempt is made to "bring literature to bear on the our-worldly by exploring our ability to *apply* aspects of the content of a literary work to extra-textual reality. Thus the reader builds the bridge between fiction and reality and so unites what the work itself cannot" (Gib-

---

22 See Kvanvig 2003.

23 See Carroll 1998, 2003; Baccharini 2010, 2018.

son, 2007, p. 18). By dealing with literature, we somehow change the way we think about the world. This kind of cognitive gain is not propositional. Rather, literature offers us raw material out of which we can build new ways of understanding our world, either by making us question our values and concepts or by teaching us to apply these concepts in new circumstances we encounter in real life. Indirect humanism powerfully defends the view that literature influences our cognitive, imaginative and emotional economy in a way that makes us approach the world differently, with enriched concepts. In a nutshell,

literature might help improve our faculty of imagination, develop our cognitive skills, discover what we would think, feel, or value if in another's shoes, become more sympathetic and adept moral reasoners, and so on. These are genuine cognitive achievements, and literature can certainly help us in our pursuit of them (Gibson, 2007a, pp. 23-4).<sup>24</sup>

To explain precisely how art can have such effects, the following claims are put forward:

- i) Literary works can invite modes of reflection, simulation, and imagination that can enable us to better understand our world;
- ii) Literature can illustrate possibilities and offer possible ways for organizing and conceiving experience;
- iii) Literature offers conceptions, stances and perspectives which we can, by using our reflective and imaginative capacities, transform into a tool for appreciating reality;
- iv) Literature can suggest ways of 'reading' the world, presenting us with new possibilities of wordly understanding and involvement.

---

<sup>24</sup> It should be pointed out though, that Gibson rejects indirect humanism as valuable from a cognitivist position, because, on his view, it focuses too much on the reader and his overall cognitive economy and too little on the literary works, thus sacrificing the humanistic aspect and the aesthetic accomplishments of the work itself. Given that our perspective here is epistemological – i.e. we are interested in exploring how engagements with art benefit readers' cognitive capacities – we can neglect Gibson's worry.

One common aspect of i) to iv) is that these cognitive benefits cannot be reduced to the acquisition of propositional knowledge. On the contrary, such achievements are only possible because the readers are already familiar with the relevant facts. If a reader is to have her conception (of, say, jealousy) transformed (i.e. in order for her to gain a deeper understanding or a new perspective) when contemplating about Shakespeare's portrayal of jealousy in *Othello*, she needs to know what jealousy is and how it feels. If that is the case though, then fictional representations "do not function epistemologically, to inform us of how things in the actual world are" (Gibson, 2007b, p. 5). An anti-cognitivist can easily turn this into a conclusive argument against literary cognitivism generally: literature does not inform us about the world, because we already need to have the relevant knowledge of the world to engage with literature. However, this would be a knock-down argument only if the cognitive value of literature and art were dependent on acquisition of new knowledge, and, on my view at least, that is not the case. Cognitive value is derived from the overall enrichment and transformation of our cognitive powers that narrative art can induce. As Frank B. Farrell claims,

literature at its best should be seen as providing complex experiences that transform the cognitive apparatus of the reader so as to adapt it for an increased sensitivity to specific features of the world. (...) Literature does not so much present us with asserted truths as make us better at tracking the truth relevant features of the world (Farrell, 2007, pp. 246-7).

In line with such ideas, David Novitz insists on recognizing multiple ways in which art can be cognitively valuable; consequently, his theory makes room for direct and indirect benefits (though he does not employ that terminology). As he sees it, there is nothing wrong with the claim that art can be a source of true propositions, provided the descriptions containing such propositions are accurate. However, he claims, the truly valuable cognitive aspect of literature is in that it can enable readers to develop various kinds of skills, practical as well as intellectual. As he explains, works "may impart intellectual strategies by enabling us to take more aspects



of a problem into account, and in this way to think more comprehensively and efficiently about it.” Artworks, he claims, “extend our thinking by drawing to our attention previously unconsidered aspects of a problem” or by offering “radically new ways of thinking about or perceiving aspects of our environment” (Novitz, 2008, p. 345).

Cognitive benefits of engaging with works of narrative art associated with indirect humanism insist on bringing together knowledge, beliefs, attitudes and opinions that readers possess, and the ‘material’ that the work provides. Some authors claim that because of such merging, narrative art works with the knowledge we already have and produces new instances of it (as we’ll see below). Peter Lamarque refers to such instances as cognitive strengthening: “Again the emphasis is away from the acquisition of newly found worldly truths towards ‘clarificationism’ (Noel Carroll), or an ‘enriched understanding’ (Gordon Graham) or an ‘acknowledgment’ (John Gibson) of beliefs readers are likely to hold already” (Lamarque, 2010, p. 381).

Graham’s notion of enriched understanding is particularly popular among literary cognitivists. As an epistemic state, understanding has only recently been recognized as valuable and we are yet to come up with an agreed-upon theory of what it consists in. Consider Wayne Riggs and Jonathan Kvanvig’s accounts:

what is involved in having understanding may well be even more obscure than what is involved in having knowledge. But it seems clear enough that it includes having a true grasp of some significant part of reality without being deeply deceived about it. ... Understanding some part of the world requires an appreciation for order, fit, and pattern. It requires that one ‘see’ how things fit together and why they are the way they are (Riggs, 2003, p. 35)

The central feature of understanding, it seems to me, is in the neighbourhood of what internalist coherence theories say about justification. Understanding requires the grasping of explanatory and other coherence-making relationships in a large and comprehensive body of information. One can know many unrelated pieces of information but understanding is achieved only when informational items are pieced together by the subject in question (...) Whereas knowledge can have as its object individual propositions, understanding may not (Kvanvig,

2003, p. 192).<sup>25</sup>

Many literary cognitivists share the idea that narrative art is well equipped to show ‘how things fit together’, that is, to unite unrelated pieces of information so that the audience comes to understand a particular process, emotion, motivation for a particular kind of action, etc. For example, given Flaubert’s depiction of social circumstances in which Emma and Charles live, and his portrayal of Emma’s mental states, a reader is brought into a position from which she can understand those impulses which induce one to act in a particularly disruptive way characteristic of Emma.

Elvio Baccarini (2010, 2018) wrote extensively on art’s capacity to foster our understanding in the context of our moral lives and actions. An aspect of understanding he stresses relates to the abilities that the cognizer has once he understands a certain moral principle. This includes grasping the reasons behind the principle. Without such grasping, Baccarini claims, a moral agent cannot be responsive to moral reasons, as she cannot make the proper relations between various moral principles, or recognize the instances covered by the normative requirement and see the relevant domain of the application of a particular principle. Art is particularly important in providing experiences that enable the agent to come to fully grasp the significance and application of moral principles.

Carroll’s notion of *clarification* refers mostly to art’s capacity to influence our ethical sensibility. According to Carroll,

Clarificationism does not claim that, in the standard case, we acquire interesting, new propositional knowledge from artworks, but rather that the artworks in question can deepen our moral understanding by, among other things, encouraging us to apply our moral knowledge and emotions to specific cases. (...) In the course of engaging a given narrative we may need to reorganize the hierarchical orderings of our moral categories and premises, or to reinterpret those categories and premises in the light of new paradigm instances and hard cases, or to reclassify barely acknowledged phenomena afresh – something we might be provoked to do by a feminist author who is able to show up injustice where before all we saw was culture as usual (Carroll, 1998 p. 142).

<sup>25</sup> See also Elgin 1993, 1996; Mišćević 2012; Vidmar 2017b.

John Gibson (2003, 2007) and Rita Felski (2008) use a distinction between knowledge and acknowledgement developed by Stanley Cavell and argue that literature enables the reader to acknowledge those facets of the world and human situation not captured by knowledge. Gibson's example is that of a person who knows someone is in pain but fails to call an ambulance or do anything else to provide help. According to him, such a person

succeeds in every case of knowledge, for he consistently reveals that he knows the 'truth' of the matter (that you are suffering, that this implies that you require aid, that the consequences are severe should we ignore this, and so on). But he has no further relation to your pain beyond his knowing it, beyond his ability to identify correctly your suffering and the rest of propositions this entails. In this sense we see that his knowledge is idle, lifeless, for his mind goes dead precisely when it ought to become animated. What we see (...) is a failure to grasp what we might call *demands* of knowledge, the claims knowledge makes on us (Gibson, 2003, p. 232).

This is where literature is at its strongest: it acknowledges the knowledge we bring to the text. For Gibson, knowing is not the highest cognitive achievement; the gap between the mind and reality is not closed by knowledge. This is why acknowledgement is important:

The concept of acknowledgment reveals the possibility of a residual gap; it shows us that the concept of knowledge alone does not express understanding as it reaches all the way into the world. And the claim the humanist wants to secure is that it is this remaining divide that literature is capable of addressing and overcoming (Gibson, 2003, p. 236).

Without such acknowledgment, Gibson claims, our knowledge is always idle and removed from the world.

Some philosophers, such as Peter Lamarque, are not happy to embrace the effects of cognitive strengthening. On the one hand – as Britt Harrison argues in her contribution to the volume – such pluralism of cognitive values attributed to narrative art says more about the conceptual confusion underlying our talk of cognition and cognitivism than about ways in which narrative art is capable of assisting cognitive growth. On the other hand, some have argued that the cognitive gains we get through art experiences necessarily lack justification, as literature offers none. Stein Olsen (1978, 1987) for example excludes literature from the domain of informative

discourses, not for its lack of informative power but for the inappropriateness of treating literature as such.<sup>26</sup> As he sees it, treating literature as a means to ends (i.e. enhanced cognition or development of moral sensibility) is a form of instrumentalization.

As a way of countering such claims, Philip Kitcher argues that there is no substantial (i.e. epistemic) difference from the perspective we develop as a result of reading, and perspectives we obtain in the course of our lives, which develop as a result of being a part of certain communities and social groups. The epistemic justification of the concepts developed as a result of engaging with literature and living in a particular time and place is the same: “it is not that we achieved our concepts and categories through some insight into their special worthiness – there was no Cartesian moment at which they were rigorously assessed and found to pass muster”, claims Kitcher. The concepts through which we perceive the world and make sense of our experience are acquired from our culture and as we go along; we change them and adjust them so that they make a more-or-less coherent whole. In that sense, literature is just one more tool available to us, but a tool that can have particularly powerful influence on us. How exactly does this happen? As Kitcher explains,

In reading a work of fiction or a poem, or in listening to a piece of music, we pass through a sequence of psychological states, partly shaped by our antecedent judgments, conceptions and emotions, partly the product of our apprehension of the words or the sounds. We imagine the actions and situations described in words, we identify the emotions and moods expressed in the music. The occurrence of these states sets up connections with other parts of our psychological lives, recalling

---

<sup>26</sup> Here s Olsen: “Literary discourse and informative discourse are two mutually exclusive classes. However, the thesis does not imply that one cannot at different points in time interpret the same piece of discourse as on one occasion literary and on another occasion informative. It is possible to change one’s point of view from an aesthetic one to one where piece of discourse is seen as informative (and to change back again at will). What is impossible is to see the informative function as being a part of the literary function. It is a category mistake to let judgments about the truth of a piece of discourse interfere with one’s aesthetic understanding or evaluation of it.” (Olsen, 1978, p. 58). See also Attridge 2015.

past judgments or emotions, sometimes modifying our established ways of conceiving and evaluating. The result is what I shall call a *synthetic complex*, whose elements may be radically disparate: memories of our own experiences, images from earlier perceptions or encounters with other works of art, judgments previously endorsed or rejected, emotions now excited by different objects, or even emotions of types we have not previously felt. The power of some works of literature and music to build synthetic complexes accounts for their enduring hold on us – as we return to them, again and again, the synthetic complexes they generate grow and change, perhaps expanding into areas of our psychological lives that were initially quite remote from their influence, so that we come to think of the pertinent works as inexhaustible (Kitcher, 2013, p.146).

An important aspect of a synthetic complex is that it becomes integrated into reader's internal cognitive economy and it mixes with all the other elements of it, becoming a part of the reader's coherent system of beliefs.<sup>27</sup> This gives additional strength to the cognitive benefits available in literature, by internalizing them and making them a stable part of reader's internal cognitive economy, they become (or have the potential to become) what internalists/evidentialists such as Matthias Steup, Richard Feldman and Earl Conee call justifiers: they serve as factors that can justify the acceptance of some future proposition or commitment.

The formation of synthetic complexes, when they persist as stable parts of our thinking and feeling, can revise our conceptions and judgments. Of particular concern are *endorsements* and *rejections*, judgments in which a subject concludes that some state of affairs is tolerable or to be resisted, or in which she takes a scenario as a serious possibility for herself, a goal to be worthy of pursuit, a course of action she has hitherto viewed as necessary to be trivial and dispensable (Kitcher, 2013, p. 147).

Kitcher's examples of how literary works can permanently in-

---

<sup>27</sup> In Kitcher's words, „Responsible building of such complexes should be *reflectively stable*: that is, as the reader or listener ponders the connections she makes in light of the full range of her antecedent attitudes and commitments, she should discover that the complex is sustainable. The reader of *Bleak House* jettisons some old convictions, but the synthetic complex that displaces them accords with quite general and fundamental commitments to avoid wishful thinking and to suspend judgment about what has been casually taken for granted, once it is clear that it can be called into question.”

form a reader's understanding or her perspective on the world are novels by Thomas Mann and James Joyce. Because Kitcher sees literature as addressing the question of the value of life, he believes that art can enable the audience to reach a point from which they contemplate their own lives from an enriched perspective. As he explains, a proper engagement with literature is to

bring (...) readers to a previously unanticipated perspective, a different Gestalt on life and on the factors that make a difference to its mattering. We envisage a process in which people are brought to see or hear or think or feel in novel ways, so that questions that had been viewed as unanswerable admit of solution (Kitcher, 2013, p. 146).

On Kitcher's view, literature has the capacity to make one change one's view on how the world is. He writes,

Through the swirling dream of Joyce's last work, readers are brought, again and again, to rejoice in the everyday, to laugh at its comic mistakes and misunderstandings, and, finally, to recognize the possibility that even flawed relationships may center lives of real value (Kitcher, 2007, p. 26).

Similarly, with respect to *Ulysses*, he says:

What makes *Ulysses* one of the greatest novels in the English language (...) is that the reconstructed thoughts of Bloom, of Stephen, and of Molly are *worth* following, showing us what it is to struggle, to aspire, to fail, to fall, to betray and be betrayed, to befriend, to forgive, showing us some of what human life is, how it is limited and confused, how it can be triumphant and worthwhile (Kitcher, 2007, p. 49).

Another coherentist account of how the impact of narrative art on the audience's cognitive (and moral) agency can be justified was provided by Elvio Baccarini in several of his papers. The gist of Baccarini's account is the claim that the input we receive from art is but one element in the processes we use in our contemplation about the world, that is, in the process of reflective equilibrium. As he sees it,

Artworks contribute better to the improvement of moral knowledge as part of gradual and reflective mutual adjustment and clarification of beliefs. In such a procedure, the moral epistemological role of experience of artworks is part of a wide reflective equilibrium and of the process of refinement of our understanding of general moral principles, in particular of the range of their application (Baccarini, 2010, p. 20).

Baccarini's work is above all insightful in bringing together art and the human moral domain. Let us now turn to a more detailed exploration of how the two interact.

## 5. Narrative Art and Ethics

Fiction is often at its most inciting, and instructs us most adequately in matters of value, when it explores moral problems and brings its readers to see them in their fullness and complexity (Novitz, 2008, p. 355).

The underlying premise in Novitz's claim is that narrative art is well equipped to help us navigate the complex issues of human morality. In fact, it is not an exaggeration to say that the central topic of much narrative art consists of moral issues and human relations and interactions that fall under domain of ethics.<sup>28</sup> It is no surprise then that many who share cognitivist intuitions see narrative art as primarily conducive to the development of moral sensibility. Berys Gaut (2007) identifies the birth of literary humanism as a response to Plato's denial of the ethical powers of mimetic art, and John Gibson ties in his humanism with the ethical content found in literature:

That literary content is often a kind of ethical content seems in one sense obvious. After all, if literary works concern themselves with the stuff of human experience, they could as much ignore the ethical as they could the psychological, familial, social, or political dimensions of life (Gibson, 2011).

The question before us now is to consider whether narrative art, given its concern with the ethical, can teach us something about those ethical issues it puts in front of us, ignoring, again, the question of how such a capacity would contribute to its aesthetic value. Above, we mentioned philosophers like Baccarini and Carroll who put a lot of faith into literature's capacity to help us reach moral understanding, i.e. understanding of what moral concepts demand of us. We also saw how more empirically oriented scholars, like Keith Oatley, credit art with expanding one's range of moral

---

<sup>28</sup> For the purposes of this essay, I will use ethics and morality, and their adjectives, as synonyms.

reactions, in light of its simulating social reality. Philosophers are, for the most part, optimistic regarding art's influence on morality and see it as particularly well equipped to assist us in our everyday moral dealings. This optimism is mostly grounded on the complexities of moral issues, when these are depicted in a widely elaborate narrative. Such complexity often eludes us in our everyday lives, because we are asked to act, not to reason, and to make choices, which are often clouded by the personal stakes we have in any given situation. No such stakes exist in our narrative engagements, which invite reflection rather than action. Consequently, narrative art gives us a chance to see firsthand how morally complex and demanding our (social) reality is.<sup>29</sup>

On the other hand, some philosophers, most famously Plato, are more cautious, claiming that art can be detrimental to our moral sensibility and behaviour, given that most art presents immoral characters which might appear appealing. It is an interesting fact that Plato fears imitative poetry primarily because he recognizes those of its influences we characterized as indirect: its capacity, in other words, to influence our cognitive economy and emotional capacities. On his view however, all such influences are by necessity dangerous, not only because art, unlike philosophy, is not dedicated to finding the truth, but because art stirs emotions which distract us from rational enquiries.

“The most serious charge against imitation is that it is able to corrupt even decent people” (*Republic*, 605c), claims Plato. Underneath this claim is his understanding of the human soul and overall human cognitive functioning, as well as the fact that imitative poetry was generally considered pleasing and that people by nature enjoy imitation. This is dangerous for morality (moral education and moral behaviour) because poets “have produced a picture of the gods that is contrary to sound moral views, and they also portray wicked human beings in a manner so attractive as to make us sympathize with them” (Rosen 2005, p. 354). Such portrayals, Plato

---

29 See in particular papers collected in Hagberg ed. 2011, 2016; Levinson 1998; Palmer 1992; Schellekens 2007.



argues, might prompt us to become such immoral creatures and develop moral vices. Consider, as an example, the allure of Emma's infidelity:

Never had her eyes been so large, so black, of so profound a depth. Something subtle about her being transfigured her. She repeated, 'I have a lover! a lover!' delighting at the idea as if a second puberty had come to her. So at last she was to know those joys of love, that fever of happiness of which she had despaired! She was entering upon marvels where all would be passion, ecstasy, delirium. An azure infinity encompassed her, the heights of sentiment sparkled under her thought, and ordinary existence appeared only afar off, down below in the shade, through the interspaces of these heights (Flaubert, 1993, pp. 101-2).

The issue, on Plato's view, is that Emma's adultery – an act in itself immoral and deplorable – is here presented in a way which might encourage someone to follow Emma's footsteps. Rather than shame and remorse, Emma feels happiness and excitement, and instead of a cruel punishment for her breach of social and marital duties, she is rewarded with an even greater physical beauty. A reader, perhaps herself unhappily married, can easily identify with Emma's loneliness. She might then succumb to the passions in the arms of her lover, having learnt from Emma a trick or two on how to hide one's infidelity from one's spouse – as Novitz reminds us, there are all kinds of practical strategies available from fiction. To Plato, that is a reason enough to sanction art. It is above all important that art only be allowed if its content and style are carefully scrutinized so as to avoid inappropriate elements. This is Plato's main argument in favour of censorship.

Plato further argues that imitative poetry influences those parts of humans which are opposite to reason, namely emotions. He claims, "imitation really consorts with a part of us that is far from reason, and the result of their being friends and companions is neither sound nor true" (*Republic*, 603c).<sup>30</sup> This is why imitative

---

<sup>30</sup> Janaway argues that it is because of this aspect of artistic influence (its capacity to move the irrational parts of human beings) that poetry should be banned from the perfect state. It is not enough to eliminate the arts and poetry from education in order to protect children who are too young to resist the charms of poetry; grownups also need protection because poetry stirs emotions and

art is particularly dangerous: it is produced by those who do not have knowledge but only deceive the audience; it plays upon those parts of the soul which are most easily fooled into believing what is not; it can easily trigger people to behave inappropriately and it stirs emotions. For these reasons, he banishes art from the perfect state,<sup>31</sup> proclaiming that “imitation is an inferior thing that consorts with another inferior thing to produce an inferior offspring” (*Republic*, 603b).

The crucial psychological mechanism operative in the experience of art (whether on the part of a poet or an audience) is what Plato calls impersonation. The idea is that a poet impersonates other characters which impacts the audience in that it can make them prone to imitating such characters. This is dangerous in two ways: in the ethical domain, it can lead to impersonation of improper (immoral and emotional) people. In the political domain, it can lead to the breaching of what he calls the ‘principle of social specialization’, according to which one person should only perform one job and have specialities necessary for doing it well.<sup>32</sup> Plato repeatedly insists that impersonating many other people, their virtues and vices, can be detrimental for the moral and political education of the youth.<sup>33</sup> Thus, young people should not

---

touches irrational parts of human souls (see Janaway, 2006, p. 391).

31 The exception being hymns to gods and dithyrambs. Rosen (2005, ch.13) claimed that Plato’s arguments are only directed against mimetic poetry, not against narrative poetry, which is allowed into the *Republic* as long as it conforms to his strict paternalistic rules. He also claims that this aspect of Plato’s theory (i.e. the allowance of ‘suitable poetry’) ultimately shows “that even the rule of philosophers cannot make do entirely without poetry” (p.353).

32 This principle is developed in Book II of *Republic*. See Halliwell (2002, p. 51) who uses the term ‘social specialization’, and Rosen who talks about „political principle of one man, one job” (2005, p. 357). Plato appeals to this principle again in *Ion*, in order to show that poets cannot have the relevant knowledge of the things they write about because, given that they have the knowledge of poetry, they cannot in addition have another set of expertise (say about medicine) that would enable them to write knowledgeably about medicine.

33 “Then, if we’re to preserve our first argument, that our guardians must be kept away from all other crafts so as to be craftsmen of the city’s freedom, and be exclusively that, and do nothing at all except what contributes to it, they must neither do nor imitate anything else. If they do imitate, they must imitate from

hear it said that in committing the worst crimes he's doing nothing out of the ordinary, or that if he inflicts every kind of punishment on an unjust father, he's only doing the same as the first and greatest of the gods (378b)...

Indeed, if we want the guardians of our city to think that it's shameful to be easily provoked into hating one another, we mustn't allow *any* stories about gods warring, fighting, or plotting against one another, for they aren't true ... If we're to persuade our people that no citizen has ever hated another and that it's impious to do so, then *that's* what should be told to children from the beginning by old men and women (378c)...

We won't allow poets to say that the punished are made wretched and that it was god who make them so. But we will allow them to say that bad people are wretched because they are in need of punishment and that, in paying the penalty, they are benefited by the gods (*Republic* 380b).

Not only does Plato, in Books II and III of *Republic*, dispel stories he considers dangerous, he also provides a list of those that should be told, as he sees them to be conducive of good behaviour (*Republic*, II, 379a-383c). In Book III, he specifies rules that bear directly on how people (should or shouldn't) feel and behave; Plato's aim here is to raise brave men, who are not afraid to fight and to die if necessary. He wants his citizens to be self-composed and to refrain from publicly expressing their emotions. Thus, children should be told stories "that will make them least afraid of death" (386a), that praise life in Hades (386b), that will make people "fear slavery more than death" (387b), that do not contain "lamentations and pitiful speeches" (387c). Particularly worrisome are stories that say that "many unjust people are happy and many just one wretched, that injustice is profitable if it escapes detection, and that justice is another's good but one's own loss" (392b). Stories should be prohibited which could inspire one to imitate "either a young woman or an older one, or one abusing her husband, quarrelling with gods, or bragging because she thinks herself happy, or one suffering

---

childhood what is appropriate for them, namely, people who are courageous, self-controlled, pious, and free, and their actions. They mustn't be clever at doing or imitating slavish or shameful actions, lest from enjoying the imitation, they come to enjoy the reality" (*Republic*, 3, 395,c).

misfortune and possessed by sorrows and lamentations, and even less one who is ill, in love, or in labour” (395e), bad men, cowards, those who ridicule one another, use shameful language, who wrong themselves in any way, who are mad in any way, those connected to craftsmen, those who indulge in sexual or other kinds of pleasurable behaviour, those who laugh, money-lovers, those who use bribery. The reason why all such stories are dangerous is that they seem interesting and alluring and it is easier for the young to be attracted to them than to those which praise virtuous people.

In contemporary discussions, Plato’s worries are predominantly directed towards art (or art-like forms, like videogames) which show particularly violent and sexual behaviour, such as pornography. As for the other ‘inappropriate’ content, philosophers, educators and psychologists are not as strict as Plato was. Some preach caution about the possibility of moral corruption (Goldman, 2013; Young, this volume), while some focus on art’s positive influence (Baccarini 2010, 2018; Nussbaum 1986, 1990, 2010; John 2010; Diamond 2010). Certainly one of the most significant changes in theorizing about art after Plato, evident already in Aristotle, is the overall dismissal of the negative view of the emotions. Aristotle places much value in art’s capacity to stir our emotions, seeing that process as one which enables the viewers to purify themselves, but also to become aware of “our own capacity for bad choice and vulnerability to reversal of fortune” (Barfield, 2011, p. 38). As Barfield further explains, when watching a tragic play, “we can actually experience fear and pity associated with great fault on the part of a person not so different from ourselves. In this way we are made aware of capacities within us that we might not otherwise acknowledge...” (Barfield, 2011, p. 38). More importantly, Aristotle sees poetry and tragedy as revealing “the ethical structure of events to which we are vulnerable” (Barfield, 2011, p. 34) which enables us to “become more immediately aware of possibility that exists and that bears on our perceptions of vulnerability and value” (Barfield, 2011, p. 34).

Another element in Plato’s theory that has been rejected in contemporary discussions relates to the idea of imitation and im-

personation. Philosophers recognize a much wider sphere of emotional reactions to art, where such reactions rarely involve imitation. The audience does not identify with the characters, i.e. they do not start believing or behaving as if they were any particular character. Rather, they develop positive or negative attitudes towards characters, i.e. sympathy or antipathy, where the development of positive (or negative) attitudes is often entwined with positive (or negative) moral assessment of characters' personae. In some cases they develop empathy towards the character, which, as some claim, is particularly relevant for understanding human beings in the real world (Gibson 2016, Hagberg ed. 2016).

While the human capacity to extend our emotional reactions to fictional entities has been confusing at least since Hume's essay *On Tragedy*, many see this as essential for art's contribution to our moral refinement. In particular, Martha Nussbaum insists on the importance of emotions in our moral reasoning, and for literature's capacity to fully expose the internal link between emotions and moral choices, thus developing our perspectives and discriminatory capacities with respect to moral issues.<sup>34</sup> Given that philosophy hardly ever takes emotions into consideration, Nussbaum argues, it is not most suitably equipped to make us understand the moral complexity of our lives. In addition, Nussbaum claims, literature takes up topics overseen by philosophers, such as the role of luck in our lives or the problem of conflicting values. That is why moral philosophy should join forces with art, and literary criticism. On her view,

The poets offer us not simply an alternative route to a contemplative or Platonic type of knowing; their disagreement with Plato is more profound. They claim to offer us an occasion for an activity of knowing that could not even in principle be had by the intellect alone. If their claim is plausible, then their works (or works like theirs) are not optional, but ineliminable in a full investigation of [ethical] matters (Nussbaum, 2010, p. 46).

---

<sup>34</sup> See also Carroll 2011. On Carroll's view, narratives evoke emotions on the part of the audience and such emotions help them realize what is at stake in any given moral situation.

The idea that art and moral philosophy should join forces has found some support, although the debate is often fierce between those who believe that literature can be helpful, at times even better than philosophy, at handling moral issues and those who see philosophy as the only domain relevant for theorizing about morality. As Eileen John expresses the clash,

In any case, in linking the ideas grounded in literary works to those offered by philosophers, my aim is not to show that the ones found in literature are right or philosophically superior. But I think that the ideas found in literature are relevant to how people who soak up this literary tradition experience and understand morality, and for that reason these ideas are worth exploring. We should not assume that explicit philosophical theorizing about morality exhausts our conceptions of morality and we should be open to the complications that may result from consulting literature as resource (John, 2010, p. 297).

When advocating literature's handling of moral issues, philosophers usually claim that it is the dedication to details and the particularities of experience that makes the difference. On Alan Goldman's view:

In presenting morally charged situations in their fully complex history and detail, novels can teach us to attend to all morally relevant features of such situations, features typically overlooked or undervalued when simply applying general rules in making decision (Goldman, 2013, p. 109).

While philosophy is abstract and aims for generalizations, literature seeks to situate moral problems within detailed contexts, thus making them more recognizable as concrete, relatable problems which demand personal, subjective responses devoid of abstractions and generalizations. For example, Kant's dictum against using other people as means to our ends seems perfectly clear, self explanatory and obviously right. However, what happens in situations when life – complex social norms entwined with complicated personal relations, goals, attitudes and desires – gets in way? What narrative art – think of Henry James' *The Wings of the Dove* or Kazua Ishiguro's *Never Let Me Go* – shows is that our everyday relations with others are far more complex than Kant's neat formula

can accommodate.<sup>35</sup> And unlike Kant's philosophical writing, narrative art can, as we saw above, help us understand reactions of others and show us possibilities for how to react in such morally gray situations.

However, not all philosophers agree with the idea that moral lessons are available to us through narrative art. Some claim that looking to literature for a moral lesson is a category mistake: whichever moral concepts feature in a literary work, their role is first and foremost literary. Attending to literature should primarily be guided by the literary stance, that is, our attention to the form, development of a theme, conjunction of subject and theme, etc., not by our desire to see if the moral claims that the work puts to view are true. Evaluating the moral lesson apparently depicted in a novel is not in any relevant sense part of our literary practices as moral lessons (or insight) do not relate to literary value.<sup>36</sup> A rather worrying issue for those who want to defend the claim that narrative art can be a source of moral knowledge or insight is a dilemma raised by Peter Lamarque. As he sees it, either the moral lesson is too close to the work to count as an independently generalizable principle applicable to the real world, or the moral lesson is too detached, too loosely connected to the specificities of the work to be perceived as part of the literary content or meaning that the work expresses.<sup>37</sup> Such claims are plausible under a particular conception of literature (and art generally), namely one that sees literature as a social practice. Underlying such a practice are conventions which govern the behaviour of those who participate in the practice, and one such convention demands that in attending to the work, we take a primarily literary stance.<sup>38</sup> In other words, we do not engage with the work to extract moral lessons from it. However, even if one were to concur with this theory, it is hard to shed our experience of making moral judgments and assessments as we go along: it is hard to imagine

---

35 See Vidmar 2016, 2017a.

36 This is in a nutshell a sketch of the argument presented by Lamarque and Olsen 1994.

37 Lamarque 1996.

38 See Olsen 1978, 1987; Lamarque and Olsen 1994, Lamarque 2009.

someone reading a work such as *Lolita* without finding Humbert Humbert morally blameworthy. Furthermore, many works develop their subject around a particular moral dilemma that characters face, where a failure to recognize the relevant parameters of such a dilemma amounts to misunderstanding the work. Consider the example from Henry James's *The Wings of the Dove*, where Kate first mentions the dilemma she faces: her rich aunt will provide for her and secure her a wealthy life. However, Kate has to give up first her father and then the man she loves:

The condition Aunt Maud makes is that I shall have absolutely nothing to do with you; never see you, nor speak, nor write to you, never go near you and make you a sign, nor hold any sort of communication with you. What she requires is that you shall simply cease to exist for me (James, 1902, 1971, p. 13/14).

As the argument is presented, attending to Kate's dilemma from the moral point of view and seeing it as a choice that one has to make, which will inevitably bring her, and others, harm and pain, partly at least amounts to a misrepresentation of the literary context within which Kate's dilemma is developed. Rather than acknowledging the demand put upon Kate and considering its ethical reverberations, readers should attend to how James develops the story, its action, characters and their interactions, against this dilemma. In that sense, the dilemma is relevant for the interpretation of the story, rather than for what it might tell us about human morality generally and the choices we are sometimes forced to make. How is a cognitivist to respond to this?

One way is to claim that one can approach literary works from any number of perspectives: from the literary/artistic, social, psychological and various others, including the ethical. But that is not an interesting response, as it does not acknowledge the fact that some novels inherently take on, as their central concern, ethical situations. The centrality of the dilemma for all that Kate does in the novel diminishes the claim that what is relevant in a literary context is the function that any given scene has for the overall literary achievement that of a work. At some point in the overall experience with this particular work, the audience is bound to make a



moral judgment regarding all the actors who create this particular situation, as well as a judgment regarding characters' reactions to it. What cognitivists claim is that, in the process of doing so, certain insights regarding human morality are available. One might understand Aunt Maud's motivation for forcing Kate to make such a demanding choice, given the hardships her father caused to the family, and given the poor social standing of Kate's lover, and see her ultimatum as a way to protect an overly sensitive, inexperienced girl. Or, one might consider the aunt as unnecessarily overprotective, arguing that Kate should be given the liberty to make her own choices, even bad ones. Contemplating on the actions of this character, the cognitivist claims, is beneficial, in that it makes us reconsider our understanding of what it means to protect someone, to do the best for one's family, of what it is to err and forgive, of what is to use and abuse another human being, etc. The cognitivist is not committed to the claim that the work will give definite answers on these issues, or that, without the book, one would never consider them on one's own initiative.

## **6. Acquiring Cognitive Benefits from Narrative Art**

So far, we have been discussing different instances of how narrative art can be cognitively valuable. Now, we have to explain how precisely the cognitive transfer between the work and the spectator takes place. We will consider two accounts: one that draws an analogy between narrative art and thought experiments (hereafter TEs) and one that states that narrative art is a form of testimony.

From the epistemic point of view, the question how we learn from narrative art arises because such works are predominantly fictional, which makes a demand to explain how we obtain real-world, factual knowledge from them. Fictionality itself does not render descriptions epistemically impotent, given that being fictional does not amount to being false or made up. However, the problem remains in that fictional narratives are exempt from the norms of epistemic reliability. One way to overcome this problem is to show that fictional elements do not diminish the cognitive po-

tential of narrative works. Noel Carroll does this by drawing the analogy between TEs and literature:

philosophy employs a gamut of techniques to produce knowledge and learning that are analogous to those found in literature. What I have in mind here specifically are thought experiments, examples and counterexamples that are often narrative and generally fictional in nature. (...) Thus, if these strategies are acceptable forms of knowledge production in philosophy and if literature contains comparable structures, then if philosophy conducted by means of thought experiments is an adequate source of knowledge and education, then so should literature be (Carroll, 2002, p.7).

Carroll explains the cognitive benefits of narrative art by invoking the similarities that exist in the way TEs are designed and the way an author creates a fictional world. However, why think that narrative art and TEs are sufficiently similar? Consider how TEs are generally understood. David Davies, another philosopher who defends the analogy, explains (scientific) TEs as taking the

form of short narratives in which various experimental procedures are described. Competent reader understands that these procedures have not been, and usually could not (for some appropriate modality) be, enacted. She is invited, however, to imagine or make believe that these procedures are enacted and to conclude that certain consequences would ensue, where this is taken to bear upon a more general question, which is the topic of the TE.<sup>39</sup>

This is not the only account of TEs,<sup>40</sup> but it is helpful in explaining why we might consider some works of narrative art as TEs.<sup>41</sup> Notice that the main structure of a TE is similar to that of a narrative artwork: one attending to a TE is asked to imagine or make-believe what the TE describes, all the while knowing that what it describes does not exist. In the same way, one who attends to a narrative work, such as Margaret Atwood's *The Handmaid's Tale*, knows that what is described does not obtain, but can still feel invited to contemplate what is described – e.g. the particular political and social system depicted in the novel, film or series –

---

<sup>39</sup> Davies 2007b, see also Davies 2007a, 2010.

<sup>40</sup> See Gendler 2000.

<sup>41</sup> Philosophers who defend the analogy include Davies 2007a, 2007b, 2010a; Carroll 2002; Elgin 2007; Dadlez 2009; Dorhn 2009; Vidmar 2014a, 2016.

and to come to a conclusion about what it would be like to live in such a system. Thus, the argument goes, by being invited to imagine something, one can come to have knowledge about it without actually having the experience. This is possible because TEs

make connections – that were hitherto recessive or obscure – between what is already known and other parts of our cognitive stock. They illuminate the relevance of what is already known to the question at hand by refocusing that knowledge in a novel way. This counts as knowledge productions, because it clarifies linkages between parts of our cognitive map. Apart from that, they can also raise counterexamples to well accepted theories, make argumentative points, motivate conceptual distinction and give counterexamples to widely accepted claims (Carroll, 2002, p. 7.).

The explanation that Carroll gives of the operative mechanism of TEs explains such cognitive benefits associated with indirect humanism, particularly with the idea of cognitive strengthening. A similar idea is defended by David Davies:

The suggestion then, is that the mental models through which readers comprehend fictional narratives also provide, through their mobilization of tacit or unarticulated knowledge of the world, a means of testing those claims to knowledge of the actual world that theorists have located in fictional narratives, and thereby validate the idea that fiction can be a genuine source of knowledge of the world (Davies, 2007b, p. 44.).

Because readers have the experience of, for example, what it is to be a parent, a spouse and a citizen, they have some notions of what such roles imply and how they determine one's goals and values. When reading a novel or watching a movie that makes salient changes to how such roles are developed, such as Margaret Atwood's novel cited above, one is forced to reconsider one's own conceptions of these roles, mostly by being challenged to respond to the situations depicted in the work. One's dormant knowledge, beliefs and opinions – e.g. of what it is to be a spouse, a parent, a rape victim – are triggered and activated in the process of evaluating them against the script offered by the work. Thus, one can understand these roles and one's conceptions of these roles through a different, arguably more developed perspective. The outcome of such process can be – as Kitcher points out in his analysis of synthetic complexes – an endorsement or rejection of one's original

conception, a recognition that one needs to readjust one's ideas, or an acknowledgment of the correctness of one's worldview.

The explanatory value of the analogy between narrative art and TEs is potent, as it surpasses the problem of fictionality, the no-expertise argument, and what is called the no-evidence argument, according to which narrative works of art do not offer supporting arguments or evidence for the claims they put forward. At best, the argument goes, works of art present a certain perspective or offer a representation of things, but, as John Gibson explains it, there is "a considerable gap between the offering of a representation and the establishment of its truth, and literary works appear to have no interest in filling in this gap" (Gibson, 2007b, p. 4). What the analogy with TEs shows is that the gap is filled in by the reader's dormant knowledge; though the proper cognitive gain is not necessarily one which can (only) be assessed as true or false. Another advantage of the analogy is its reliance on the general human capacity to employ counterfactual thinking and imaginative engagements to reach knowledge and make sense of experience.

The analogy however is not without its limits and problems, some of which are discussed by James Hamilton's contribution to this collection.<sup>42</sup> Consider again Davies' definition: TEs use procedures that have not been and could not be enacted. Thus, when imagining the content presented in this way, the reader has to be aware that it is precisely because of this inability to actually carry out what is described that the whole set of propositions is put forward in the first place. If this is the defining principle of TEs, and if the same principle applies to narrative art, it seems that the analogy cannot be applied to all works. Those left out include realist works, including historical novels which rely upon historical accuracy; works that (set out to) present or portray a certain political regime, historical event, social circumstance, etc, paying close attention to this accuracy; works which, although unrealistic from a contemporary perspective, reflect certain elements or beliefs of the target audience at the time they were written. All such works

---

<sup>42</sup> See Vidmar 2014 for exploring the scope of the analogy.

rely on the accuracy of what has been depicted, rather than on presenting scenarios which have not, and cannot be, enacted. The conclusion then is that the analogy can only be applied to works that in some sense deviate from the verisimilitude principle, where this is obvious in breaching the laws of physics, in portraying patterns of human behaviour, and relations or social and biological structures that are saliently different from our world. The relevant background against which a reader addresses these works by necessity includes an awareness of this breach. Examples here are works pertaining to science fiction, works that rely on state-of-the-art-scientific developments and magnify them, the fantasy genre, horror stories, utopias and dystopias, etc.

In another sense, the analogy seems explanatorily powerful if we think of narrative art as inviting reflection on the thematic concerns it puts to view. A realist novel and a science fiction dystopia can both deal with the same set of thematic concerns. Think, for example, of the political issues underlying *Fathers and Sons* and *1984*, which both invite reflection on political regimes, the role of the state in determining one's personal liberties, etc. In that sense, works of narrative art are extended TEs that work by triggering the spectators' dormant knowledge, experience, intuitions and the like.

### 6.1. Literature as a Form of Testimony

To epistemically ground the cognitive value of literary works of narrative art, I suggest to treat these works as a form of testimony – for terminological reasons, I will refer to it as fictional testimony.<sup>43</sup> The motivation for this is multiple. On the one hand, there is a structural symmetry between testimony and literature, as in both cases we are told various sorts of things that can deliver knowledge

---

<sup>43</sup>I developed this analogy in Vidmar and Baccharini 2010, Vidmar 2010, Prijić-Samaržija and Vidmar 2012, Vidmar 2012. Treating films, plays and series as a form of testimony is not as straightforward as treating literature as such, as these forms rely on vision and visual perception. Therefore, this chapter might be more narrowly construed as depicting the particular mechanism which grounds the cognitive value of literature.

and other cognitive benefits such as understanding. Authors invite us to attend to what they are telling us through their stories in the same way as our informants in everyday conversations invite us to attend to theirs. In both cases, an informant is reporting something and by attending to the content, we may expand our body of knowledge and our conceptual framework. With respect to literary fiction, we are invited to attend to artistic and aesthetic features of how the story is told, but that does not stand in the way of readers being moved in cognitive and emotional ways – after all, our informants in everyday contexts can use poetic language just as well.

Another reason to take the analogy seriously comes from the conception of author as the observer of society. As we saw above, particularly in realism, but also in other periods, authors often write about their own society, from the perspective of a participant in the community. In that sense, their works can be taken as first-person reports about what is going on in any given culture. Thus, they satisfy the most intuitive account of what testimony is: a report about things one has witnessed first-hand. It might be objected here that authors distort their reports for aesthetic, artistic, and other reasons, which might raise suspicion about the reliability of their reports and ultimately render them cognitively impotent. However, unlike in cases of insincere testimony where listeners remain unaware of the informer's intention to deceive, or in the cases of unreliable testimony where the grounds for justification are disturbed given that the informer is not an expert or is mistaken in what he is saying, in the case of fictional testimony readers are aware of the fact that they are reading fiction, since they are aware of the practice of story-telling that underlines our literary creations and engagements. By choosing to write literary works, authors express their intention to write a kind of work where fidelity to facts is not the governing maxim. Therefore, authors cannot be said to purposefully deceive. Readers' awareness of participating in this literary practice includes their awareness of the possibility that the work distorts facts. But it is not by default that all works distort all facts, and so the statement that literature is unreliable is not justified and the judgment of unreliability cannot be applied to all liter-

ature. The fact that individual works might be unreliable does not render literature *as such* unreliable.

Let us specify the details of the analogy between testimony and narrative art. For it to work, we need to abandon the idea that testimony has to do with the informer transmitting her beliefs to the audience.<sup>44</sup> Jennifer Lackey (2008) opposes this view and shows that we learn from other people's words, from what they are saying, not from what they believe. Once testimony is in this way separated from one's beliefs, it is easier to see why we should think of literary fiction as testimony: authors do not have to believe what they are saying – or, to put it differently, following David Davies, authors are not bound by the fidelity constraint – in order for their works to count as potential source of cognitive gain. Second, we have to abandon the so-called Narrow View of testimony, on which testimony is limited to reporting factual information to those in need and it serves as evidence for the claims expressed (Coady, 1992). Such an account rules out literature, since neither is literature a report of factual information, nor is it offered as such, nor is the work itself evidence for the truth of what it presents, as Stolnitz (1992) pointed out. However, we can abandon this view in favor of a less restrictive one. Several epistemologists have claimed that testimony is broader than Coady takes it to be and that it includes “not only cases of telling but also cases of the expression ... of judgments, views, and opinions with no restriction either on the subject matter, or on the speaker's epistemic relation to it” (Fricker, quoted in Lackey 2008, p.50). This Broad View of testimony is more in line with our everyday communicative practice, which does justice to the benefits available through such exchanges. According to this account, one can testify simply by telling things, without the condition that the testimony be taken as evidence of what is stated or that it resolves the question that the audience is in need of answering. The act of testimony then is not defined according to the needs of the audience, or according to the content of the utterance, but by the act of telling itself. Therefore, testimony is much wider than

---

44 See Pritchard 2004.

Coady's view suggests. It includes cases of "tellings in general (i.e. with no restriction either on the subject matter, or on the speaker's epistemic relation to it)" (Fricker, 1995, pp. 396-397) and it can take the form of an extended narrative.<sup>45</sup>

The main motivation for accepting the Broad View of testimony is our everyday conversational practice. Rather than thinking that the paradigm of testimony is an act of witnessing in the courtroom, as Coady's view indicates, we should consider our everyday communication. As participants in communicative practice, we are told not only factual reports about what happened, when, where, etc., but also the testifier's opinions, judgments, attitudes and similar reflections about their lives, emotions, experiences and the like. The cognitive gains from a testimonial exchange are wider than the Narrow View suggests; i.e. we gain more than what is captured by propositional knowledge. Informed by what other people are saying, we change our ways of thinking, much as the indirect humanist argues in explaining the cognitive benefits of narrative art. Other people's stories might change our views, inspire us to consider new possibilities previously unnoticed, or show us that our views are (or are not) strong enough to face potential counterarguments. Testimony, understood in the broad sense, gives us opportunities to develop our cognitive economy because it influences our conceptual repertoire, the perspective from which we think about things, and the depths to which we are forced to consider them. As we saw above, narrative art does the same. So the broad view of testimony, when applied to literature, helps us ground its cognitive benefits without committing us to the propositional theory of cognitive value. Instead, it allows us to embrace all the cognitive benefits associated with narrative literary art.<sup>46</sup>

---

45 See also Robert Audi 1997 and 2006 for the Broad View of testimony.

46 James Young for example rejects testimony as the underlying mechanism that grounds fiction's cognitive value because, as he argues, such a view presupposes a propositional theory of cognitive value, according to which literary works are cognitively valuable in the light of informative propositions contained explicitly or implicitly in the work. Accepting the Broad View of testimony allows us to overcome such worries.



Emphasizing the structural and operative similarities between everyday testimony and fictional testimony is not enough to show that narrative art can be an epistemically valid source of knowledge. Even if we understand literature via the analogy with TEs and feel convinced that we gain benefits along the lines of indirect humanism, we need to show that, if literature is a form of testimony, it satisfies two crucial criteria for it to transmit knowledge and other cognitively valuable states: sincerity and reliability of the speaker. This is another way of acknowledging Plato's worries regarding the epistemic authority of the poet: why should we trust what a poet is saying, given his lack of knowledge of the things he is writing about? As we saw above, the worry is given an extra push by the freedom from the fidelity constraint. However, there are ways to surpass it. First of all, the problem is not unique to fictional testimony, given that, as listeners, we are always vulnerable to insincere and unreliable informers. Snježana Prijić Samaržija has written extensively on this, and her solution can help us apply the analogy to works of narrative art.<sup>47</sup> She argued that in testimonial exchange, listeners should 'do their part' to make sure they do not accept something that is not true. According to such evidentialist position, the listener should have the evidence which shows that the speaker is sincere and trustworthy. Overall, her empirical and inferential evidence should not speak against the reliability of the speaker but in favour of accepting the testimony.<sup>48</sup> Two relevant factors can help the listener obtain such evidence: her background beliefs and the context of the testimonial exchange.<sup>49</sup> These factors play an important role in providing listeners with information regarding the aspects of the situation in which testimony is being offered, as well as the

---

47 See Prijić-Samaržija 2000, 2001, 2003, 2006, 2007.

48 See Prijić-Samaržija 2007, particularly pp. 678-679.

49 See Prijić-Samaržija 2006 and 2007. Prijić-Samaržija claims that our experience of participating in the testimonial exchange helps us differentiate between acceptable, non acceptable and partly acceptable testimony. It also teaches us that the practice of communication is characterized by stability and uniformity. Both of these help us recognize the aspects of situation in which we are being delivered a testimony; this is what our evidence consists in. (See Prijić-Samaržija 2007, particularly p. 680).

relevant features of the speaker. In addition, Prijić-Samaržija points out, through practice and experience, listeners develop the abilities to differentiate between those topics on which informers generally are less trustworthy and less reliable.

We can accommodate this evidentialist position so as to make it hospitable to the contextual specificities of narrative art understood as a form of testimony. Primarily, we need to adjust the sense in which an author is to fulfil conditions of sincerity and reliability. We already saw that readers' awareness of participating in the practice of literature renders issues of authors' sincerity irrelevant. However, criteria listed by Prijić-Samaržija are helpful in assessing the reliability of authors. We saw above, when we compared Dreiser and Hedwig Courths-Mahler, that readers, familiar with literary practice and different genres, can easily differentiate reliable from non-reliable informants, i.e. authors. However, it is important that we do not presuppose, as Plato does, that readers are not making the relevant judgments when it comes to evaluating the content of the story they are attending to. In other words, readers are – or should be, if they are to be responsible with respect to their sources – reflective when attending to the work and sensitive to what the work puts to view, primarily when it comes to accepting certain morally problematic perspectives. For example, there are several instances in *Lolita* where Humbert seeks to explain and justify sexual relations with minors and young adults. Were a reader to unreflectively accept these and form a belief that paedophilia is not morally blameworthy and damaging for children, she would be held responsible for her moral corruption. While nothing in literature serves as a safeguard against this possibility, it is up to the readers to make sure they do not accept faulty or morally problematic stands. However, the same applies with respect to everyday communicative practice, where our informants can also push us into accepting something morally disturbing.

Given that works of narrative art tend to be a patchwork of reports, reflections and characters' conversations, dialogues and monologues, they are complex and multilayered, arguably more so than our everyday conversations. This raises the question of how

precisely to account for the analogy, and how to evaluate different aspects of a narrative with respect to reliability. Notice however that our everyday testimonial exchanges can be just as complex and consist of multiple types of information. Breaking the testimony into smaller bits and evaluating these separately is thus something that we generally do, and narrative art does not count as an exception.

Readers need to pay attention to reports which state what is true in the fictional world, but, as we saw, such reports often contain truths *per se*, i.e. truths about the real world. Differentiating between the two is not always easy, and neither is recognizing distortions of factual reports done for artistic purposes, which is certainly one of the reasons which motivate anti-cognitivist position. However, anti-cognitivism is too quick to abandon the cognitive value of works because of this problem. Resources that readers have to evaluate which descriptions and reports are true, or which might be true, include their background knowledge, knowledge of the history, sociology, politics and other domains, their familiarity with literary genres and styles, etc. In addition, they can consult other resources, such as history books, to find additional evidence. Readers can also differentiate authors' reliability with respect to the subject/theme of a work.

What about the reliability of those parts of a work which are not factual, but which report characters' (or its author's) reflections? Given that a literary work is multilayered and complex, and often long and detailed in what it puts to view, readers should, ideally, use different methods to evaluate different parts of the work. Consider an extract from Dreiser's *Sister Carrie*

Among the forces which sweep and play throughout the universe, untutored man is but a wisp in the wind. Our civilisation is still in a middle stage, scarcely beast, in that it is no longer wholly guided by instinct; scarcely human, in that it is not yet wholly guided by reason. On the tiger no responsibility rests. We see him aligned by nature with the forces of life -- he is born into their keeping and without thought he is protected. We see man far removed from the lairs of the jungles, his innate instincts dulled by too near an approach to free-will, his free-will not sufficiently developed to replace his instincts and afford him per-

fect guidance. He is becoming too wise to hearken always to instincts and desires; he is still too weak to always prevail against them. As a beast, the forces of life aligned him with them; as a man, he has not yet wholly learned to align himself with the forces (Dreiser, 1981, p. 73).

What matters here is that the reader engages with the thematic concepts that Dreiser is interested in – reason, instinct, passion, will, human being – and that he recognizes concerns Dreiser raises – that of the role that instincts, on the one hand, and will and reason on the other, have with respect to a person’s actions and behaviour, as well as the question of moral responsibility. Such reflections are not to be evaluated with respect to truth, but in terms of how they contribute to the way the reader himself thinks of these concepts. Readers come to works with their own set of beliefs and opinions, judgments and attitudes, and they rely on that set to evaluate the perspective offered by the work – recall the explanation that Baccharini provides of reflective equilibrium or Kitcher’s notion of a synthetic complex. Thus, they will consider whether the extract seems plausible, coherent, sensible, challenging, revelatory and/or illuminating with regard to the aspects of the world (people, society, etc.) it brings to view, and will either concur with the author or reject his views.

Conversations are another important element in literary works which bear cognitive value. Attending to characters’ thoughts and statements can be illuminating with respect to how people reason, so much so that some philosophers and cognitive scientists argue that via engagements with art, we can improve our capacity to understand and predict other people’s behaviour. Furthermore, different fictional characters can advocate different perspectives, thus enabling the reader to get an insight into the dialectical process of arguing in favour of or against something. To go back to Plato’s concern about art presenting something under one aspect, dialogues often give an opportunity to compare and contrast two or more such aspects of the same thing, i.e. two or more possible ways of thinking about an issue.

One final point regarding the role of readers and their epistemic position in fictional testimony: those epistemologists, like Pri-

jić-Samaržija, who accept some kind of reductionism with respect to testimony claim that a hearer should independently confirm the testimony before accepting it as true, which means searching for additional evidence that support the claims made by the speaker. Such an independent confirmation presupposes an active and engaged reader, rather than one who, in the manner presupposed by Plato, uncritically accepts all that a work presents because of the pleasures of mimesis and the alluring power of verse. Wolfgang Huemer describes such a reader:

the cognitive value of literature depends not only on the text, but also on the receptive reader. We must not see the reader as an empty sheet of paper on which the author inscribes truths, but as a rational agent who weighs the author's opinion against hers, who reads critically, and who has the freedom to accept or dismiss insights from literary texts, and even if she dismisses the insights, she is invited to form an informed judgment on a new topic. (...) Literature (...) negotiates with reader, as it were. By doing so, it enriches our reflective abilities. Narrative texts focus our thoughts on a topic; they enrich our understanding by inviting *acknowledgment*, and urge us to arrive at an informed judgment about topics we might otherwise have neglected (Huemer, 2007, p. 242).

An illustration of such independent confirmation comes from the theory developed by David Novitz. As he claims, readers extract various beliefs from a fictional world and apply them to real world situations which resemble those from the fictional world: "If a particular factual belief acquired from fiction and tentatively projected on to the world does not enable us to negotiate the world better, we will reject this mode of thinking and observing" (Novitz, 1984, p. 63). On the other hand, if the beliefs acquired from fiction prove useful and valuable to how we understand the world and other people, then indeed we can say we have learnt something from fiction. The necessary justification for this will come from the fact that readers rely on their 'real world experience' to confirm beliefs derived from fiction. In a similar vein, Peter Kivy defends what he calls a *theory of literary plausibility*. His defense of the cognitive value of literature is based on the claim that literary works give us hypotheses about how to approach our experience. Drawing on the work of William James, Kivy develops his distinction between

live and dead hypotheses. The difference is that live hypotheses are recognized by the audience as worthy of further consideration and thought, while dead ones are recognized as of no interest and probably untrue.<sup>50</sup> The important aspect of a live hypothesis is the content: what the hypothesis suggests has to ‘matter to us’, in the sense that it concerns matters of “deep and abiding significance” (Kivy, 2006, p. 103).

Naturally, a significance of a live hypothesis is relative to the reader’s interest, and the ‘liveness’ itself can change over time, due to the “passage of time, the advancement of learning” (Kivy 2006, p.103), and, we can add, changes in social, political, scientific, cultural, ideological, etc. aspects of our circumstances.

## **7. Conclusion**

The aim of this essay was to present an overview of aesthetic cognitivism: a view according to which art is cognitively valuable. Our focus was on narrative art, and more narrowly, on literature. Our perspective was strictly epistemological, in that we did not aim to establish the connection between the cognitive dimension of works and their artistic or aesthetic value – though the hope remains that the analysis presented here can be inserted into some such account. The assumption was that the most recent developments in epistemology, primarily the pluralism of epistemic values, enable a cognitivist to offer more convincing arguments in favour of her view, arguments which are not available if propositional knowledge is seen as the only cognitive gain generally and the only one available in literature and narrative art.

The main idea behind this approach to aesthetic and literary cognitivism is that narrative art is primarily concerned with people, with our experience, with who we are and how we live. There-

---

<sup>50</sup> Here is Kivy: “A live hypothesis is one that appears to the person who contemplates it as at least a viable candidate for belief, even though he or she might not presently believe it. A dead hypothesis, on the other hand, is one that has no such appeal at all, but is taken to be not a possible option, that is to say, not possibly true” (Kivy, 1996, p. 102).

fore, whatever it has to say – provided of course we are not dealing with trivial, trashy, or formulaic art devoid of psychological, social, ethical and other complexities distinctive of human experience – should matter and should be taken into consideration. This is not to suggest that we should not look for artistic and aesthetic values in art, or that we should not hedonistically enjoy the pleasures art provides; it is only to underline some of the crucial ways in which engagements with art can be beneficial to us. On the view defended here, such an approach does not diminish artistic value and does not presuppose a mistreatment or instrumentalization of art; rather, it celebrates all the values, benefits and enjoyment that art has to offer.

Art's way of being cognitive is as unique as it is irreplaceable, and for all the strength that anti-cognitivism derives from such claims as the no expertise argument or the no fidelity constraint, there is plenty that we can gain from art. As responsible and reflective cognizers, we should aim to know as much as we can about things that matter. Narrative art not only deals with such issues, but it does so in a manner that enables us to think better about them and to develop cognitively, imaginatively and emotionally more sensitive repertoires for appreciating them. Works of narrative art ask us to consider options we did not think relevant or did not recognize as possible – because of that, we might have missed something important in our experience of the world. Literary works show that things might be different than they seem – that might save us the pain of learning such lessons in hard way. In some cases, literary works will make us wonder about the values we place on different aspects of life or connections we make with people. All these and many more lessons can be extracted from narrative art, if one is open to looking for them.

## Bibliography

- Aristotle (1989). "Poetics." In J. L. Ackril, ed., *A New Aristotle Reader*, Princeton: Princeton University Press.
- Attridge, D. (2015). *The Work of Literature*. Oxford: Oxford University Press.
- Audi, R. (1997). "The Place of Testimony in the Fabric of Knowledge and Justification." *American Philosophical Quarterly* 34(4): 405-22.
- Audi, R. (2006). "Testimony, Credulity, and Veracity." In J. Lackey & E. Sosa, eds., *The Epistemology of Testimony*, 25-49. Oxford: Oxford University Press.
- Austen, J. (2005). *Pride and Prejudice*. In *The Wordsworth Collection of Classic Romance*, 7-268. Wordsworths Editions.
- Baccarini, E. (2010). "Reflective Equilibrium, Art, and Moral Knowledge." *Proceedings of the European Society for Aesthetics* 2: 20-33.
- Baccarini, E. (2018). "Art, Moral Understanding, Radical Change." *Rivista Di Estetica* 69: 40-53. DOI : 10.4000/estetica.3666
- Barfield, R. (2011). *The Ancient Quarrel between Philosophy and Poetry*. Cambridge: University of Cambridge Press.
- Carroll, N. (1998). "Art, Narrative, and Moral Understanding." In J. Levinson, ed., *Aesthetics and Ethics; Essays at the Intersection*, 126-60. Cambridge, Cambridge University Press.
- Carroll, N. (2001). *Beyond Aesthetics*. Cambridge University Press.
- Carroll, N. (2002). "The Wheel of Virtue: Art, Literature, and Moral Knowledge." *Journal of Aesthetics and Art Criticism* 60(1): 3-26.
- Carroll, N. (2007). "Literary Realism, Recognition, and the Communication of Knowledge." In J. Gibson, W. Huemer & L. Poggi, eds., *A Sense of the World: Essays on Fiction, Narrative and Knowledge*, 24-41. New York, London: Routledge.
- Carroll, N. (2011). "Narrative and the Ethical Life." In G. Hagberg, ed., *Art and Ethical Criticism*, 33-62. Oxford: Blackwell.
- Carroll, N. (2013). *Minerva's Night Out*. Wiley Online Library.
- Carroll, N. (2016). "Character, Social Information, and the Challenge of Psychology." In G. Hagberg, ed., *Fictional Characters Real Problems: The Search for Ethical Content in Literature*, 83-101. Oxford: Oxford University Press.
- Carter, R. and McRae, J. (2001). *The Routledge History of Literature in English*, 2nd edition, London, New York: Routledge.
- Crane, G. (2007). *The Cambridge Introduction to the Nineteenth-Century American Novel*. Cambridge: Cambridge University Press.
- Dadlez, E. (2009). *Mirrors to One Another: Emotion and Value in Jane Austen and David Hume*. Oxford: Willey-Blackwell.
- Davies, D. (2007a). *Aesthetics and Literature*. London: Continuum.
- Davies, D. (2007b). "Thought Experiments and Fictional Narratives." *Croatian Journal of Philosophy* 7(9): 29-45.



- Davies, D. (2010). "Learning through Fictional Narratives in Art and Science." In R. Frigg & M. Hunter, eds., *Beyond Mimesis and Convention: Representation in Art and Science*, 51-70. Dordrecht: Springer.
- Davies, D. (2015). "Fictionality, Fictive Utterance, and the Assertive Author." *British Journal of Aesthetics* 55(1): 39-55.
- Davies, D. (2016). "Fictional Truth and Truth through Fiction." In N. Carroll & J. Gibson, eds., *The Routledge Companion to Philosophy of Literature*, 372-81. London, New York: Routledge.
- Dohrn, D. (2009). "Counterfactual Narrative Explanation." *The Journal of Aesthetics and Art Criticism* 67(1):37-47.
- Dreiser, T. (1981). *Sister Carrie*. Philadelphia: University of Pennsylvania Press.
- Eby, C.V. (2005). "Theodore Dreiser and the Force of the Personal." In R. P. Lamb & G. R. Thompson, eds., *Companion to American Fiction 1865-1914*, 572-85. Oxford: Blackwell.
- Elgin, C. (1993). "Understanding: Art and Science." *Synthese* 95(1): 196-208.
- Elgin, C. (1996). *Considered Judgment*. Princeton: Princeton University Press.
- Elgin, C. (2000). "Reorienting Aesthetics, Recovering Cognition." *The Journal of Aesthetics and Art Criticism* 58(3): 219-25.
- Elgin, C. (2007). "The Laboratory of the Mind." In J. Gibson, W. Huemer & L. Pocci, eds., *A Sense of the World: Essays on Fiction, Narrative and Knowledge*, 43-54. London, New York: Routledge.
- Farrell, F. B. (2007). "'The Way Light at the Edge of a Beach is Learned': Literature as Learning." In J. Gibson, W. Huemer & L. Pocci, eds., *A Sense of the World: Essays on Fiction, Narrative and Knowledge*, 246-65. New York, London: Routledge.
- Feagin, S. (2007). "How *Could* you? Deeper Understanding through Fiction." In J. Gibson, W. Huemer & L. Pocci, eds., *A Sense of the World: Essays on Fiction, Narrative and Knowledge*, 55-66. London, New York: Routledge.
- Felski, R. (2008). *The Uses of Literature*. Oxford: Blackwell Publishing.
- Fricke, E. (1995). "Critical Notice: Telling and Trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony." *Mind* 104(414): 393-411.
- Friend, S. (2007). "Narrating the Truth (More or Less)." In M. Kieran & D. Lopes, eds., *Knowing Art: Essays in Aesthetics and Epistemology*, 35-50. Springer.
- Gaut, B. (2005). "Art and Knowledge." In J. Levinson, ed., *The Oxford Handbook of Aesthetics*, 436-50. Oxford: Oxford University Press.
- Gaut, B. (2006). "Art and Cognition." In M. Kieran, ed., *Contemporary Debates in Aesthetics and the Philosophy of Art*, 115-25. Oxford: Wiley-Blackwell.
- Gaut, B. (2007). *Art, Emotion and Ethics*. Oxford: Oxford University Press.
- Gibson, J. (2003). "Between Truth and Triviality." *British Journal of Aesthetics* 43(3): 224-37.
- Gibson, J. (2007a). *Fiction and the Weave of life*. Oxford: Oxford University Press.
- Gibson, J. (2007b), "Introduction. The Prospects of Literary Cognitivism". In J. Gibson, W. Huemer, L. Pocci eds., *A Sense of the World; Essays on Fiction, Narrative and Knowledge*, 1-9. Routledge.

- Gibson, J. (2009). "Literature and Knowledge." In R. Eldridge, ed., *The Oxford Handbook of Philosophy and Literature*, 466-85. Oxford: Oxford University Press.
- Gibson, J. (2011). "Thick Narratives." In N. Carroll & J. Gibson eds., *Narrative, Emotion and Insight*, University park: Pennsylvania State University Press. <http://louisville.edu/faculty/jogibs02/Thick%20Narratives.pdf> (accessed June 2011)
- Goldman, A. (2013). *Philosophy and the Novel*. Oxford: Oxford University Press.
- Graham, G. (1995). "Learning from Art." *British Journal of Aesthetics* 35(1): 26-37.
- Graham, G. (1996). "Aesthetic Cognitivism and the Literary Arts." *Journal of Aesthetic Education* 30: 1-17.
- Green, M. (2010). "How and What We can Learn from Fiction?" In Hagberg G. & W. Jost, eds., *A Companion to the Philosophy of Literature*, 350-66. Oxford: Wiley-Blackwell.
- Hagberg, G. ed. (2011). *Art and Ethical Criticism*. Oxford: Blackwell.
- Hagberg, G., ed. (2016). *Fictional Characters, Real Problems. The Search for Ethical Content in Literature*. Oxford: Oxford University Press.
- Halliwell, S. (2002). *The Aesthetics of Mimesis*. Princeton: Princeton University Press.
- Hogan, P. (2003). *The Mind and Its Stories: Narrative Universals and Human Emotion*. Cambridge: Cambridge Up.
- Hogan, P. (2013). *How Authors' Minds Make Stories*. Cambridge: Cambridge University Press.
- Huemer, W. (2007). "Why Read Literature? The Cognitive Function of Form." In J. Gibson, W. Huemer & L. Pocci, eds., *A Sense of the World: Essays on Fiction, Narrative and Knowledge*, 233-45. London, New York: Routledge.
- Janaway, Ch. (2006). "Plato and the Arts." In H.H. Benson, ed., *A Companion to Plato*, 388-400. Oxford: Blackwell Publishing Ltd.
- John, E. (2010). "Literature and the Idea of Morality." In Hagberg G. & W. Jost, eds., *A Companion to the Philosophy of Literature*, 285-99. Oxford: Wiley-Blackwell.
- Kaufmann, W. (1992). *Tragedy and philosophy*. Princeton: Princeton University Press.
- Kieran, M. (1996). "Art, Imagination, and the Cultivation of Morals." *The Journal of Aesthetics and Art Criticism* 54(4): 337-351.
- Kieran, M. and Lopes, D.M. eds. (2003). *Imagination, Philosophy and the Arts*. London, New York: Routledge.
- Kieran, M. (2005). *Revealing Art*. Oxon: Routledge.
- Kitcher, P. (2007). *Joyce's Kaleidoscope: An Invitation to Finnegans Wake*. Oxford: Oxford University Press.
- Kitcher, P. (2013). *Deaths in Venice: The Cases of Gustav von Aschenbach*. Columbia University Press.
- Kivy, P. (2006). *The Performance of Reading: An Essay in the Philosophy of Literature*. Oxford: Blackwell.
- Krailshaimer, A.J. (1977). "Introduction to *Salamambo*." In G. Flaubert, *Salamambo*, London: Penguin Classics.

- Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge University Press.
- Lackey, J. (2008). *Learning from Words, Testimony as a Source of Knowledge*. Oxford: Oxford University Press.
- Lamarque, P. and Haugom Olsen, S. (1994). *Truth, Fiction, and Literature: A Philosophical Perspective*. Oxford: Oxford University Press.
- Lamarque, P. (1996). *Fictional Points of View*. Ithaca: Cornell University Press.
- Lamarque, P. (2007a). "Learning from Literature." In J. Gibson, W. Huemer & L. Pocchi, eds., *A Sense of the World: Essays on Fiction, Narrative and Knowledge*, 13-23. New York, London: Routledge.
- Lamarque, P. (2007b). "On the Distance between Literary Narratives and Real-Life Narratives." *Royal Institute of Philosophy Supplements* 82(60): 117-32.
- Lamarque, P. (2009). *The Philosophy of Literature*. Oxford: Blackwell Publishing.
- Lamarque, P. (2010). "Literature and Truth." In Hagberg G. & W. Jost, eds., *A Companion to the Philosophy of Literature*, 367-84. Oxford: Wiley-Blackwell.
- Levinson, J. ed. (1998). *Aesthetics and Ethics, Essays at the Intersection*. Cambridge: Cambridge University Press.
- Matravers, D. (2014). *Fiction and Narrative*. Oxford: Oxford University Press.
- Mikkonen, J. (2013). *The Cognitive Value of Philosophical Fiction*. London, New York: Bloomsbury Academic.
- Mikkonen, J. (2015). "On Studying the Cognitive Value of Literature." *Journal of Aesthetics and Art Criticism* 73(3): 273-82.
- Miščević, N. (2012). "Wisdom, Understanding and Knowledge: a Virtue-Theoretic Proposal." *Acta Analytica* 27(2): 127-44.
- Newman, I. (2009). "Virtual People: Fictional Characters through the Frames of Reality." *The Journal of Aesthetics and Art Criticism* 67(1): 73-82.
- Novitz, D. (1984). "Fiction and the Growth of Knowledge." In J. Margolis, ed., *The Worlds of Art and the World*, 47-68. Amsterdam: Rodopi.
- Novitz, D. (2008). "Fiction and the Growth of Knowledge." In D. Davies & C. Matheson, eds., *Contemporary Readings in the Philosophy of Literature*, 343-58. NY: Broadview press.
- Nussbaum, M. (1986). *The Fragility of Goodness*. Cambridge: Cambridge University Press.
- Nussbaum, M. (1990). *Love's Knowledge*. Oxford: Oxford University Press.
- Nussbaum, M. (2010). "Perceptive Equilibrium: Literary Theory and Ethical Theory." In Hagberg G. & W. Jost, eds., *A Companion to the Philosophy of Literature*, 241-67. Oxford: Wiley-Blackwell.
- Oatley, K. (2017). "On Truth and Fiction." In M. Burke & E. T. Troscianko, eds., *Cognitive Literary Stance, Dialogues between Literature and Cognition*, 259-78. Oxford: Oxford University Press.
- Olsen, S. H. (1978). *The Structure of Literary Understanding*. Cambridge: Cambridge University Press.

- Olsen, S. H. (1987). *The End of Literary Theory*. Cambridge: Cambridge University Press.
- Palmer, F. (1992). *Literature and Moral Understanding; A Philosophical Essay on Ethics, Aesthetics, Education, and Culture*. Clarendon Press.
- Plato (1997). "Apology." In J. M. Cooper, ed., *Plato, Complete Works*, 17-36. Cambridge: Hackett Publishing Company.
- Plato (1997). "Ion." In J. M. Cooper, ed., *Plato, Complete Works*, 937-49. Cambridge: Hackett Publishing Company.
- Plato (1997). "Republic." In J. M. Cooper, ed., *Plato, Complete Works*, 971-1223. Cambridge: Hackett Publishing Company.
- Porter, L. M. and Gray, E. F. (2002). *Gustave Flaubert's Madame Bovary*. Greenwood Press.
- Prijic-Samaržija, S. (2000). "Blind Trust." *Synthesis Philosophica* 1(2): 97-106.
- Prijic-Samaržija, S. (2001). "Trust and Epistemic Cooperation." *Croatian Journal of Philosophy* 1(2): 147-58.
- Prijic-Samaržija, S. (2003). "Trust – between the Disposition to Cooperation and Evidence." *Logic and Philosophy of Science* 1(1): .
- Prijic-Samaržija, S. (2006). "Povjerenje i društveni kontekst." *Društvena Istraživanja* 15(3): 295-317.
- Prijic-Samaržija, S. (2007). "Evidencijalizam i povjerenje." *Filozofska Istraživanja* 27(3): 671-83.
- Prijic-Samaržija, S. (2011). "Trusting Experts: Trust, Testimony and Evidence." *Acta Histriae* 19(2): 249-62.
- Prijic-Samaržija, S. and Bojanić, P. (2012). "Socijalna epistemologija između revizionizma i ekspanzionizma." In S. Prijic-Samaržija & P. Bojanić, eds., *Nenad Mišćević – sva lica filozofije*, 93-107. Beograd : Univerzitet u Beogradu, Institut za filozofiju i društvenu teoriju.
- Prijic-Samaržija, S. and Vidmar, I. (2012). "Fiksijsko Svjedočanstvo." *Prolegomena* 11(1): 65-82.
- Prijic Samaržija S. (2019), 'Democracy and Truth. The conflict between political and epistemic virtues. Mimesis Internationa.
- Pritchard, D. (2004). "The Epistemology of Testimony." *Philosophical Issues* 14(1): 326-48.
- Putnam, H. (1976). "Literature, Science, and Reflection." *New Literary History* 7(3): 483-91.
- Riggs, W. D. (2003). "Balancing Our Epistemic Goals." *Nous* 37(2): 342-52.
- Rosen, S. (2005). *Plato's Republic*. Yale University Press.
- Row, M. (2010). "Literature, Knowledge and Aesthetic Attitude." In S. Schroeder, ed., *Philosophy of Literature*, 1-23. Oxford: Willey-Blackwell.
- Scofield, M. (2006). *The Cambridge Introduction to the American Short Stories*. Cambridge: Cambridge University Press.
- Shippey, T. (2005). "Hard Reading: The Challenges of Science Fiction." In D. Seed, ed., *A Companion to Science Fiction*, 11-26. Oxford: Blackwell Publishing.

- Stock, K. (2006). "Fiction and Psychological Insight." In M. Kieran & D. Lopes, eds., *Knowing Art: Essays in Aesthetics and Epistemology*, 51-66. Dordrecht: Springer.
- Stock, K. (2011). "Unpacking the Boxes: The Cognitive Theory of Imagination and Aesthetics." In P. Goldie & E. Schellekens, eds., *The Aesthetic Mind: Philosophy and Psychology*, 268-82. Oxford: Oxford University Press.
- Stolnitz, J. (1992). "On the Cognitive Triviality of Art." In P. Lamarque & S. H. Olsen, eds., *Aesthetics and the Philosophy of Art*, 337-43. Oxford: Blackwell Publishing Ltd.
- Gendler, T.S. (2000). *Thought Experiment: On the Powers and Limits of Imaginary Cases*. NY: Garland Press (Routledge).
- Schellekens, E. (2007). *Aesthetics and Morality*. London: Continuum.
- Towner, T. (2008). *The Cambridge Introduction to William Faulkner*. Cambridge: Cambridge University Press.
- Treščec, N. D. (2005). *Platonova kritika umjetnosti*. Zagreb: Naklada Jurčić.
- Troscienko, E. and Burke, M. eds. (2017). *Cognitive Literary Science*. Oxford: Oxford University Press.
- Vidmar, I. and Baccarini, E. (2010). "Art, Knowledge, and Testimony." *Synthesis Philosophica* 25(2): 333-48.
- Vidmar, I. (2010). "Against the Cognitive Triviality of Art." *Proceedings of the European Society for Aesthetics* 2: 516-31.
- Vidmar, I. (2012a). "Communicative View of Literature: A Case for Fictional Testimony." *Balkan Journal of Philosophy* 4(1): 81-92.
- Vidmar, I. (2012b). "Plato's Ion in the Context of Literary Cognitivism." *Glasnik za Društvene Nauke* 4: 111-52.
- Vidmar, I. (2013). "Literary Cognitivism and Epistemic Aims." In M. Arsenijević, ed., *Beogradsko-riječki susreti: zbornik filozofskih radova*, 147-166. Beograd: Institut za filozofiju Filozofskog fakulteta Univerziteta u Beogradu.
- Vidmar, I. (2013b). *The Nature of Fictional Testimony and Its Role in Reaching, Fulfilling and Promoting Our Epistemic Aims and Values* (Doctorate thesis).
- Vidmar, I. (2014a). "Thought Experiments, Hypotheses and the Cognitive Dimension of Literary Fiction." *Synthesis Philosophica* 28(1-2): 177-93.
- Vidmar, I. (2014b). "Istine i trivijalnosti književnih djela." *Kultura* 143(2014): 11-33.
- Vidmar, I. (2016). "Identity, humanity and bioethics: philosophical aspects of Never Let Me Go." In A. Prnjat & V. Kolarić, eds., *Shadows in the Cave: Film and Philosophy*, Beograd: Alfa Univerzitet.
- Vidmar, I. (2017a). "On Literary Cognitivism from the Perspective of Difference between Literature and Philosophy." *Synthesis Philosophica* 32: 371-86.
- Vidmar, I. (2017b). "Epistemic Game of Thrones." *Synthesis Philosophica* 32(1): 215-34.
- Vidmar, I. (2019). "Literature and Truth – Revisiting Stolnitz's Anti-cognitivism." *Croatian Journal of Philosophy* 56: 351-70.

- Vidmar, I. and Pektor, L. (2019). "Dangerous Passions of Fictional Characters: Why We Need Them in Our Classrooms." In *Language, Literature, and Emotions*; Proceedings of the seventh international conference, 44-61. Belgrade: Alfa University.
- Walsh, D. (1969). *Literature and Knowledge*. Middletown, Conn.: Wesleyan University Press.
- Walton, K. (1990). *Mimesis as Make-Believe*. Harvard: Harvard University Press.
- Young, J. O. (2001). *Art and Knowledge*. London, New York: Routledge.
- Zamir, T. (2002). "An Epistemological Basis for Linking Philosophy and Literature." *Metaphilosophy*, 33(3): 321-336.
- Zamir, T. (2006). *Double Vision: Moral Philosophy and Shakespearean Drama*. Princeton: Princeton University Press.



# Literary Fiction and the Cultivation of Virtue

James O. Young

---

## 1. Introduction<sup>1</sup>

The hypothesis that reading literary fiction cultivates virtue is an old one. Its origins can be traced to Aristotle and it was widely adopted in the eighteenth century, when Charles Batteux (1746/2015) and Adam Smith (1759/2002) defended it. More recently, Gregory Currie (1995), Martha Nussbaum (1990), Elisabeth Schellekens (2007), and other philosophers have defended the view. Even more recently, psychologists have turned their attention to the hypothesis and sought empirical evidence for it. This essay will critically examine the psychological literature. It will conclude that psychologists have succeeded in mustering considerable evidence for the claim that reading literary fiction cultivates virtue. At the same time, however, this essay will conclude that some of the claims that philosophers have made about literary fiction and the promotion of virtue may need to be qualified.

Two preliminary points are in order. The first is that evidence that individuals display increased empathy and prosocial behaviour will be taken as evidence that individuals have become more virtuous. Certainly, arguments can be given against the view that empathy and prosocial behaviour are indicators of good character or virtue. Nevertheless, this essay will assume that they are. This assumption is widely held and certainly not outlandish.

The second preliminary claim, on which this essay depends, is that the hypothesis that reading literary fiction cultivates virtue is

---

<sup>1</sup> This paper was originally published in the *Croatian Journal of Philosophy* (XIX, 56, 2019).



an empirical hypothesis. (For the sake of brevity, I will henceforth call this hypothesis *H*.) As an empirical hypothesis, *H* ought to be testable by empirical and, indeed, experimental means. That is, an examination of readers of literary fiction should show that they are virtuous, relative to those who do not read literary fiction. Moreover, it should be possible to establish a causal relationship between reading literary fiction and virtuous actions.

While it may seem obvious that *H* is empirically testable, some writers have denied that it is. The argument for denying that *H* is testable runs as follows. According to *H*, readers of literary fiction do not acquire, or do not only acquire, propositional knowledge about how they ought to act. On the contrary, as we shall see, the moral benefits accruing to readers of literary fiction are largely non-propositional. Readers of such fiction become better able to understand other people, more able to empathize with others, and better able to recognise the mental states of others. These capacities, in turn, make them more inclined to engage in prosocial behaviour. Putnam suggested that the sort of knowledge acquired from the reading of literary fiction is of a sort different in kind from that provided by science and, consequently, “inaccessible to scientific testing” (Putnam, 1978, p. 89). Mikkonen (2015) endorsed the view that reading literature does not provide propositional knowledge. On his view, literature provides a sort of understanding or an ability to see significance. He is sceptical about the suggestion that we can test whether readers of literary fiction have this understanding or ability. He writes that, “The enhanced understanding gained by reading fictional literature is akin to happiness, marital satisfaction, or a mechanic’s comprehension of carburetors in that it can be conceived only from inside” (Mikkonen, 2015, p. 277). Some things, he holds, simply do not lend themselves to empirical investigation and the sort of understanding acquired from literary fiction is one of them. We are invited to conclude that empirical investigation, at least of the sort in which psychologists engage, cannot confirm *H*.

This argument is unsuccessful. Grant that reading literary fiction provides readers with a non-propositional knowledge: a way of understanding, or certain abilities, of the sort that Mikkonen and

Putnam have in mind. Grant, moreover, that this sort of understanding or ability is what makes readers of literary fiction more virtuous. The argument shows at most that we cannot express in words what it is like to have this understanding or ability. This is not surprising. Many things cannot be expressed propositionally. For example, it is not possible to capture in words what it is like to be able to ride a bicycle or what being happy is like. Nevertheless, the argument still fails. It is obviously still possible to determine empirically whether someone is able to ride a bicycle or whether someone is happy. This is done on the basis of a person's actions and other observable factors. Similarly, one can determine whether readers of literary fiction become more virtuous by reading literary fiction. We just need to observe a correlation between reading literary fiction and virtuous behaviour. This will not tell us what it is like to have the understanding that makes virtue possible, but it will give us reason to believe that reading literary fiction makes people virtuous.

## 2. The Philosophical Origins of *H*

Although *H* has only recently received strong experimental support, it has long been widely adopted by philosophers. As already noted, *H* can be traced to Aristotle but it was widely held in the eighteenth century. Smith, for example, was of the opinion that literary fiction could make an important contribution to moral education. Moral education, he believed, was largely a matter of cultivating emotional responses. Imagination plays a role in the cultivation of sympathy and other innate moral responses. Smith writes that fellow feeling is not only aroused by the actual suffering of one of our fellows. Rather,

an analogous emotion springs up, at the thought of his situation, in the breast of every attentive spectator. Our joy for the deliverance of those heroes of tragedy or romance who interest us, is as sincere as our grief for their distress, and our fellow-feeling with their misery is not more real than that with their happiness (Smith, 1759/2002, p. 13).

In this way, Smith believes, literary fiction cultivates fellow-feeling

and makes readers more virtuous.

Batteux was another eighteenth-century writer who believed that reading literary fiction can cultivate virtue. In part, poetry (by which he means literature or literary fiction) does so, on his view, by cultivating a capacity for fine-grained perception of social reality. Batteux also agrees with Smith that literature can arouse the emotions required by a virtuous person of good character. Batteux writes that,

in order to give us a perfect and enduring pleasure, it [literature] should only arouse emotions that it is important that we feel intensely and that are not enemies of wisdom. Abhorrence of crime followed by shame, fear, and repentance among other tortures; compassion for the unfortunate, which has an application nearly as extensive as that of humanness; admiration for great exemplars, which inspire virtue in the heart; heroic and, consequently, proper love: these, everyone allows, are the emotions that poetry should address (Batteux, 1746/2015, p. 77).

Batteux suggests several things in this and related passages. For a start, literature represents certain situations or actions and these situations arouse certain emotions. These emotions track the moral qualities of the actions represented. Most importantly, poetry inspires virtues in its readers. Batteux also holds that literature can set up valuable exemplars, worthy of emulation.

Contemporary philosophers have also considered the possibility that reading literary fiction promotes virtue. Nussbaum (1990) was among the first contemporary philosophers to maintain that literary fiction is a valuable source of moral knowledge. On her view, reading literary fiction helps readers understand social situations and understand the complexities of making moral decisions. Similarly, Currie (1995) believes that imagining ourselves in the situations of fictional characters can lead to moral growth. Other philosophers have also suggested that dealing with the hypothetical situations presented in fiction can assist in the acquisition of an ability to act morally. For example, Elisabeth Schellekens holds that reading works of fiction, readers simulate experiences that they can encounter in real life. This experience prepares readers to respond appropriately. Schellekens takes the example of Flaubert's *Madame Bovary* and writes that, persons like Emma Bovary "have, do, and

will exist in reality.” After reading the novel, readers “stand a greater chance of coming to know those who in real life show similarities with Emma Bovary, and may alter [their] actions and judgements accordingly” (Schellekens, 2007, p. 51).

Several themes emerge from the philosophical literature. Philosophers have maintained that, in reading literary fiction, people acquire insight into the lives of others by walking a mile in their shoes. In other words, readers simulate participation in social interaction. They gain practice in such interaction and, consequently, understand others and their motivations. Literary fiction can also provide exemplars of moral behaviour. Practicing social interaction leads to increased understanding of, and empathy with, others. Moreover, readers of literary fiction emulate moral exemplars. As a result, readers of literary fiction are more inclined to engage in prosocial behaviour. In short, they are more virtuous. Let us turn now to the question of whether the empirical literature supports *H* and the conclusions of philosophers.

### 3. The Empirical Evidence

In recent years, many experimenters have found that reading literary fiction is associated with increased empathy. Often the psychological literature distinguishes between cognitive empathy (or a capacity to see matters from other people’s perspective) and affective empathy (or a feeling of sympathy for other people). Various experiments have found that reading literary fiction leads to increases in both cognitive and affective empathy. Experiments have also found evidence that reading literary fiction promotes prosocial behaviour. In short, the empirical evidence seems to support *H*.

A typical experiment is that conducted by Johnson (2012). Test subjects were given the Positive and Negative Affect Schedule (PANAS) to assess their initial mood. Next they read a story designed to provide readers with a good example of prosocial behaviour and to arouse feelings of compassion for the characters in the story. After the subjects read the story, then the PANAS was administered again, together with an instrument measuring affec-

tive empathy. Test participants were asked to use a five-point scale to rate the degree to which they had been moved and experienced compassion, sympathy, soft-heartedness, tenderness, and warmth while reading the story. Next, the degree to which readers had been transported by the story was measured. (Transportation is the feeling of being lost in a book. William James was among the first psychologists to speak of this phenomenon. Referring to Sir Walter Scott's novel, *Ivanhoe*, he wrote that, "Whilst absorbed in the novel, we turn our backs on all other worlds, and, for the time, the *Ivanhoe*-world remains our absolute reality" (James, 1891, vol. II, p. 292–3).) Finally, the subjects were told that they had to retrieve the debriefing forms. As they returned, the experimenter pretended, in full view of the participants, to accidentally drop six pens. He then recorded which of the participants helped pick up the pens.

Johnson (2012) found that test subjects experienced increased affective empathy. Those who experienced higher degrees of transportation into the story showed higher degrees of empathy. Increased empathy translated into increased prosocial behavior: those test subjects who experienced the highest degree of empathy were significantly (almost twice) more likely to engage in the prosocial task (assisting with retrieving the pens that the researcher had pretended to accidentally drop). However, it should be noted that another study did not confirm all of Johnson's results. It found an increase of cognitive empathy after reading a literary short story, but only for subjects with certain personality traits. This study did find that people who frequently read fiction perform better on the Interpersonal Reactivity Index, which measures affective and cognitive empathy (Djikic, Oatley, and Moldoveanu, 2013).

Johnson's results have received support from a series of experiments by Kidd and Castano (2013). Their experiments were designed to distinguish between the effects of literary fiction and popular or genre fiction. They randomly assigned subjects the task of reading works of literary fiction (in this case, winners of literary prizes such as the PEN/O. Henry Award). Control groups read genre fiction (selected from among Amazon.com bestsellers) and works of non-fiction. The subjects who read the works of literary

fiction scored higher on tests of cognitive and affective empathy (the Reading the Mind in the Eyes Test (MIE) and the Yoni test). It is worth noting, however, that scepticism has been expressed about the value of these tests as predictors of prosocial or compassionate behaviour (Koopman, 2015, p. 63).

The studies just considered measured the increase of empathy and prosocial behaviour as a result of exposure to a single piece of fiction. It seems unlikely that reading a single piece of literary fiction will have a huge impact upon a person's character and virtuousness. Kidd and Castano (2013) suggest that reading a single story is unlikely to teach subjects much about other people. Instead, they speculate that reading literary fiction "recruits" (or starts working) their Theory of Mind (ToM). ToM is the "capacity to identify and understand others' subjective states....It allows successful navigation of complex social relationships and helps to support the empathetic responses that maintain them" (Kidd and Castano, 2013, p. 377). More recently, other experimenters have duplicated these results (Black and Barnes, 2015). These authors also found that the benefits of reading literary fiction seems to be limited to improved capacity to understand and respond to social situations. In particular, they found that reading literary fiction does not improve results on the Intuitive Physics Test.

Kidd and Castano (2013) only studied the effects of reading a single piece of short literary fiction. They suggest, however, that extensive reading of literary fiction improves ToM. Let us consider the possibility that regular reading of literary fiction increases empathy and improves character.

Several experiments have measured the impact of a habitual practice of reading fiction. One such study (Mar et al., 2006) began by administering the Author Recognition Test (ART), which provides a measure of what, and how much, individuals read. As revised for this test, the ART provided a measure of how much fiction and how much non-fiction test subjects read. Subjects were also assessed by the Interpersonal Reactivity Index, which measures empathy, the MIE-revised, and the Interpersonal Perception Task-15 (IPT-15). The IPT-15 has subjects view a series of videos of

unscripted interactions between two or more individuals. Subjects then answer a series of questions to determine whether they understand the interactions. It is regarded as a good test of sensitivity and social skills. The Interpersonal Reactivity Index measures, among other things, engagement with narrative (which is akin to transportation).

The researchers found that reading a lot of fiction was correlated with the ability to perform tasks such as the IPT-15 and MIE-revised test. Readers with a high degree of narrative engagement (or transportation) performed particularly well. Reading a lot of non-fiction was correlated with poorer performance on these tests. It should be noted, however, that this test did not distinguish works of fiction and works of literary fiction. Moreover, this experiment does not rule out the possibility that empathetic people are more likely to read literary fiction and that the readers of fiction do not owe their empathy to their reading of fiction. We will consider this possibility below.

A complex study by Koopman (2015) also suggests that familiarity with literary fiction is correlated with increased empathy. In this study, test subjects read texts on either depression or grief. Three sorts of texts were used for each sort of emotion: literary narratives, non-fiction first person narratives, and expository texts. Koopman hypothesized that personal narratives would lead to increased empathy and prosocial behaviour as well as literary fiction does. She also predicted that the texts concerned with grief would have more marked effects on persons dealing with grief. Readers were hypothesized to find it easier to imagine themselves in a position where they feel grief than they can imagine feeling depressed. The experiment controlled for a number of factors, including antecedent empathy, exposure to literature, and personal experience of grief or depression. A questionnaire was used to measure empathetic understanding.

Subjects were then asked about the extent to which they agreed that insurance policies should cover treatment for grief and depression and the extent to which they understood the plight of those suffering from grief and depression. The experiment also built in a

practical measure of prosocial behaviour. Test subjects were given the option of donating some or all of the fee (€10) they received for participating in the study to a charity serving those who suffered from grief or depression.

Koopman found several interesting results that are relevant to present concerns. Those who read personal narratives of depression or grief and (to a somewhat lesser extent) those who read a fictional narrative were more likely to engage in prosocial behaviour than those who read an expository text. This gives limited support to the hypothesis that reading literary fiction promotes prosocial behaviour. Personal familiarity with grief or depression was positively correlated with donations. While the type of text the subjects read was correlated with prosocial behaviour, no correlation was found between familiarity with literature and prosocial behaviour. Exposure to literature did, however, predict increased empathetic understanding. Those with a high exposure to literary fiction were inclined to be in favour of insurance coverage for treatment for depression. On the whole, Koopman's findings are in keeping with those of other researchers. (The number of test participants contributing to charity was small in all conditions. Likely the small number of people donating was affected by the fact that all were students for whom €10 is a significant sum and a considerable incentive to participate in the study.)

Philosophers and psychologists have hypothesized that reading literary fiction makes readers more empathetic and prosocial since, readers of this genre simulate experience of social situations and practice dealing with them. This hypothesis receives support from the study of the brains of people engaged in reading literary fiction. Our brains have what psychologists call the "default network," a collection of regions of the brain that are responsible for simulation. Simulations include mental constructions of social contexts while reading. If reading literary fiction involves simulating experience of social situations, and practicing dealing with social situations, we would expect that the default network would be engaged. This turns out to happen.

In a recent study, test subjects underwent fMRI (functional



magnetic resonance imaging) scans while reading passages drawn from novels and a variety of non-fiction sources, including newspapers, magazines, and self-help books (Tamir et al., 2016). The passages were contrasted along two dimensions: from vivid to abstract and from descriptive of a person's mental content (social) to non-social. Vivid and social works are taken to be literary. (I take it that, in this context, to say that a work is vivid is to say that it employs figurative language.) The fMRI results revealed that vivid passages and passages that describe the mental content of a person or persons recruited the default network. This adds to the empirical evidence in favour of *H*.

#### **4. Criticisms of *H***

While the empirical evidence seems to suggest that reading literary fiction makes people virtuous, someone might object that this evidence is misleading. Possibly highly empathetic people read literary fiction, and this is why reading literary fiction is associated with higher degrees of empathy. In other words, perhaps the causal arrows lead from high empathy to the reading of literary fiction rather than from reading literary fiction to increased empathy. As well, some philosophers have objected to *H* on grounds that reading literary fiction takes readers away from the real world in which they can practice virtuous behaviour.

The possibility the causal arrows lead from being empathetic to reading literary fiction has been anticipated and ruled out in the experimental literature. In one experiment, the empathy of test subjects was measured prior to the experiment, immediately after they had read the text (either a work of fiction or, in the control group, a work of non-fiction), and one week after reading the text. The researchers found that higher empathy measurement post-experiment was correlated with the degree to which subjects were transported into the story. They ruled out the hypothesis that increased empathy post-experiment can be explained by higher empathy pre-experiment (Bal and Veltkamp, 2013). Another study arrived at a similar result. This study tested subjects for the “Big Five”

personality traits: extraversion, agreeableness, conscientiousness, neuroticism/stability, and openness. The subjects were then given the ART and the MIE test. Openness was the only personality trait associated with reading fiction. Performance on the MIE test was also correlated with reading literary fiction. The researchers concluded that they needed to control for openness when gauging the impact of reading fiction on empathy. Analysis of the experimental data revealed that, after controlling for gender (women are more empathetic than men), age, English fluency, and openness, the degree of people's exposure to fiction predicts they will perform better on a test of empathy (Mar, Oatley, and Peterson, 2009).

Some philosophers have also objected to *H*. Candace Vogler has criticised the hypothesis that reading literary fiction leads people to become more virtuous. She believes that, on the contrary, time spent reading literary fiction is, from a moral point of view, wasted. Time spent reading literary fiction is time not spent engaging with one's fellow human beings. The only way to become more virtuous she believes, is to perform virtuous acts. She writes that if, for example, "I seek to cultivate generosity, I *give*....Since silent reading induces *retreat* from my circumstances, silent reading is the *opposite* of habituating myself to noticing what's going on in my world *by* noticing" (Vogler, 2007, p. 33).

The flaw in this sort of reasoning is now apparent. To a certain extent, at any rate, simulating engaging in virtuous and prosocial acts assists people in becoming more virtuous. This should not be surprising. One becomes a better pilot by flying aircraft. But one can also become a better pilot by training on a flight simulator. Similarly, the empirical evidence suggests that a person becomes more virtuous by reading literary fiction and simulating acts of empathy with other people. By reading literary fiction and simulating interacting with other people, readers can learn how to interact better with others.

## 5. How Literary Fiction Makes People Virtuous

The mechanisms by which literary fiction makes readers more

virtuous, and improves their characters, are likely imperfectly understood. Still, the psychological literature is beginning to provide insight into these mechanisms. This section will address three mechanisms that appear to be at work. For a start, readers become caught up in a story and imagine themselves in a social situation. This gives them practice in dealing with, and reflecting on, social situations, especially when readers are transported into a story. In particular, readers can practice “perspective-taking,” seeing the world from the perspective of others. This practice, in turn, helps readers understand other people (that is, it increases cognitive empathy). This leads to increased emotional empathy with a wide variety of people and, in particular, people unlike ourselves. Secondly, literary fiction provides opportunities for self-reflection. In other words, fiction provides readers with the opportunity to examine their own lives and this leads to improved character. Emulation is the third mechanism whereby literature leads to the cultivation of virtue. Humans have a tendency to imitate the actions of others, including others imitated in fiction.

As we have seen, empirical evidence indicates that readers of literary fiction simulate social interaction. This evidence includes fMRI results that show that reading fiction recruits the default network. In simulating social interaction, readers of literary fiction are led to engage in what is known as perspective taking. Perspective taking involves adopting the perspectives of others and imagining what it is like to see the world from their points of view. Readers have the experience of walking a mile in the shoes of a variety of people, and of people quite different from themselves. Having imagined themselves living the lives of others, they acquire more cognitive and affective empathy for a variety of people. That is, they understand the perspective of, and feel for, these people.

That seeing the world from the perspective of others promotes virtue, is supported by the research of Kaufman and Libby (2012). These authors conducted an experiment in which three versions of a story were used. In one, the protagonist was revealed early in the story to be gay (gay-early story). In another, he was revealed late in the story to be gay (gay-late story). In the final version, he was

revealed to be heterosexual (straight-story). All test subjects identified themselves as straight. The experimenters found that readers of the gay-late story were more transported than were readers of the gay-early story. Likely this was because readers found it easier to enter into the life of someone they perceived to be similar to themselves. Most interestingly, the readers of the gay-late story, having walked in the steps of a gay man, manifested positive attitudes towards gay people after reading the story. On a five-point scale of beliefs about gays, they had significantly more positive beliefs compared to readers of the gay-early and straight stories. Similar results were found with stories in which the protagonist was revealed early and late in a story to be African-American. Readers of the story in which the character was revealed late to be African-American were found to score significantly lower on a test of racist attitudes (Kaufman and Libby, 2012). Another study indicated that readers transported into a story about a Muslim woman had increased empathy for Muslims, compared to those who did not read the story (Johnson, 2013).

The effect of simulating social interaction is increased by transportation into a story. Several writers, including Johnson (2012) and Bal and Veltkamp (2013), have noticed that reading literary fiction is particularly associated with increased empathy when readers are transported into the story. When readers are transported, they “let go of key components of their own identity—such as their beliefs, memories, personality traits and ingroup affiliations—and instead assume the identity of a protagonist” (Kaufman and Libby, 2012, p. 2). These protagonists can be quite various and different from the readers, in personality, characteristics, and situation in life. The experience of transportation makes perspective taking more compelling. The experience of reading fiction becomes almost like being another. When these others are diverse, the extent of one’s fellow feeling and empathy can be considerably extended.

Abundant evidence indicates that literary fiction’s focus on the experience of individuals is one of the factors that increases its impact on readers’ characters. Literary fiction focuses on individuals, while non-fiction tends to focus on groups of individuals. Human

beings seem to be constituted in such a way that we are more affected by a story about an individual than a non-fiction report about a group of individuals. Consider for example an experiment that had one group of subjects read a chapter from Malikia Mokkeddem's novel *L'Interdite* (1994). This novel is concerned with the sexist treatment of an Algerian woman who returns to her homeland. Another group read an essay on the condition of women in Algeria (Hakemulder, 2000). Readers of the selection from *L'Interdite* were significantly more concerned about, and inclined to resist, the condition of women in Algeria than were readers of the essay. The opportunity to see the world from the perspective of another human, to be transported, is plausibly held to be the factor that makes literary fiction contribute to increased empathy and prosocial attitudes.

Literary fiction provides readers with a better opportunity to practice simulation of social behaviour than does popular fiction. The fictional worlds of literary fiction have the complexity of the real world. They are not over-simplified and full of caricatures such as Mary (or Marty) Sues. (A Mary Sue (masculine: Marty Sue) is an implausible, over-idealised character.) Since the worlds of literary fiction are realistic, negotiating them is like negotiating the real world.

Consider now the second mechanism whereby literary fiction contributes to the cultivation of virtue. Recently Koopman and Hakemulder (2015) have suggested that reading literary fiction enables readers to engage in contemplation and self-reflection. Here they are building on a remark by Yann Martel, the author of *The Life of Pi* (2001) and other novels. Martel suggested that literary fiction provides readers with the opportunity to reflect on their lives. In particular, Martel spoke of the "stillness" provided by reading literary fiction. It is hypothesized that readers who are more reflective are more likely to avoid purely self-regarding behaviour.

Some evidence indicates that readers of literary fiction are reflective and Koopman and Hakemulder (2015) canvass some of this evidence. Other evidence is provided by an experiment that tracked the sorts of memories evoked by the reading of literary fiction as opposed to other sorts of texts. This experiment had one

group read a short story by Pär Lagerkvist, a winner of the Nobel Prize for Literature. Another group read an essay on the growth of the world's population. As subjects read the text they were asked to record the sorts of memories they experienced. Memories were divided into three categories: memories of events in which readers had actively participated; memories of events which readers had observed without participating in them; and memories of events that the readers knew only by report (Seilman and Larsen, 1989). Readers of the short story were significantly more likely to recall memories of events in which they had actively participated than were readers of the expository essay. Another study has confirmed these results, and found that the memories evoked by reading fiction are more vivid than those aroused by reading non-fiction (Mar and Oatley, 2008). The sorts of memories evoked by reading literary fiction, in comparison to those aroused by non-fiction, is evidence that reading fiction promotes self-reflection.

The question of how self-reflection assists readers in becoming virtuous remains to be addressed. Koopman and Hakemulder (2015) suggest that by promoting self-reflection, by leading readers to take a moment to think, readers avoid knee-jerk reactions. Readers of literary fiction have an increased opportunity to see some matter from a range of perspectives. If this is right, this capacity of literary fiction works in concert with its capacity to promote perspective taking and transportation. By engaging in self-reflection, readers of literary fiction are more likely to engage in perspective taking. As already noted, perspective taking is associated with empathy and prosocial behaviour.

The third mechanism whereby literary fiction improves character is by the setting of good examples that readers can emulate. As we have seen, Batteux long ago suggested that literary fiction functions by setting good examples, examples that readers can emulate. Certainly, a great deal of evidence suggests that humans tend to emulate or imitate the behaviour of other people. As two psychologists note in a survey of the experimental literature, "there is substantial evidence for facial, emotional, verbal, and behavioural mimicry. We mimic virtually everything that we can observe

another person do, and even “catch” their affective states as well” (Chartrand and van Baaren, 2009, p. 226). A good deal of evidence suggests that we do not only mimic real people. We also mimic fictional characters (Eder, Jannidis and Schneider, 2010, p. 55, 57).

Given that mimicry is so common in human behaviour, and that there is evidence that readers mimic characters, it seems likely that part of the effect of literary fiction on character is due to the setting of good examples. This conclusion is suggested in one of the studies already discussed in this essay. Johnson suggests that the prosocial behaviour detected in his experiment was promoted by the fact that the main character in the story used in his experiment “modeled prosocial behaviour” (Johnson, 2012, p. 152). Presumably, readers then mimicked this prosocial behaviour.

Other mechanisms are likely at work when reading literary fiction cultivates character. Several philosophers, including Young (2001) have suggested that the emotions evoked by works of literary fiction, and other works of art, can assist readers in understanding individuals and social situations. This understanding has the potential to increase empathy and prosocial behaviour. Unfortunately, the role of emotions in cultivating virtue has not been subjected to sufficient empirical study. Some tantalizing pieces of information are available. For example, one study has found that reading a short story by Chekhov is associated with the changing of readers’ self-perception of their personality traits and these changes were also correlated with emotional arousal (Kjikic, 2009). Johnson (2012) has also suggested that arousal of compassion, sympathy, soft-heartedness, tenderness, and warmth play a role in promoting virtue. The relationship between emotional arousal by literary fiction and the cultivation of virtue deserves further attention.

## **6. Fiction and Harm to Character**

Many philosophers have suggested that literary fiction can make readers more virtuous, and we have seen that this hypothesis enjoys considerable empirical support. Few recent philosophers have, however, considered the possibility that reading literary fic-

tion could make readers less virtuous, that is, more inclined to make moral errors. Currie is among the few who have considered this possibility. He writes that, while literary fiction has the potential to increase moral understanding, it also has “the capacity to induce moral error” (Currie, 1995, p. 257). Almost no psychologists have entertained or tested this possibility. There is, however, reason to be concerned that some works of fiction could lead readers to be less virtuous.

If we carefully examine the psychological literature, we find that there is reason to worry that literary fiction could make people less virtuous. The problem is that reading literary fiction is a complex activity. In reading literary fiction, affective empathy seems to be induced and this leads to prosocial behaviour. Johnson (2012) is one of many empirical studies that supports this view. But he also found that the affective empathy aroused by a work of literary fiction is unable to fully explain the effect of reading fiction on prosocial behaviour. Another factor, namely the mimicking of prosocial behaviour, must play a role.

The problem is that fiction need not always set good examples. If it does not, then there is a chance that it would sometimes make people less virtuous. Surprisingly little effort has been made to test the hypothesis that works of fiction with immoral characters, who are treated sympathetically, could lead readers to emulate their behaviour and act immorally. Experimental results in other realms suggest that this worry is not groundless. Representations of violent behaviour on television have been shown to increase violence and antisocial behaviour in test subjects. A meta-analysis of the many studies of the effects of television violence on behaviour concludes that regardless of the ages of the test subjects, there is a strong co-relation between television violence and aggression and antisocial behaviour. The combination of violence with erotica has even worse effects on viewers and leads to “sexual callousness” (Paik and Comstock, 1994, p. 537). A meta-analysis of the psychological literature on violent video games found that exposure to such games was “positively associated with aggressive behavior, aggressive cognition, aggressive affect” (Anderson et al., 2010, p. 167). Ex-



posure to such games is also associated with antisocial behaviour and decreased empathy. These effects are found across cultures. In contrast, prosocial video games lead to prosocial thoughts and behaviour (Greitemeyer and Osswald, 2010).

As already indicated, little empirical evidence is available to test the hypothesis that literary fiction that approvingly represented persons who are engaged in violent, aggressive, or antisocial behaviour could make readers less virtuous. However, given that violent television and violent video games have deleterious effects on empathy and prosocial behaviour, it seems likely that literary fiction that favourably or sympathetically represents immoral characters will similarly be associated with aggressive and antisocial behaviour. This is a concern that has been around since Plato's *Republic*. Plato was deeply concerned that people would imitate immoral behaviour that poets depict. Although Plato is sometimes ridiculed, we should not be surprised if some novels, like television programming and video games, lead to reduced empathy and prosocial behaviour. Ayn Rand's *Fountainhead* and *Atlas Shrugged* have undoubtedly had a deleterious effect on the characters of generation after generation of American teenagers.

Someone might deny that literary fiction has the potential to make readers less virtuous. One could deny, for example, that works that harm character are works of literary fiction. The suggestion that *Atlas Shrugged* is literary fiction is certainly tendentious. It is not a carefully observed, insightful exploration of society or personality. It is "morally incoherent." It is characterized by pontification, bombast, and "a naïve attitude towards history and philosophy that at times can only be described as sophomoric." It has been suggested that it is "an effective rather than a literary novel" (Bertonneau, 2004, p. 296, 298 and 306). In this way, it can be argued that *Atlas Shrugged* is a work of fiction, but not an example of a work of literary fiction that harms character since it is not a work of literary fiction. One might similarly argue that any work that harms character is not literary fiction. On this view, works of literary fiction, by their very nature, express a genuine understanding of society and provide insight into morality. On such a view,

reading literary fiction cannot lead people to become less virtuous. I am sympathetic to this view. Literary fiction will typically be the product of careful observation. Any good observer of society and persons is likely to grasp moral facts. Nevertheless, I am not confident that we can so easily rule out the possibility that some works of fiction, plausibly classified as literary fiction, can harm readers' characters. At any rate, it still seems possible that some works of literary fiction could harm the characters of some readers by modeling immoral behaviour in a positive light.

## 7. Conclusion

The recent psychological literature provides empirical support for *H*, the hypothesis that reading literary fiction makes people more virtuous. At least, reading some literary fiction makes some people more virtuous. The mechanisms whereby literary fiction makes people more virtuous deserve more careful attention. Perhaps such attention will help address the concern that some literary fiction could have a deleterious effect on the characters of some readers.

## Bibliography

- Anderson, C., Ihori, N., Bushman, B., Rothstein, H., Shibuya, A., Swing, E., Sakamoto, A. and Saleem, M. (2010). "Violent Video Game Effects on Aggression, Empathy, and Prosocial Behaviour in Eastern and Western Countries: A Meta-Analytic Review." *Psychological Bulletin* 136(2): 151-73.
- Bal, P. and Veltkamp, M. (2013). "How Does Fiction Reading Influence Empathy? An Experimental Investigation on the Role of Emotional Transportation." *PLoS ONE* 8 (1): e55341. <https://doi.org/10.1371/journal.pone.0055341>.
- Batteux, C. (1746). *The Fine Arts Reduced to a Single Principle*, trans. J. Young. Oxford: Oxford University Press, 2015.
- Bertonneau, T. (2004). "Ayn Rand's Atlas Shrugged: From Romantic Fallacy to Holocaustic Imagination." *Modern Times: A Conservative Review* 46(4): 296-306.
- Black, J. and Barnes, J. (2015). "The Effects of Reading Material and Social and Non-social Cognition." *Poetics* 52: 32-43.
- Currie, G. (1995). "The Moral Psychology of Fiction." *Australasian Journal of Philosophy* 73(2): 250-59.
- Djikic, M., Oatley, K. and Moldoveanu, M. (2013). "Reading other Minds: Effects of Literature on Empathy." *Scientific Study of Literature* 3(1): 28-47.
- Chartrand, T. and van Baaren, R. (2009). "Human Mimicry." *Advances in Experimental Psychology* 41: 219-74.
- Eder, E., Jannidis, F. and Schneider, R. (2010). *Characters in Fictional Worlds: Understanding Imaginary Beings in Literature, Film, and Other Media*. Berlin: De Gruyter.
- James, W. (1891). *Principles of Psychology*. London: MacMillan.
- Johnson, D. (2012). "Transportation into a Story Increases Empathy, Prosocial Behaviour, and Perceptual Bias toward Fearful Expressions." *Personality and Individual Differences* 52(2): 150-55.
- Johnson, D. (2013). "Transportation into Literary Fiction Reduces Prejudice against and Increases Empathy for Arab-Muslims." *Scientific Study of Literature* 3(1): 77-92.
- Greitemeyer, T. and Osswald, S. (2010). "Effects of Prosocial Video Games on Prosocial Behaviour." *Journal of Personality and Social Psychology* 98(2): 211-21.
- Hakemulder, J. (2000). *The Moral Laboratory: Experiments Examining the Effects of Reading Literature on Social Perception and Moral Self-Concept*. Amsterdam: John Benjamins.
- Kaufman, G. and Libby, L. (2012). "Changing Beliefs and Behavior Through Experience-Taking." *Journal of Personality and Social Psychology* 103(1): 1-19.
- Kidd, D. and Castano, E. (2013). "Reading Literary Fiction Improves Theory of Mind." *Science* 342(6156): 377-80.
- Koopman, E. (2015). "Empathetic Reactions after Reading: The Role of Genre, Personal Factors and Affective Responses." *Poetics* 50: 62-79.
- Koopman, E. and Hakemulder, F. (2015). "Effects of Literature on Empathy and Self-Reflection: A Theoretical-Empirical Framework." *Journal of Literary Theory* 9(1): 79-111.

- Mar, R., Oatley, K., Hirsh, J., dela Paz, J. and Peterson., J. (2006). "Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social abilities, and the simulation of social worlds." *Journal of Research in Personality* 40(5): 694-712.
- Mar, R. and Oatley., K. (2008). "The Function of Fiction is the Abstraction and Simulation of Social Experience." *Perspectives on Psychological Science* 3(3): 173-92.
- Mar, R., Oatley, K. and Peterson., J. (2009). "Exploring the Link between Reading Fiction and Empathy: Ruling out Individual Differences and Examining the Outcomes." *Communications* 34(4): 407-28.
- Mikkonen, J. (2015). "On Studying the Cognitive Value of Literature." *Journal of Aesthetics and Art Criticism* 73(3): 273-82.
- Nichols, S. (2004). *Sentimental Rules: On the Natural Foundations of Moral Judgement*. Oxford: Oxford University Press.
- Nussbaum, M. (1990). *Love's Knowledge*. New York: Oxford University Press.
- Paik, H. and Comstock., G. (1994). "The Effects of Television Violence on Antisocial Behaviour: A Meta-Analysis." *Communication Research* 21(4): 516-46.
- Putnam, H. (1978). *Meaning and the Moral Sciences*. Boston: Routledge and Kegan Paul.
- Schellekens, E. (2007). *Aesthetics and Morality*. London: Continuum.
- Seilman, U. and Larsen., S. (1989). "Personal Resonance to Literature: A Study of Reminders while Reading." *Poetics* 18(1-2): 165-77.
- Smith, A. (1759/2002). *The Theory of the Moral Sentiments*, K. Haakonssen. ed. Cambridge: Cambridge University Press.
- Tamir, D., Bricker, A., Dodell-Feder, D. and Mitchell., J. (2016). "Reading Fiction and Reading Minds: the Role of Simulation in the Default Network." *Social Cognitive and Affective Neuroscience* 11(2): 215-24.
- Vogler, C. (2007). "The Moral of the Story." *Critical Inquiry* 34(1): 5-35.
- Young, J. (2001). *Art and Knowledge*. London: Routledge.



---

# Better Worlds and Mark Twain's Submarines: Utopian Literature as a Stimulus for Social Engagement<sup>1</sup>

*Ana Maskalan*

---

## 1. Literary Utopias and Social Change

There is an amusing, though most likely untrue story about Mark Twain and submarines told by Bertell Ollman, a professor of politics at the New York University in his 2005 article *The Utopian Vision of the Future (Then and Now)*. Apparently, Mark Twain was once asked what could be done about the invention of a new, and seemingly very dangerous, weapon – submarines. After a short deliberation, Twain answered: the only way to deal with enemy submarines is to heat the oceans to the boiling point, thus incapacitating the submarines. When questioner persisted on a further elaboration of his proposal, demanding details regarding the boiling of the oceans, Twain allegedly said – *You asked me what we should do; do not expect me to tell you how to do it.*<sup>2</sup>

Ollman uses Mark Twain's submarine analogy with the intention of presenting a Marxist critique of utopias and utopian thinking as producing impossible ideals without giving any explanation about the process that leads to them.<sup>3</sup> This is the way all utopias

---

1 In this paper, as a starting point I have used some insights already published in my book *Budućnost žene: Filozofska rasprava o utopiji i feminizmu* [Woman's Future: A Philosophical Treatise on Utopia and Feminism]. Zagreb: Plejada; Institute for Social Research in Zagreb, 2015.

2 Ollman, 2005.

3 The rich Marxist tradition of utopian criticism begins with Marx's and Engels' views on the attitudes of socialist utopians outlined in *The Communist Manifesto* (1848) and *Anti-Dühring* (1878). There they reproach utopians' ideas on the lack of practical value that manifests itself in their reluctance to carry out

work, claims Ollman. They are distinguished by “a desirable goal drawn from hopes and daydreams, unrealistic means, and ignorance of existing conditions”.<sup>4</sup> And this is how we usually understand the word *utopia*, as a naïve little story, popular with naïve people – daydreamers and wishful thinkers, living in some distant naïve times. Today, utopias are considered old-fashioned and outdated, deprived of the power they were once believed to have – that of making people become socially engaged, of making people change the world.

In this paper, I tell a slightly different story of utopia, a story that, although critical towards many aspects of utopianism, points out positive traits of utopias that are constantly and persistently forgotten. In doing so I primarily concentrate on so-called literary utopias. This is not to devalue other utopian forms and manifestations. However, my choice of literary utopias as stimuli of social engagement is intentional, since I argue that they are the original keepers of the true utopian spirit, a spirit that is, unfortunately, today often under attack, and occasionally lost completely.

---

serious (revolutionary) social changes, as well as in applying the wrong methods and mechanisms of change. See for example *The Communist Manifesto*: “The significance of Critical-Utopian Socialism and Communism bears an inverse relation to historical development. In proportion as the modern class struggle develops and takes definite shape, this fantastic standing apart from the contest, these fantastic attacks on it, lose all practical value and all theoretical justification. Therefore, although the originators of these systems were, in many respects, revolutionary, their disciples have, in every case, formed mere reactionary sects. They hold fast by the original views of their masters, in opposition to the progressive historical development of the proletariat. They, therefore, endeavour, and that consistently, to deaden the class struggle and to reconcile the class antagonisms. They still dream of experimental realisation of their social Utopias, of founding isolated “phalansteres,” of establishing “Home Colonies,” of setting up a “Little Icaria” -- duodecimo editions of the New Jerusalem -- and to realise all these castles in the air, they are compelled to appeal to the feelings and purses of the bourgeois. By degrees they sink into the category of the reactionary conservative Socialists depicted above, differing from these only by more systematic pedantry, and by their fanatical and superstitious belief in the miraculous effects of their social science”. (Karl Marx and Frederic Engels, *Communist Manifesto*, <https://www.marxists.org/archive/marx/works/download/pdf/Manifesto.pdf> (retrieved on June 23)).

4 Ollman, 2005.

The first literary utopia, *De optimo reipublicae statu deque nova insula Utopia*, was written in 1516 by Sir Thomas More, an English Renaissance philosopher, Catholic saint, scholar and great chancellor in the service of King Henry VIII. More used a very simple storyline that many writers of literary utopias later inherited. There was usually the main utopian hero, a traveller who discovers a remote and isolated place, an island or a country. There he meets its citizens and is acquainted with their social, political, economic and religious institutions, their private and public lives and their relationships. Literary utopias usually end with the hero returning home, conveying a message about the existence of an alternative and better society.

Although utopia is of Renaissance origin, many of its attributes were not novel for that period. Utopia was born under ancient and Judaeo-Christian influence, borrowing from the first a polis-like vision of a perfect society, prospering on some remote island, usually surrounded by tall walls. Judaeo-Christian tradition gave further value to the original remoteness of the imagined place by adding a temporal dimension – utopias were often not only spatially but temporally distant as well.<sup>5</sup> Like a Christian paradise, utopias were to be expected at some point in the future, with one crucial difference: unlike in paradise, man had an active, even primary role in the creation of the future with utopia in it. That being said, it should be noted that I do not understand the creation of utopia as some coincidental event happening to one of the most brilliant Renaissance men. Although Thomas More wrote his utopia with the intention of subtly criticizing British society and royal government, leaving his utopian work to be a mixture of satire and fantasy without taking it, or himself, too seriously, he did resonate the spirit of the time marked by the awakening of human confidence in his intellectual and moral abilities to change the world and its future. In other words, utopia became such an immensely popular form because it downright tackled something that was already present in the European societies of the time and that can be described

---

5 Manuel and Manuel, 1979; Kumar, 2003.



as a utopian spirit or a utopian thought. It is unlikely that literary utopia would have brought so much joy in some earlier times, or inspired generation of writers, readers and lawmakers, had they not themselves already nurtured the utopian spirit.

Literary utopias during and after More became immensely popular, affecting the development of other utopian forms, such as utopian intentional communities and utopian social theory.<sup>6</sup> Ernst Bloch was among the first to see utopianism in many myths and fairy tales, art, music and architecture, in social and in revolutionary thought. Bloch is also a philosopher who went the farthest in understanding utopian thought and its influence on the modern world, contributing to the development of utopian philosophy and utopian studies as well.<sup>7</sup>

But what is it that makes utopias, especially literary utopias, socially engaging in the first place? In what follows, I try to explain what I mean by socially engaging and what factors made literary utopias function in such a way. I understand social engagement in a slightly different way from how the notion is employed in the usual sociological and psychological definitions concerned with our degree of involvement in community life. For the purposes of this paper I will narrow this definition, concentrating on those aspects of social engagement that are oriented toward transcending the existent and creating a better life, a better society and a better world. In other words, I am talking about social engagement as a means leading to a social change. I argue that utopias had and potentially still have a capacity for encouraging people to change themselves, the community they live in and the future that awaits them. I also argue, and I am aware that I am not alone in that argument, that the undesirability of utopias in the eyes of many does not stem from their alleged naivety or outdatedness, but arises because they are considered dangerous due to their promise and incitement of social change.

---

6 Sargent, 1994.

7 See for example: *Geist der Utopie*, 1918; *Das Prinzip Hoffnung*, 1938-1947; *Tübinger Einleitung in die Philosophie*, 1963.

So, what is it that literary utopias have that incites (or at least, that used to incite) social engagement? In my work on literary utopias I have detected at least four factors explaining their social capacities. These are *feasibility*, *criticality*, *democracy of authorship* and *seductiveness*. I will briefly explain each one starting with the last.

## 2. Seductiveness

The seductiveness of literary utopias can be traced to their capacity, so eloquently described by Bertrand de Jouvenel, to paint beautiful pictures of everyday life.<sup>8</sup> De Jouvenel refers to the utopian *likeability* and *honesty* that give utopia an advantage over other forms of textual persuasion on social engagement. Immersing in the utopian story – and it is the story part where the persuasion holds its strongest position – is compared to dreaming by de Jouvenel. Dreams, although not real, give us a sense of reality that no political or philosophical elaboration of the best possible worlds can give.

A dream, while less than reality, is much more than a blueprint. A blueprint does not give you the “feel” of things, as if they existed in fact: a dream does so. If you can endow your “philosophical city” with the semblance of reality, and cause your reader to see it, as if it were actually in operation, this is quite a different achievement from a mere explanation of the principles on which it should rest. This “causing to see” by means of a feigned description is obviously, what More aimed at: it is also the essential feature of the utopian genre<sup>9</sup>.

For de Jouvenel utopian description of a daily life is not just an ornament, a literary device to entertain idle readers. Rather, it is a necessary element of incitement since it does not only give its readers the conceptual contours of the imagined system, but the ways in which the system affects the everyday lives of ordinary citizens under it. For some utopian writers, utopia meant a world without private property led by communist ideals. For others, it designated

<sup>8</sup> de Jouvenel, 1965.

<sup>9</sup> de Jouvenel, 1965, p. 437-438.

some scientific community where people lived for knowledge and truth, or a little village where life was in harmony with nature, or a land of carnal pleasures and sexual liberties. Precisely this utopian insistence on "what" and not so much on "how" is what Ollman resents in utopian writers. What can be said in defence of those writers is that the *wishing* part usually precedes the *knowing* part. Thus, before giving answers on *how* to achieve utopia, utopian writers – if they are to be held responsible for giving those answers – should have the right to first and foremost imagine their utopias. In my opinion, the final judgment not only on the quality of the utopian picture but also on the means and ways of achieving it, should not be forced on the writer. The readers are those who decide if a utopian world is worth living and worth fighting for, and the most talented among them can maybe even have ideas of how to build them.

### 3. Feasibility

My next argument regarding the power of literary utopia to incite social engagement concerns its feasibility. Feasibility should not be in any way mistaken for a criterion used in many discussions on the (im)possibility of utopian realisation. Much of utopian criticism is often exhausted on arguments about the utopian impossibility, where such alleged impossibility is regularly considered a justification for their worthlessness. Utopias, it is said, are impossible because usually, there are no existing scientific, technological or engineering means of their realization. The problem with this kind of argument lies in the fact that by using it, we impose on utopia the benchmarks of reality which, by definition, it seeks to transcend. Utopias, especially literary utopias, work with non-existent worlds, sometimes using non-existent preconditions of their own existence. By putting the *impossible* label on the utopias, we are reducing human existence to the immediate, negating the horizons to which human creativity and imagination can lead us. A long time ago, when flying was considered humanly impossible, to dream of flying was a utopian dream. There were also times when

the same could have been said about diving to the bottom of the ocean or walking on the moon. In the meantime, man took off, reached the bottom of the ocean and walked on the moon, proving that the ancient impossibility of utopia imagining such things was nothing more than an empty word that, if one believed in it, one would probably not have flown, reached the bottom of the ocean or walked on the moon. It is precisely the belief in the utopian picture, or, even better, the feeling such a picture produces, that makes people engage, which is why many philosophers and some poets believe in the close relationship between utopia and human progress. In the words of Oscar Wilde:

A map of the world that does not include Utopia is not worth even glancing at, for it leaves out the one country at which Humanity is always landing. And when Humanity lands there, it looks out, and, seeing a better country, sets sail. Progress is the realisation of Utopias.<sup>10</sup>

There is however one criterion that utopia should meet, not only to be considered practical, relevant and socially engaging but to be considered utopia at all. I call it the feasibility criterion. When I say that feasibility is an attribute of literary utopias, what I am in fact suggesting is that, unlike many descriptions of beautiful worlds that have been present in human civilisation and culture from their very beginning, Thomas More's utopia, and many others that followed, didn't leave the responsibility for a good life, a good society and a good future to gods, monsters, forces of nature, or aliens. Utopias give the responsibility over human destiny only to humans themselves. And that is why utopias are feasible. Their feasibility is grounded in the fact that they are man-made i.e. they are dependent on human's engagement with his own and with social life. That is why, in my opinion, Plato's *Republic* is not a utopia and this is why stories of paradise are not utopias. In the first case, the creator of the perfect world possesses capabilities that Plato is convinced ordinary man does not; in the second, the creator is God himself.

As has already been said, literary utopias depend on the his-

---

<sup>10</sup> Wilde, 1891.

torical emergence of the necessary spiritual, intellectual and social prerequisites for a different understanding of human history. In that understanding, the past, the present and the future are possible only because of the human agency – both positive and negative. Utopias' feasibility makes them socially engaging because they appoint the main actor of all events – man. In such a constellation, it is impossible for him to look at the world's and the society's inadequacies externally and disinterestedly, blaming gods or destiny for them. As a sole creator of his circumstances, he is the only one responsible for shaping them, making his engagement with the world and the future desirable or even obligatory. And though this responsibility seems like a divine gift, it is also a great burden, since, we see that today more than ever before, human engagement is not always wise or just and the consequences of such engagement can be terrifying.

#### 4. Criticality

Closely related to the above-mentioned topic is utopia's criticality. Utopia is never only about imaginary worlds or imaginary futures. An integral part of every utopia is its relationship to the present, which is always found inadequate. It can even be said that ever so often, a careful analysis of the existing world's shortcomings has a primacy over the construction of the imaginary utopian world. Even in More's case, one of the most interesting parts of his book refers to his insightful analysis of the social crisis in the country, which finally led him to conclude on the necessity of abolishing private property. In fact, as I mentioned earlier, probably the main reason for writing his book was the ability it gave him, to criticise his government without suffering the sanctions for it.

There is no description of a non-existent world without some form of a critical review of the existent one. This critical review in literary dystopias (descriptions of non-existent worst possible worlds) usually occurs directly, since the main effort of dystopian writers is to make the closest possible connection between the imagined world and the world in which they live, making their

dystopian world a direct consequence of the events in the existing world. In literary utopias, the relationship between the imagined and the existent is usually indirect, since the imagined is considered not to be necessarily a direct consequence of current events in the existing world. This is why I do not agree completely with Marxist accusations of utopian writers for their ignorance of existing conditions. It is true that many of them did not understand or did not want to understand the complexity of injustices, inequalities and unhappiness of the world in which they lived. They even sometimes closed their eyes to the causes of suffering, especially when they themselves were contributing to it. It is also true that many of them did not know, or did not care to suggest, practical ways of dealing with injustices, inequalities and unhappiness. What they did know was that something was fundamentally wrong, even if they could not pinpoint it. Utopias are born out of understanding such existing conditions, no matter how insufficient this understanding occasionally is. To put it even more bluntly, there is no utopia without some form of critique.

Criticality grows out of injustice and dissatisfaction with it. Utopian criticality makes it a tool for social engagement and social change, which is why many proclaim its utility in the modern world obsolete. They usually remind us either of needlessness of change or of human incapability to change anything for the better. The modern crisis of utopian critique is, I suggest, followed by a crisis of political, social and other kinds of critique in contemporary society. In the absence of critique, there is a lack of engagement and consequently, a lack of social change. Without such social change, the world remains the same and man remains the same. Without social change, there can be no progress, no elimination of injustice, no breakthroughs of human ingenuity and no hopes for a different kind of living for those who are suffering. This is why Francis Fukuyama calls the end of history, understood as the end of progress and fundamental change, sad.<sup>11</sup>

---

11 Fukuyama, 1992.

## 5. Democracy of Authorship

It is interesting to notice that the main carriers of the utopian spirit today are social movements which, being occasionally the only critics of reality, carry the power to change it. Social movements are continuing something that utopias started earlier – giving voices to those who were usually unable to speak; hence, utopian democracy of authorship. The utopian ability to pinpoint many of the life's complexities by telling powerful and sad stories of injustice and disempowerment encouraged different voices to use precisely the utopian literature to express their own attitudes, hopes and perseverance. Experimenting with different kinds of living, black, feminist, queer or postcolonial utopias gave insights into the, for the most part, hidden lives and adversities of the oppressed, thus giving hope to the ones living unbearable lives.

In claiming the utopian democracy of authorship, a certain *naïveté* should be avoided at all costs, by keeping in mind two things. The first one pertains to the fact that, throughout history, utopian writers were educated and financially secured white men, who spoke about the injustice from their mostly privileged positions. Like all forms of writing, writing utopias required time, money and, in the words of Virginia Woolf, *a room of one's own* – circumstances, in other words, associated with a privileged life. To claim that anyone could write a utopia sounds very cynical if we know that for the biggest part of the history, the majority could not write or did not have sufficient resources to write anything, let alone a utopia. Even if a woman, a black man or a proletarian did write, what odds did they have of reaching a wider audience, an audience composed mostly of rich white men? Of course, one can argue that the privileged utopian writer could have had an understanding, sympathy or even compassion for the oppressed ones – after all, utopian socialists definitely had such an understanding, making an effort to write a utopia addressing issues of existent social injustices. That being said though, I nevertheless have to agree with Fredric Jameson's claim that, historically, the “view of those who

are oppressed is ontologically more fundamental"<sup>12</sup>, believing that the true understanding of oppression comes only from the experience of oppression.

The second thing that should be kept in mind is that historically, for many writers, literary utopias were not so much a medium for combating injustice, but a means of expressing oppressive views. Quite often, utopian writers built their utopias precisely on the sacrifice of the weak and on the exploitation and suffering of those who were different. Not even Thomas More himself could have imagined a society without slaves. In addition, although throughout history, their inadequate social position remained unchanged, most utopian writers did not only fail to notice the oppression of their female contemporaries, but many of them believed that women were already living in utopia.<sup>13</sup> By making oppression utopian, many utopian writers silenced the voices and hopes of the oppressed.

By the contemporary democratization of utopian voices, women writers (and the other 'Others') entered the world of literary utopias showing all the shortcomings of the existing and imaginary worlds. They did so by giving alternatives to political and other public institutions on one side, and by critically considering everyday private life with all its little intricacies on the other. By doing that they made the utopian worlds, previously often reduced to descriptions of public space and life, even more complex and, in my opinion, valuable.

Paradoxically, modern authors, by paying attention to the elaborate layers of contemporary lives, have shown all the difficulties not only of realizing utopian dreams but of engaging in that realization as well. Defining *others* as goals and never as tools, they questioned many of the historical assumptions of utopian life – one of them being that it necessarily depends on some form of exploitation of others – people, animals, and other living or inanimate creatures. They have also significantly influenced what it means to

---

12 Jameson, 1971.

13 Sargent, 1981.



have a good life and under what circumstances that life is possible. By showing us the world that is becoming increasingly irreparable, they have made utopia ambivalent and sometimes hard to swallow. Their utopias are often sober and dark; they rest on the delay of pleasure and on sacrifice; they describe hardships and struggles; they are global and universal. Proclaiming their work naïve is nothing more than cynical.

The history of utopia so far is based on the dialectic of the utopian and the dystopian. This dialecticism is nullified in contemporary utopias, as the boundary between a worse and a better world has dissolved. Modern utopias are even less of a blueprint than they have ever been, since they are no longer perfect places. It is up to us to decide what kind of places they are and whether we want to live in them. Democracy of authorship does not only rely on the freedom of writing utopias but on the freedom of choosing what utopia is and whether it is worth dreaming. Some would even say that utopia from the start was just that – a nonexistent place that could be better or worse than the existent one and that the audience should have been the only arbitrator of its (im)perfection. By giving everybody that choice, utopias are once again becoming grim reminders of their responsibilities towards the world and the future, again turning attention to the possibility and to the necessity of engaging with them.

## 6. Conclusion

To conclude, in this paper I have pointed out the link between literary utopias and social engagement. I did so by singling out a few traits of literary utopias that I consider crucial to the ability of a utopian text to incite its readers towards social engagement and social change. The seductiveness of literary utopias lies in their ability to immerse readers, by painting beautiful pictures of everyday life, into their imaginary worlds, describing life as it could and should be. Even though many utopian critics claim that literary utopias represent an escape from the woes of the existing, I only partially agree with them. Utopia can represent an escapist

fantasy only during the act of reading. When the reading is over, a necessary return to the real world occurs, followed by the painful realization of a differences between the two worlds. Utopian feasibility reminds utopian readers of their responsibility for creating a better (or worse) world and future. The feasibility criterion refers to the emancipatory move against any kind of authority, celebrating human ingenuity, independence and bravery. Utopian criticality represents an articulation of awareness and an attitude towards the real world. It challenges the logic of order, allowing the critics to move out of their daily lives and re-evaluate them. Utopian democracy of authorship refers to the opening up of a space for different voices, different experiences and different visions. It also points to the complexity of the lives and coexistence of different people that traditional literary utopias often sought to ignore.

The contemporary attack on utopia is, in my opinion, an attack not so much on the usual understanding of utopia as an insignificant backlog from some naive times, but on utopia as a stimulus for social engagement. Such utopia encourages all social groups to criticize the existing world and to offer an alternative one, while reminding them of their responsibility for their own destiny. Repudiation of utopia also means giving up criticism of existing injustices, building alternatives to them, and giving everyone a voice to actively contribute to social change. Recognition of these aspects of the modern attack on utopia opens up a whole range of social and political issues that have long outgrown the boundaries of literary utopia, issues regarding the possibility and necessity of social change in general. I hope that this analysis will contribute to dealing with, and solving, these issues.

## Bibliography

- de Jouvenel, B. (1965). "Utopia for Practical Purposes." *Daedalus* 94(2): 437-53.
- Fukuyama, F. (1992). *The End of History and the Last Man*. New-York: Macmillan, Inc.
- Jameson, F. (1971). *Marxism and Form*. Princeton: Princeton University Press.
- Kumar, K. (2003). "Aspects of the Western Utopian Tradition." *History of the human sciences* 16(1):63–67.
- Manuel, F. E. and Manuel, F. P. (1979). *Utopian Thought in the Western World*. Cambridge: Belknap Press of Harvard University Press.
- Marx, K. and Engels, F. (1848). *Communist Manifesto*. Retrieved from <https://www.marxists.org/archive/marx/works/download/pdf/Manifesto.pdf>
- Maskalan, A. (2015.) *Budućnost žene: Filozofska rasprava o utopiji i feminizmu* [Woman's Future: A Philosophical Treatise on Utopia and Feminism]. Zagreb: Plejada; Institute for Social Research in Zagreb.
- Ollman, B. (2005, July). "The Utopian Vision of the Future (Then and Now)." *Monthly Review*. (retrieved on May 12) <https://monthlyreview.org/2005/07/01/the-utopian-vision-of-the-future-then-and-now-a-marxist-critique/>
- Sargent, L. T. (1981). "An Ambiguous Legacy: The Role and Position of Women in the English Eutopia." In M. S. Barr, ed., *Future Females: A Critical Anthology*, 88–100. Bowling Green: Bowling Green University Press.
- Sargent, L. T. (1994). "The Three Faces of Utopianism Revisited." *Utopian Studies* 5(1): 1–37.
- Wilde, O. (1891). *The Soul of Man under Socialism*. Retrieved from <http://www.gutenberg.org/ebooks/1017>

# (When) Is ‘Engaged’ Art Existentially Authentic? de Beauvoir contra Sartre

David Collins

---

To yield subjectively, not merely to a party machine, but even to a group ideology, is to destroy yourself as a writer. [...] Group loyalties are necessary [in life], and yet they are poisonous to literature, so long as literature is the product of individuals. As soon as they are allowed to have any influence, even a negative one, on creative writing, the result is not only falsification, but often the actual drying-up of the inventive faculties.

George Orwell<sup>1</sup>

## 1. Introduction

This essay examines the account of engaged or committed literature<sup>2</sup> found in Jean-Paul Sartre’s *What is Literature?* (1948; hereafter *WL*) in relation to his existentialism, and specifically his notions of freedom, authenticity, and bad faith. The commitment in question is a social-political one: Sartre is concerned with how writing literature can be a way of actively engaging with the material and social conditions of one’s lived situation, and with how artists can contribute through their works to the political betterment of their age. Because of how he conceives the essence of literature, Sartre argues that writers have an obligation to be engaged in their works, where this seems for him, in brief, to be a matter of advocating for socially progressive positions and critiquing instances of oppression. And many of Sartre’s remarks strongly imply that he requires *good* literature to be engaged in this way, and that works

---

1 Orwell, 1948, p. 463-64.

2 Since the French ‘*engagée*’ in Sartre’s phrase ‘*littérature engagée*’ has been translated into English as either ‘committed’ or ‘engaged’, I will use both terms synonymously here, though I generally favour talk of engagement.

that aren't overly political or that endorse positions other than socialist ones will be bad *qua* literature.

While this account of what it is for literature to be engaged is the most prominent consistent position found in *What is Literature?*, Sartre also frequently makes claims that are in tension with this account and sometimes appear to contradict it. While these claims aren't substantial or developed enough to be considered a second, alternate account of engaged literature, they are in line with Simone de Beauvoir's positions on literature's nature, its value, and its relation to human freedom. One of the things I aim to show in this essay is how reading these remarks of Sartre's together with de Beauvoir's helps to develop an alternative to Sartre's more prominent account. I also aim to show that this alternative account of how literature can be engaged and of its social—as well as moral and cognitive—value is more plausible than Sartre's, both on independent grounds and because it avoids certain problems that his main account faces: specifically, worries concerning his account's consistency, and whether what it calls for would lead authors into what Sartre dubs 'bad faith', which, as the contrary of authenticity, could be considered the 'cardinal sin' of existentialism.<sup>3</sup>

Because Sartre and de Beauvoir both discuss literature, I will keep my focus on artworks that are, or that include, fictional narratives. However, my argument in favour of the conception of engaged art that can be found in de Beauvoir and that is implied in the remarks of Sartre that run counter to his main position could also be applied to artworks in other media and genres, albeit with modification depending on the characteristics of the medium or genre in question. Ultimately, I think that the question of the value of art in general—both in the sense of what makes a work good/better or bad/worse *qua* art, and why good art is valuable to create and to engage with as spectators, readers, listeners, etc.—is closely tied to the particular values that I argue obtain for engaged art on

---

3 Cf. Grene, 1952 on the ethical role of authenticity in existentialist thought. See also de Beauvoir, 1948 on how living an authentically human existence is an ethical matter.

this second account. Unfortunately, it is beyond the scope of this essay to say more about a broader application of the points argued for, and beyond the space I have to give examples of specific artworks that are engaged or in bad faith. This will have to be left as material for future work, for which this essay will hopefully pave the way.

In addition to developing those remarks of Sartre's that are in tension with his primary position and showing how they relate to, and can be supplemented by, de Beauvoir's remarks on literature, another reason for my focus in this essay is that the political content and perspectives of artworks and artists is a currently pressing issue in what Sartre would call my own historical situation as an author, given rising concerns with social justice and resisting perceived systemic oppression. In my experience, a view of what it is for art to be politically engaged that is very close to Sartre's primary account is often assumed by artists, critics, educators, and organizations that fund the arts, with the political or moral value of a work's 'message' often being taken to be synonymous with the work's *artistic* value, i.e. its goodness or badness *qua* art. If this way of thinking about art is flawed, as I argue it is, and if I am right that it is frequently assumed by contemporary artists and art critics, as well as by many in their audience, there is a problem not just in Sartre's thought but in the ways that many currently think about art and its value. And if there is a problem with the conception that artists and professional critics, especially, have of art, these practices are themselves likely to suffer.<sup>4</sup> So, while my primary focus here is on Sartre and de Beauvoir, my argument has strong implications for current discourses surrounding art practices and their relation to social-political issues.

I shall proceed as follows. Sections 2 and 3 explicate Sartre's accounts of human existence and freedom and of the nature of literature, respectively, since these are in the background of his account of engaged literature, which I outline in Section 4, and set up

---

4 On the implications for art practices for a problem in the theory of art that is presupposed by the practitioners, see Collingwood, 1931.

the problems discussed in Sections 5 and 6. In Section 7 I outline Simone de Beauvoir's view of literature and the cognitive and moral values it can realize when written authentically, which involves an alternative account of promoting freedom. And in Section 8 I show that her view is compatible with some of Sartre's better claims which can be developed to give an account of art's potential social value, and of how art can be *authentically* engaged.<sup>5</sup>

## 2. With Great Freedom Comes Great Responsibility: Sartre's Existential Ontology

*What is Literature?* was published initially as six articles in the journal *Les Temps modernes* that were then published together by Gallimard in 1948, with an English translation following two years later. Sartre discusses what writing—specifically, prose writing—is, why and for whom an author writes, and the situation that European writers find themselves in at the time. (The 1967 reprint of the English translation by Methuen contains an appendix, 'Writing for One's Age', that appeared in the June 1948 edition of *Les Temps modernes* but not in the original Gallimard edition.) Sartre's call for literature's engagement in the later chapters must be understood in relation to the theory of writing proposed in the first chapter, which in turn must be understood in the context of the accounts of freedom and responsibility and the ontology of human existence that Sartre develops in his major work of existential phenomenology, *Being and Nothingness* (1943; hereafter *BN*).<sup>6</sup>

Sartre's account rests on a distinction between being in-itself (*en soi*) and being for-itself (*pour soi*). Unlike the rest of being—

---

5 Section 2 gives some of the background necessary to understand the later argumentative sections. Readers already familiar with the basics of Sartrean existentialism will have this background and may wish to skip on to Section 3.

6 The following summary is general enough to make it hard to give non-trivial references to specific places where Sartre makes these points. Interested readers are invited to consult the introduction, chapter two of part one, and the first section of the conclusion of *BN*. Also, the introduction to the excerpts from Sartre in *Existentialism: Basic Writings* (Guignon and Pereboom, eds., 2001, p. 255-75) presents a clear and thorough overview of Sartre's account.

anything that in any sense *is*—conscious beings, or beings who are aware of their own existence, exist in the mode of the for-itself. Not only are they self-conscious but they care about, can take a stand on, and interpret their existence. Everyone's life is an issue for them, in relation to which other beings (people, things, events, etc.) take on meaning and value. And one's life is an issue not just in the sense of having a 'survival instinct' to want to continue living, but in the sense that *how* one's life goes will matter and will be the ultimate purpose for the sake of which particular projects are pursued, actions are taken and choices made, things are valued, etc.

Following Husserl, consciousness for Sartre is *intentional*, in the sense of always being conscious *of* something toward which it is directed. Because intentional directedness is perspectival, conscious beings always occupy a perspective on the parts of being they intend, i.e. their intentional objects. Perspectival consciousness not only intends its objects but intends them *as* something-or-other, i.e. makes them *determinate* objects, which in Sartre's terms is to say that it 'discloses' being. For Sartre, this is how distinctions and meanings are introduced into existence. Somewhat dramatically, he says that consciousness makes distinctions by secreting its 'nothingness' into the world. Put more plainly, the idea is that any determination of something *as* something, and any distinction between it and other things, implies negation; something being an X entails that it is *not* a Y or a Z, and something being a distinct object means that it is self-identical, and that it's *not* identical to anything else.

For Sartre, the world of being-in-itself doesn't contain negation or absence; it is a plenum of mere undifferentiated *stuff* that he describes as superfluous (*de trop*) in the sense both of overflowing and of being contingent, having no reason for being other than that ascribed to it by consciousness. In order for consciousness to introduce nothingness into this world, Sartre argues that consciousness must at its core *be* the nothingness it 'secretes'. Less dramatically, consciousness can never get behind itself to be aware of itself along with its intentional object; one can be *self*-conscious only by reflecting on aspects of what Sartre calls one's '*facticity*', e.g. one's past



actions and experiences and historical facts about oneself, or about one's present bodily condition and situation, by making these one's intentional object. This self-reflection is itself an act of consciousness, with the 'intender' behind such an act remaining outside of consciousness itself, which is to say that in any conscious act, the consciousness 'performing' this act is never an object for itself, which is another way of saying that it is 'no-thing'.<sup>7</sup>

This constitutes a double gap in the ontological structure of being for-itself. Because consciousness needs to be distinct from the objects it intends in order to take a perspective on them, there must be a gap between consciousness, or being-for-itself, and its actual or potential objects, which includes all of being-in-itself. This makes beings-for-themselves free from the causal determinacy and quantifiability that apply to beings in the mode of the in-itself. This freedom from determinism is part of what Sartre calls our '*transcendence*', where this includes our *self-transcendence*. This involves a gap between who we have become at any particular moment in our existence and who we are 'overall', including who we *will* become as our lives unfold temporally and as we continue to develop through our choices and actions. At every moment and in every situation we will have a finite number of determinate factual characteristics, and there will be certain things that will either be true or false about our pasts and our social situatedness—one's class, gender, race, and nationality, one's occupation, one's relations to others, e.g. being a mother or a friend, etc.—but we are never reducible or identical to these aspects of our facticity. As being-for-itself, we *are* never mothers or friends, men or women, rich or poor, etc. in the way that, say, a table *is* a table or a rock *is* a rock. This is what Sartre means by his seemingly paradoxical claim that we are what we are not, and are not what we are (*BN*: 28).

Since the for-itself is not determined in its being, we *are* only what we make of ourselves through our actions and through how we disclose ourselves to ourselves in self-consciousness. For us,

---

7 This is similar to Hume's argument that one never experiences one's *self*; see Hume, 1740, I, IV, § VI.

“existence precedes essence” (Sartre, 1946, p. 20); what we are isn’t something given in advance that causes what we do; rather, who and what we are is shaped by, and follows from, our actions. For Sartre, what we do is always the result of a free choice, even if only insofar as we ‘choose’ it by not choosing to do otherwise; e.g. as long as suicide or passivity is an option, not choosing these means we’ve chosen whatever it is we do instead. This makes us radically responsible for our choices and actions, and insofar as what we *are* is formed by what we *do*, for our very existence. Moreover, because consciously intending or disclosing things is something we do, even if not always deliberately, and because we are not limited in terms of what we disclose things as, but can always conceive of them differently, we are similarly responsible for the meanings and values that things have for us. Without God to give us an essence and in the absence of a biologically or psychologically fixed ‘human nature’, we are left with no excuses.<sup>8</sup>

*Bad faith* is a matter of trying to avoid this responsibility for who we are, either by denying our radical freedom or by denying our past choices and current situatedness. This can either be a matter of denying our transcendence by identifying with some aspect of our facticity that we take to define our being in the manner of a fixed essence or to determine our choices and actions, or of denying our facticity by identifying with our transcendence as itself a kind of object, or as Sartre puts it, taking our being as for-itself in the manner of the in-itself. The first kind of bad faith can be seen in the example of the waiter who identifies with his occupation, taking himself to *be* a waiter in the way a table *is* a table (BN: 102), while the second would be exemplified by someone who denies that their facticity characterizes who they are at all, e.g., by holding that they *are not* a waiter in the way that a table *is not* an inkwell (BN: 103).

*Authenticity*, in contrast, is a matter of living and acting with an awareness of the kind of being one is as for-itself—of acknowl-

---

<sup>8</sup> Along with *Being and Nothingness*, Sartre develops the consequences of our freedom for our responsibility for our actions, for ourselves, and ultimately for the world in *Existentialism is a Humanism* (1946); see especially p. 17-25.

edging ourselves as always both facticity and transcendence—and accepting and embracing our radical responsibility. For example, an authentic waiter will acknowledge that while it doesn't exhaust his identity, he 'is' a waiter in some sense, based on the position he occupies as the result of his past choices, and that this is contingent on his continuing to perform this role with nothing determining how he will perform it or that he will continue to do so. One could also say that authenticity involves understanding oneself as a dynamic process of becoming, or in de Beauvoir's phrase, "a constantly renewed upspringing that is opposed to the fixed reality of things" (de Beauvoir, 1945, p. 212).

Now that Sartre's accounts of human existence, freedom, and responsibility have been outlined, we can turn to his theory of literature and his call for its engagement.

### 3. Sartre on Literature and the Writer's Responsibility

Sartre begins to answer what literature is by inquiring into the nature of writing and the written word. Unlike other arts such as painting or music, the materials with which writers work, i.e. words, are not themselves the focus of the aesthetic interest one takes in a novel or short story; rather, the meaning of the words are. While the colours used by painters and the sounds produced by musicians are things to be appreciated for themselves, words are *signs*: they refer to something beyond themselves, where their meaning lies in the things, people, events, etc. to which they refer (*WL*: 1-4). Or at least this is where the meaning lies in prose writing, when words are used as signs. Any sign is, of course, also a thing, and Sartre distinguished prose from poetry; in the latter, words are treated as things, with poets and readers being interested less in their literal referents and more in their properties, e.g. how they sound when spoken aloud, the syllables and stresses they contain, their homonymous resonances, etc. That is, in poetry words are opaque to the consciousness of both writer and reader (*WL*: 5), but in prose writing words are "transparent" (*WL*: 15), with writer and reader intending 'through' them to the things, events and ideas

they describe and to which, as signs, they point.

As such, prose is utilitarian, a means for communicating ideas, with this communication and not the words themselves being the end aimed at by the practice (WL: 10). “The man who talks is beyond words and near the object,” Sartre writes, “whereas the poet is on this side of them” (WL: 5); i.e. is focused on language rather than on the world to which language refers. This is why Sartre limits his notion of *littérature engagée* to prose, or communicative speech. Language, as a tool to communicate ideas, functions in the manner of other tools by extending our capacities to act. As Sartre puts it, for the prose writer, his words “are the prolongations of his meanings, his pincers, his antennae, his spectacles [...] he is surrounded by a verbal body which he is hardly aware of and which extends his action upon the world” (WL: 6; see also 11). In this way, prose puts both the writer and the reader into contact with the world beyond language: it “tears him away from himself and throws him out into the world” (WL: 7-8).

Another way to explain this distinction between prose and poetry is to note the difference in the criteria that prose writers will tend to employ when it comes to their selection of words and shaping of phrases in contrast with those employed by poets. Prose writers will choose words and structure sentences largely based on what will best convey the idea they want to get across. The aim is to say what one means, and words or phrases that don't contribute to this ought, on this standard, to be eliminated for the sake of concision; writers should, as Faulkner said, kill their darlings. On the other hand, poets will be concerned with more than the literal meaning of words and phrases. A different word with the same meaning will be just as good as its synonym for a prose writer, but for a poet a change in literal meaning might be less important than finding a word with the right sibilant sound or the right number of syllables, or one that begins with a certain letter, etc. This is what Sartre is getting at when he writes that for the prose writer, “it is not first of all a matter of knowing whether they please or displease in themselves, but whether they correctly indicate a certain thing or a certain notion” (WL: 11).

To use words to communicate implies an addressee in principle, even if one's writing is never actually read by anyone else.<sup>9</sup> Communicating with others, even if the writer is addressing a general audience rather than specific others, is a way of acting in and on the world through language, insofar as these others are in the world alongside the writer, at least in principle. Thus, to write is to act (*WL*: 12),<sup>10</sup> over and above the actions of choosing the words that will express what one means, putting them down on paper or typing them on a keyboard, publishing them, etc. These actions are part of writing poetry as well, but prose writers also act in the sense of performing 'speech acts'; in using prose, one "designates, demonstrates, orders, refuses, interpolates, begs, insults, persuades, insinuates", etc. (*WL*: 10), where this is neither intrinsic to nor standard for poetry.

Most crucially, prose *discloses* the things, persons, events, etc. that it points to; writing is "action by disclosure" (*WL*: 13), portraying things from a perspective or under a certain aspect, and as meaningful in some way. In an important passage, Sartre writes: "If you name the behaviour of an individual, you *reveal* it to him; he *sees* himself" (*WL*: 12, my emphasis), where this also applies to naming, describing, or portraying some object or element of the world other than behaviour. Writing reveals the object to readers in a certain way, and so makes this way of 'seeing' available to them. Revealing a new way of seeing something leads to change, not only in how the person to whom it is disclosed experiences it, but possibly in that person's behaviour in the world, insofar as how we conceive the things around us and the relations in which they stand to other things and to ourselves affects what we do with and towards them. Continuing his example of a hitherto-unconscious aspect of someone's behaviour being revealed by being named, Sartre asks:

---

9 To see why this is intrinsic to language use, one might think, following Collingwood, that the writer serves as her or his own 'first reader' insofar as he or she understands what he or she is writing, and so has an idea of what another might understand by reading it (see Collingwood, 1938, p. 247-52).

10 Sartre's phrase is "To speak is to act" (*WL*: 12), but by 'speak' he just means to communicate in prose, whether in writing or by uttering words aloud.

“After that, how can you expect him to act in the same way?” (*WL*: 13).<sup>11</sup>

So, by describing or naming something one inevitably characterises that thing in some way and communicates a view of it as having some meaning or value, where communicating this can change—specifically, by adding to—the ways in which one’s readers think of, relate to, and value it themselves, and so can change how they act with respect to it. How one discloses the things one describes can be more or less deliberate or self-aware, but any communicative use of language will still disclose it in some way or other, with the choice to speak or write rather than remaining silent being one that the author had. Thus, for Sartre, a writer bears responsibility for the meanings he or she gives to things by writing about them, i.e. for how they are disclosed. This is not to say that writers must be sincere and should only characterize things as they believe them to be; one can, of course, be ironic, but one will still be responsible for presenting a view that one advances ironically or insincerely as a possible way to conceive of and value one’s subject, even if one doesn’t endorse that conception or those values oneself.

As explained above, Sartre takes consciousness’s ability to disclose things in multiple ways—to interpret things and situations, including itself, differently—to be an essential part of its freedom. While, for him, our ontological freedom as beings-for-ourselves can never be negated, we can be alienated from it, e.g. in cases of bad faith where we identify with our facticity and deny our transcendence. By disclosing things and the world in new ways, revealing new perspectives along with new interpretations, writers exemplify humanity’s freedom and show us ways in which the world can be meaningful. Thus, literature can liberate readers from alienation and promote their freedom (see Cauter, 1967, p. ix), especially since for Sartre reading is a creative act wherein the reader re-enacts the writer’s speech acts for herself; when we read our own work, Sartre writes, “we create it again, we repeat mentally the operations which

---

<sup>11</sup> On writing being an act of disclosure that reveals the world and things as meaningful, see also *WL*: 26, 43-44.

produced it” (*WL*: 28; see also 29-30). Thus, reading a text allows us mentally to go through the process of meaning-making by which it was produced, and so to re-think for ourselves what its author thought. “To write,” then, “is to make an appeal to the reader that he lead into objective existence the revelation which I have undertaken by means of language” (*WL*: 32).

Just as authors exercise their freedom in writing and so need not to be alienated, readers must be free in order to engage with what authors have written and re-create it in their imagination (*WL*: 32-33). Sartre puts this by saying that the author needs to “address himself to the freedom of readers” so that they in turn “recognize his creative freedom” (*WL*: 36). To do this, the author requires his readers to be free; thus, writing in any way that promotes readers’ alienation rather than their freedom is self-defeating. Sartre insists, moreover, that authors must write so as to actively promote their readers’ freedom: “the end to which [prose] offers itself [as a means] is the reader’s freedom” (*WL*: 33). So, in his view, literature is characterized by two related aims, “both to disclose the world and to offer it as a task to the generosity [i.e. freedom] of the reader” (*WL*: 43), with authors being obligated and responsible for both.

Two worries can be raised here, one about Sartre’s distinction between poetry and prose, and one about the question of disclosure and of literature’s capacity to connect the reader to the world when the objects, people, and events disclosed are fictional. Regarding the first worry, one might object that the difference between poetry and prose as forms of writing is not as hard-and-fast as Sartre makes out. Even if poets are typically more concerned with properties of words than prose writers are, with different criteria for selecting and arranging words, it doesn’t mean they don’t also aim to communicate ideas and never perform speech acts such as asserting, interrogating, confessing, pleading, etc. through their poetry. Likewise, prose writers can also attend to the visual or sonic qualities of words and choose them not only in order to communicate ideas but for reasons of style, even if this is a secondary concern. And while poems typically involve more linguistic self-reflexivity by foregrounding aspects of the words they contain, they can

also disclose objects or the world as meaningful in certain ways: Shakespeare's sonnets, for instance, disclose various aspects of the phenomenon and experience of love.

Even if his distinction between poetry and prose isn't tenable, it doesn't pose a problem for Sartre's position that literature can and ought to be engaged. If poetry can disclose things as well as prose can, it only shows that what Sartre says about a writer's responsibility can apply to poetry as well.<sup>12</sup> This leads us to rethink what Sartre means by 'poetry' and 'prose' if his distinction doesn't plausibly map on to the kinds of writing to which these terms are typically applied. It suggests that Sartre employs them not in their ordinary sense, but as technical terms for ways of using language. This can be seen in his claim that what he means by prose "is first of all an attitude of mind" (*WL*: 11), i.e. one of 'intending' the world *through* language, and so by 'prose' and 'poetry' we can take him to mean ways of approaching and relating to language as 'transparent' or 'opaque', respectively. Hence, a line in what would ordinarily be called a poem that is written to communicate clearly an idea about some object, with little emphasis on the words themselves, would count as a line of 'prose' in Sartre's technical sense.

As for the worry about how Sartre's theory will apply to fiction, one might think that if the objects—characters, settings, events, etc.—that words refer to are fictional, language could no longer be said to put the writer and reader into contact with the actual world, rather than an imaginary one. It becomes less obvious how an author could disclose anything about the non-fictional world by writing about fictional characters and events, since whatever is communicated about these characters or events might only apply within the bounds of the fiction. A film of a botched robbery, for instance, may communicate perspectives on loyalty, greed, self-inter-

---

<sup>12</sup> Christina Howells argues that Sartre came to accept that non-utilitarian or 'pure' art (i.e. an art concerned more with 'form' than 'content'), can also be committed, where this can be seen in his changed position on Flaubert, whom he first condemns for a supposed lack of engagement and for anti-socialist views (*WL*: 92-93), but whom he later counts as a kind of engaged writer in *The Family Idiot* (1971-72); see Howells, 1978, esp. p. 178-81.



est, etc., and while the story may show these perspectives to be true within the fictional world (e.g. maybe there really is honour among thieves—in the fiction) there is no reason to take the story to reveal anything that holds true of any actual person, their behaviour and motivations, or the meanings or values that actual actions and events can have in the world outside the fiction.

This would be a problem if Sartre thought that literature conveyed propositional truths about the world, but his concern is with how it conveys perspectives—i.e. ways of conceptualizing, interpreting, and valuing—where a perspective towards a fictional object can also in principle be taken towards non-fictional ones. Fictional entities or events are instances of general kinds that also have instances in the non-fictional world, and so these kinds can be said to 'be' in the world.<sup>13</sup> In leading the reader to imagine things about an instance of a kind, a text also has the reader imagine something about the kind.<sup>14</sup> This is because we understand particulars in terms of kinds or categories, with every new experience of something that we take to belong to a category potentially adding to or modifying our conception of that category. Sartre is getting at this when he writes that “[j]ust as one perceives things only against the background of the *world*, so the objects represented by art appear *against the background of the universe*” (WL: 40-41, my emphasis). That is, we understand objects represented in fictional works as possible instances of kinds that we have concepts of through our familiarity with some of their actual instances.

Since what is disclosed through the portrayal of a fictional object can pertain to the kind of which it is an instance, if the kind is

---

<sup>13</sup> I remain neutral here on the metaphysical question whether kinds (or types, etc.) are themselves 'entities' or can properly be said to 'exist' if one or more instances (tokens, etc.) of them exist. My point is that fictional objects will be comprehended in terms of categories (kinds, types) that also apply to things in the non-fictional world.

<sup>14</sup> For the view that fictions prescribe the readers to imagine what is true in the fiction, see Walton, 1990, p. 39. Most contemporary philosophers of fiction endorse this view, but for a dissenting view, see Matravers, 2014. Nothing of Sartre's view of literature hinges on Walton being right, so it is not threatened by Matravers's opposing argument.

a part of the actual world then the fiction *does* disclose an aspect of this world, and not only of the fictional one. A portrayal of a fictional action as generous or brave doesn't show that there actually are generous or brave people or actions, but it does demonstrate ways of being generous or brave, where this can add to how a reader conceives of generosity or bravery. Moreover, portraying an event or character in a certain way—as having some value or meaning, etc.—discloses a way in which things of that kind can be understood or valued. Representing a robbery as admirably brave reveals how such activities *could* be seen as brave and how one *could* admire their participants.

Thus, works of fiction can disclose aspects of the world to readers, and depend on the freedom of their readers in order for the author's ideas to be re-created in the readers' imaginations. In both cases a author will be responsible for their choice of what to write—and their choice to write at all—and so will be responsible for what their writing discloses about its subject-matter, e.g. for the ways of interpreting or valuing that it shows to be possible. Broadly speaking, what Sartre calls engagement or commitment has to do with the writer accepting this responsibility. But what exactly does this involve?

#### 4. Engaged Art as Disclosing Social Problems

Since Sartre advocates for writers to engage with their social and historical circumstances, it is useful and appropriate to consider his own context and what he was aiming at by writing, i.e. what change he hoped his writing would help bring about. The war, the Nazi occupation of France, and the resistance were important elements of this context and fuelled Sartre's concern with freedom. Specifically, *What is Literature?* was written against the backdrop of the prosecution of French authors who had collaborated with the occupation, with some being sentenced to death for their involvement, which publicly raised the question of what specific responsibilities, if any, they held as writers (see Sapiro, 2006, p. 40-41). As Gisèle Sapiro recounts, some of these writers denied that they were

responsible, appealing to an art-for-art's-sake view of literature to contest the claim that what they wrote committed them politically. However, at least two accused authors, Robert Brasillach and Pierre Drieu La Rochelle, accepted their responsibility *qua* writers—the former during his 1945 trial at which he was condemned to death, and the latter prior to his suicide a few weeks after—and so became “the negative figures in relation to whom Sartre elaborates his definition of the social role of the writer” (Sapiro, 2006, p. 41).

This suggests that Sartre sought to condemn these authors' failures to engage with the conditions of their fellow French under the occupation, and how they either used their writing to support the occupation and its values, whether directly or indirectly, or made no mention of this situation as it was occurring. This further suggests that for Sartre, engaged literature is writing that *doesn't* remain silent on pressing issues in an author's factual (political, historical, etc.) situation, but instead addresses these issues, and not simply by mentioning or describing them but specifically by disclosing them *as* problems, i.e. by naming what is problematic about them and taking a stance against them. So understood, Sartre isn't just offering a theory of what literature is, but of what *good* literature is, insofar as arguing that authors *ought* to write in a certain way presents a view of how literature *ought* to be written, where such a view seems equivalent to a view of what constitutes good literature: viz. that literature (and by extension, art) is good *qua* literature (or art) to the extent that it raises awareness of and critiques social problems, and bad *qua* literature (or art) if it does not.

This view of good literature as needing to be politically engaged, and of political quietism making a work worse *qua* literature, is easy to attribute to Sartre based on several of the things he writes, and has been noted by others. Charles Whiting, for instance, takes Sartre to be saying that any “novel should imply corrective measures for the solution of current social and political problems” (Whiting, 1948, p. 84), while Iris Murdoch writes that Sartre “makes the connexion of ‘good writing’ with ideological commitment” (Murdoch, 1953, p. 112). Much of *What is Literature?* supports this view, and moreover requires *good* literature's ideolog-

ical commitments specifically to be socialist and progressive, and in some passages explicitly Marxist, though he is careful to point out that he is not calling for the sort of social realist art officially promoted by the Soviet Union.

This Marxist orientation can be seen in his discussion of the development of French literature from the seventeenth to the twentieth century, considered in relation to the class interests of the nobility, the bourgeoisie and the proletariat. According to Sartre, in the seventeenth century an author's readership was to be found in the upper classes and the nobility, with literature serving the interests of this audience by promoting and maintaining the values of this class. "The authors of the seventeenth century," he writes, "had a definite function because they addressed an enlightened, strictly limited, and active public which exercised permanent control over them. Unknown by the people [i.e. the lower classes], their job was to reflect back its own image to the *élite* which supported them" (*WL*: 67, original italics). In the eighteenth century, with the displacement of the aristocracy by the bourgeoisie, literature came to serve the interests of this newly dominant class—or so Sartre's story goes—with writers acting as the 'bad consciousness' of the nobility but failing to give their bourgeois readers "a clearer class consciousness" of themselves (*WL*: 78). By not also acting as *their* 'bad consciousness', limiting their critiques of oppression to that of the old nobility without raising awareness of the bourgeoisie's own oppression of the masses, literature became the 'good consciousness' of this now-dominant class, maintaining their values and implicitly their ideology, thereby becoming complicit in their oppression of the proletariat (*WL*: 82-85, 177-78).

Whether or not this is accurate as history, it reveals Sartre's concern with literature as a means not only of raising awareness of social conditions in which people are living at a certain time, but of giving readers the right sort of 'class consciousness' with the aim of inspiring concrete actions to change said conditions. Moreover, Sartre's faulting of authors *qua* authors for not "giving [their] readers a clearer class consciousness" (*WL*: 78) shows that this is something he thinks authors *ought* to do, and hence that he takes

the value of literature to be tied to its use as such a means. This can also be seen in what he writes elsewhere (e.g. *WL*: 86-87), where he implies that *good* writers will make their works reflect the social order of their era instead of focusing on the psychology and experiences of individuals. That Sartre aims for literature's consciousness-raising to inspire changes in actual social conditions is made explicit when he states that "writing conceived as a concrete and historical phenomenon [...] want[s] the material improvement of [the proletariat's] lot, and ... the end of man's exploitation by man" (*WL*: 90), and when he contends that a writer's strength "lies in his *direct action upon the public*, in the anger, the enthusiasm, and the reflections which he stirs up by his writings" (*WL*: 141, my emphasis), and goes on to declare that "our writings would have no meaning if we did not set up as our goal the eventual coming of freedom by means of socialism" (*WL*: 211).

Further support for this view can be found throughout *What is Literature?* but there is only room here to note some of the more telling remarks. For instance, Sartre defines a bad novel as one that "approves or accepts or simply abstains from condemning the subjection of man by man" (*WL*: 45, my emphasis), with the implication that a novel is good only if it actively condemns this. He goes on to state that "the essence of the art of writing" involves "clarifying and supporting the claims of the proletariat" and that, as such, writers must think of themselves as "united with the oppressed masses by a solidarity of interests" (*WL*: 110). Similar claims are easy to find, e.g. that writers, "by the subject of [their] writing," should "direct [the reader's] attention upon ... the oppressed of the world" (*WL*: 204); that writers "must militate, in [their] writings, in favour of the freedom of the person *and the socialist revolution*" (*WL*: 205, my emphasis); and that writers not only should present solutions for social problems, but that "solutions which are not rigorously inspired by socialist principles" are to be rejected (*WL*: 206).

The view that a work's artistic value—e.g. a novel's literary quality—depends on or is reducible to the perceived value of its political perspective, whether this perspective is explicit or implicit, is at the heart of what has been called orthodox Marxist aesthetics.

In *The Aesthetic Dimension* (1978; hereafter *AD*), Herbert Marcuse outlines the six core theses of this theory, where versions of the second, third, and fourth theses are expressed in or presupposed by many of Sartre's claims. As formulated by Marcuse (*AD*: 2), these are that:

2. There is a definite connection between art and social class. The only authentic, true, progressive art is the art of the ascending class. It expresses the consciousness of this class.
3. Consequently, the political and the aesthetic, the revolutionary content and the artistic quality tend to coincide.
4. The writer has an obligation to articulate and express the interests and needs of the ascending class. (In capitalism, this would be the proletariat.)

These theses aren't, of course, necessarily tied to Marxism—except perhaps for the parenthetical remark about capitalism and the proletariat—but are compatible with other political theories.<sup>15</sup> However, because of Sartre's commitment to Marxism in the period of his life when he wrote *What is Literature?*, it is safe to read his claims as expressing this commitment. Moreover, Sartre doesn't simply endorse the fourth thesis but goes further, holding authors equally obligated to refrain from writing what might seem like 'apolitical' literature of psychological rather than sociological interest, and from dealing with characters' personal problems without tying them into the political problems of her or his society.

Even if art can raise awareness of social problems and endorse social change, and even if the effects of a given work are politically progressive or morally positive, the question remains: why think

---

<sup>15</sup> They may not even be necessary for a committed Marxist to endorse. As Marcuse notes, the position he calls 'orthodox Marxist aesthetics' that endorses these theses arguably "does not do justice to the views of Marx and Engels" on art and literature (*AD*: 11), and he cites Hans-Dietrich Sander (1970) for the view that, for Marx and Engels, "the essence of a work of art [is] precisely not in its political or social relevance" (Sander, 1970, p. 174). See also Marx and Engels (1973) for a view of art that's closer to Kant's or Schiller's than to the views of Marxist theorists such as Brecht, Lukács, Berger, etc. Marcuse himself shows how one can be committed to Marxism as a socio-political theory but not endorse these theses.

that a work will be more *artistically* valuable, i.e. good or better *qua* art, as a result? For what Marcuse calls orthodox Marxist aesthetics, the reduction of artistic to political value is grounded in the positioning of art on the side of the social 'superstructure', seeing artworks as part of a culture's ideology (cf. Marcuse, 1978, p. 1, 12-5). If artworks are manifestations of ideology—with traditional views of aesthetic or artistic value as distinct from political value being dismissed as forms of 'false consciousness'—then the identification of an artwork's positive value with the degree to which it is revolutionary or counteractive to the dominant ideology follows. But, likely because he wants to avoid Marxism's historical determinism due to the incompatibility with his existentialist notion of consciousness's radical freedom, Sartre doesn't ground the link between engagement and artistic value and his claims about the writer's responsibility in the idea of a 'superstructure' that is determined by a 'base' of material conditions and that in turn determines how these conditions are understood, and so this answer isn't available to him.

Instead, Sartre aims to ground his position in the nature of literature itself. He claims there is an "internal relationship between the demands of the lower classes and the principles of the art of writing" (*WL*: 91), where unsurprisingly this relationship has to do with freedom. As a medium of communication, literature allows authors to disclose aspects of the world through their writing. In order to disclose some part of being, one's consciousness must be free in the sense that it is able to go beyond received ideas and fixed or habitual ways of experiencing and understanding what one is disclosing, in order to take up a new conceptual, perceptual, or emotional perspective on it. And in order to communicate this new perspective, language must be free to describe things in multiple ways, e.g. through new metaphors, while remaining comprehensible, rather than being limited to a finite number of descriptions already on hand.

This is what Sartre calls literature's formal autonomy (cf. *WL*: 113), or its freedom from being fully determined by the situation of the author within which a literary work is produced. However,

more than this formal autonomy is needed for literature to realize its essence. Communication is a two-way process that requires understanding on the part of those addressed as well as an utterance addressed to them by a speaker or writer (cf. *WL*: 111-12). Thus, just as writers must be free in their consciousness of the world and in their use of language to communicate what they disclose through this consciousness, readers must be free in order to grasp the new perspectives the author's work conveys: i.e. free to understand how words can refer to objects in ways that go beyond what is already familiar to them and beyond any finite, codifiable set of meanings available outside of and prior to the particular communicative act. This capacity for free understanding is part of the 'transcendence' that humans have *qua* conscious beings-for-themselves in Sartre's ontology, and while for Sartre this is an inherent condition of personhood that can never be fully negated, one can nevertheless be alienated from one's freedom, e.g. in conditions of social oppression, or in the form of 'bad faith' wherein one identifies with one's facticity.

Sartre's argument for the necessity of engaged writing can be summarized as follows. Because writing requires an audience and *free*, i.e. disclosive, writing requires the freedom of this audience as well as of the author, and because socially oppressive conditions limit freedom, anything that contributes to such conditions works against an end that is internal to literature. And, since *not* speaking out against a problem in one's social circumstances counts for Sartre as endorsing it, if only implicitly, any literature that is *not* overtly committed to advocating for concrete social freedoms and opposing social oppression works against its own nature *qua* literature, and so counts as bad literature. This is why Sartre writes that "there is a coincidence not only between formal freedom of thought and political democracy, but also between the material obligation of choosing man as a perpetual subject of meditation and social democracy" (*WL*: 110), and why he insists that a literary work, to be *good* as the kind of thing it is, *ought* to be engaged in the promotion



of social freedom.<sup>16</sup>

## 5. Is Sartre's Account Internally Consistent?

If the argument above—which has been reconstructed from several remarks that Sartre makes throughout the third chapter of *What is Literature?* (see especially *WL*: 51-52, 91-92, 111-13)—is his justification for making artistic value dependent on political value, flaws in this argument will weaken Sartre's call for literature's engagement. Without some intrinsic connection between literature and social freedom, and without appealing to something like the orthodox Marxist reduction of art to ideology, it wouldn't be clear why a work is worse, *qua* literature, for not explicitly endorsing freedom or engaging with the social problems of its day, even if we might still hold an author generally blameworthy as a person, but not specifically as an author, for remaining silent when there might have been an opportunity to do good.

Iris Murdoch raises problems for Sartre's argument here, and by extension for the consistency of his position overall. For one thing, she notes that if disclosure or communication is taken to require both the speaker's and addressee's freedom, and if the speaker is therefore obliged to promote freedom through what is disclosed or communicated, it isn't clear why *all* art and *all* forms of discursive communication wouldn't also have the same obligation (Murdoch, 1953, p. 115). It is implausibly strong to hold that all communication must confront and critique social oppression and explicitly promote human freedom, or else it is flawed *as speech*, and Murdoch takes it to be similarly implausible to require that works of literature do these things, even if literature depends on its readers' freedom in the way Sartre describes.

For another, Murdoch suggests that Sartre's account might conflate two notions of freedom, and also two ways in which liter-

---

<sup>16</sup> In fact, in places his position seems to be even stronger than this, with written works being required to promote freedom in order to count as literature at all; see, e.g., *WL*: 73 on eighteenth century religious literature—"by ceasing to be a free appeal to free men, *it was ceasing to be literature*" (my emphasis).

ature connects one to the world. The latter conflation concerns the point about language being a tool for communication with words being signs, i.e. means to refer to objects in the world. As discussed above, Sartre takes this referential function of language to entail that literature connects authors and readers to the world, and Murdoch notes that he uses this “to support his contention that prose literature is naturally and properly ‘committed’” (Ibid.). The possible conflation here is between language ‘engaging’ the author and reader in the world in the sense of connecting them to it, and a literary work being ‘engaged’ in the sense of addressing or taking a stand on specific elements in the world to which it is so connected, e.g. social problems. We might also suspect a conflation between the sense in which literature could be said to ‘commit’ the reader to the world by referring to it, insofar as this leads the reader to consciously intend this world, and a specifically social or political sense of ‘commitment’, i.e. an obligation to make an effort to solve these problems.

The other possible conflation Murdoch notes is between the notions of ‘freedom’ involved in the claim that “the proper activity of the prose writer is to invite a free and selfless response from his reader” and the claim that literature is required “to commend the cause of freedom for all mankind” (Ibid.). If the sort of freedom that literature appeals to in its readers is the ontological freedom of being-for-itself or what Sartre calls literature’s *formal* autonomy, it doesn’t obviously follow that literary works, in their narrative *content*, must discuss and endorse freedom from social oppression in order to be good *qua* literature.<sup>17</sup> In particular, it isn’t clear why an explicit condemnation of oppression in a story, or a positive representation of freedom, counts as inviting a free response from readers any more than other kinds of content would; or, moreover, why other narrative content not explicitly promoting social freedom or

---

<sup>17</sup> That Sartre is concerned with a commitment to freedom at the level of a work’s content—i.e. the events and actions depicted, the perspectives it represents and endorses, etc.—is clear from certain of his remarks, e.g. his contention that he has “never spoken of anything but the content” of literature in his call for engagement (*WL*: 15).

condemning social oppression would fail to invite this response. As Murdoch notes, there is “no reason to prefer one sort of artistic subject-matter to another” when it comes to addressing one’s writing to the ontological freedom of one’s readers, and that “to suggest that ... there is a *contradiction* involved in lending one’s imagination (as author or spectator) to a work which approves of tyranny is to lean too heavily upon the word ‘freedom’” (Murdoch, 1953, p. 116, her emphasis; cf. *WL*: 46).

While Murdoch doesn’t push this worry further, one could ask why concrete political or social freedom is needed for a reader to have the sort of free response to a work that Sartre requires. If, according to Sartre, we are always ontologically free and can never be unfree but only alienated from our freedom, then in order to maintain that readers who are alienated due to social oppression will be unable to respond in the way that literature requires, one would have to presuppose that oppressive social conditions will alienate a person *entirely* so that she is no longer able to realize *any* of her ontological freedom. If alienation is always only partial, it is unclear why a reader in oppressive conditions couldn’t be receptive to a work’s appeal to her ontological freedom, and why she wouldn’t be able to understand the new perspective that the work discloses and communicates.

If Sartre does equivocate between two senses of the word ‘freedom’ and conflates two ways in which literature could be said to be ‘engaged’ with the world, his argument is internally inconsistent. Also, it can be noted that in places Sartre appears to presuppose that readers must *already* be free from social oppression in order to respond to the appeal to their freedom that literature makes. This can be seen in his insistence that authors should promote social freedom because they need their readers to be socially free and that because their works, as language, aim to communicate to a universal reader or an ‘every-person’ rather than to specific individual others, so they should want the widest possible audience and hence should aim at universal social freedom (see *WL*: 49). And it is especially evident in his remarks to the effect that it is only in a classless society that literature can only fully realize its essence as universal

communication that appeals to, and can be received by, the freedom of all (*WL*: 116, 177).

This leads to another inconsistency: if ‘real’ literature is possible only in a classless society in which everyone is free from alienation and so can be part of the audience to whose freedom an author needs to appeal, the works in which authors need to aim to bring about these conditions can’t themselves count as ‘real’ literature; but once the conditions in which literature can realize its ‘essence’ are achieved, there wouldn’t seem to be a need for it to address and critique social oppression and to promote human freedom, since these issues will no longer be problems. Thus, if literature can only realize its essence in a classless society, it can’t be part of its essence that it explicitly address social problems and advocate for concrete conditions that promote human freedom. Rather, what Sartre seems to be urging the authors of his day to produce seems to be propaganda designed to motivate readers to work against oppression in order that his ideal of ‘real’ literature can become possible. However, this is in tension with his repeated insistence that committed literature is not propaganda (see, e.g., *WL*: 153, 198, 220).

A related worry is that this limits literature’s liberatory potential, since it seems to preclude literature from itself being able to combat readers’ alienation, leaving it to appeal to those who are less oppressed and less alienated and to try to get them to act, outside of the realm of literature, to reduce the conditions of *others’* alienation. Literature’s potential to enhance and promote freedom would be greater if it could communicate by appealing to the inherent freedom that even an alienated reader will have as a being-for-itself, and thereby work to counteract this alienation, making readers more aware of their inherent freedom through the very act of disclosing new perspectives without requiring them already to be non-alienated in order to be receptive. This is more or less what Ralph Waldo Emerson is getting at when he writes in his essay “The Poet” that artists “are free ... and make free” (Emerson, 1844, p. 301), and there is no reason why Sartre couldn’t accept this instead of the presupposition noted above, as it is compatible with his broader account of human consciousness and freedom while al-

lowing for the intrinsic connection between literature and freedom that he wants to maintain.

These apparent inconsistencies in Sartre's position and his argument in support of it are instances of a greater tension in Sartre's thought at the time he wrote *What is Literature?*, between the existentialist position he developed in *Being and Nothingness* with its view of radical freedom, and his political commitment to Marxism with its historical determinism and its view of most peoples' beliefs, thoughts, and actions being caused by material conditions and social forces, or what Marx calls 'ideology'. The notions that literature ought to combat oppression and that people can only truly be free in a classless society are influenced by Marx, but the view of literature as the free disclosure of new meanings and perspectives and the notion that it appeals to readers' freedom in communicating with them align more with existentialism. Sartre struggled in his later life to reconcile these two sets of ideas,<sup>18</sup> and if he succeeded it wasn't until after *What is Literature?* was written, in his *Search for a Method* (1957) and *Critique of Dialectical Reason* (1960).

As noted above in the introduction, there are a number of remarks to be found in *What is Literature?* that are in tension with the account of engaged literature that is most prominent and most coherently developed, which could be seen as a further internal inconsistency in *What is Literature?* These remarks will be discussed in Section 7, below, along with the alternative account of literature found in de Beauvoir, with which Sartre's 'outlier' remarks are compatible. But first I want to note another problem for Sartre's account on his own standards, which is that literary engagement, understood as explicated above, arguably leads to what he calls bad faith.

## 6. Is Sartre's Account in Bad Faith?

Bad faith is usually understood as a form of self-deception about one's own being involving identifying as one's facticity or

---

18 For an early attempt to show that existentialism is compatible with socialist politics and with Marxist principles, see Sartre, 1946.

transcendence while denying the other, but this is not the only way that Sartre thinks a person can be in bad faith. While it is under-acknowledged, certain of his examples show that he takes bad faith to be possible in relation to the being of others, conceiving of and relating *to them* as either pure facticity or pure transcendence. The former—reducing another to her or his facticity—can be called objectification insofar as it takes someone *to be*, say, their class, race, gender, occupation, etc. in the manner of the in-itself, i.e. a thing that has this characteristic as part of its essence, where this determines how they will be and what they will do. An example would be taking a person to be ‘a blonde’ or ‘a nerd’ and thinking that anything about them follows from this. Treating another as pure transcendence, on the other hand, might involve trying to ease someone’s worries by telling them not to be affected by their circumstances, or trying to encourage someone by telling them they can do or be anything they want, where such empty advice overlooks how actions, feelings, and possibilities are always conditioned and limited—though never determined—by facticity.

That Sartre thinks we can be in bad faith about others and not only ourselves is seen in his examples of the woman on a date (*BN*: 96-98) and the closeted homosexual, or rather the friend who wants him to admit to being a homosexual (*BN*: 107-09). In the woman’s case, her bad faith includes her attempt to withdraw from or deny part of her own factual situation, but another part of her bad faith involves her thinking of her suitor in terms of fixed qualities. While his behaviour may be sincere or respectful, her taking *him to be* “sincere or respectful as the table is round or square” (*BN*: 97), attributing these qualities to him and not just to his actions, counts as bad faith on her part towards him; it involves her “project[ing] the strict present of the qualities into the temporal flux” (*Ibid.*), which takes his being to be fixed in the mode of the in-itself, effectively denying that ‘what’ he is will surpass any determinate qualities.

In the case of the homosexual, the bad faith exhibited by the friend whom Sartre calls the ‘champion of sincerity’ is clearly an instance of bad faith regarding another’s being. By urging his friend to admit to being a homosexual in light of his desires and sexual

history, the champion of sincerity “demands ... that he constitute himself as a thing” (BN: 108). Sartre asks “Who can not see how offensive to the Other and how reassuring for me is a statement such as ‘He’s just a [homosexual]’” (Ibid.). This suggests that we might be inclined to be in bad faith about others because regarding them as definable and their behaviour as determinable—and so, more easily predictable and controllable—lets us feel more secure in our interactions with them, seeing them as manageable, if only for our understanding, in ways that a free being won’t be. “The champion of sincerity,” Sartre concludes, “is in bad faith to the degree that in order to re-assure himself, he pretends to judge, to the extent that he demands that freedom as freedom constitute itself as a thing” (BN: 109).

So, for Sartre, to regard others as ‘tokens’ of some type, and hence as definable and explainable in terms of one or more static factual properties, is to regard them in the manner of the in-itself and so puts one into bad faith towards them. For example, conceiving of people in terms of their membership in a social group—e.g. their race, class, gender, etc.—and taking their words and actions to be products of social ‘laws’ or forces, or to express the group’s perspective rather than their own as particular consciousnesses, reduces them to their membership in this group or category and is bad faith, even when the intention is to diagnose and critique forms of oppression that are themselves based on reducing others to this factual quality. Since Sartre’s account of engaged literature calls for writers to do this in their treatment of the characters and dramatic conflicts in their narratives, it follows that engaged literature, so understood, leads to and promotes bad faith, where this can be seen from several parts of Sartre’s account. Most striking is his call for “[e]ach character [to] be nothing but the choice of an issue and [to] equal no more than the chosen issue” during his discussion of what he calls a “theatre of situations” as opposed to a “theatre of characters” (WL: 217).<sup>19</sup> Writing that he hopes “*all literature* will become

---

<sup>19</sup> Cf. his call for writers *not* to focus on the psychology and ‘subjective’ experience of their characters (WL: 86-87).

moral and problematic like this new theatre” (Ibid., my emphasis), where he likely has Brechtian theatre in mind, shows that he takes the reduction of fictional persons to social issues—or specifically, their reduction to the social types they are taken to exemplify, where the issues in question are understood as conflicts between these types—to be the model for how fictional narratives can be socially engaged.<sup>20</sup>

This makes sense of his claim that Vercors’ novel *The Silence of the Sea*, with its ‘realistic’ or psychologically nuanced portrayal of characters on both sides of the war, “lost its effectiveness” in 1942 when France and Germany were in active combat, since what was “necessary” at that moment was to present the characters and sides of the conflict in simplified terms—the German soldier as brutish, the French soldier as noble, etc.—so readers could “be either for them or against them” (WL: 53-54). In other words, Sartre sees the commitment to human freedom in this historical situation as calling for characters to be portrayed in such a way that readers can easily take them in the manner of determinable things rather than persons or beings-for-themselves who will always be more than any type or category in which they could be classified. And the fact that the ‘persons’ here are fictional doesn’t preclude such a treatment of them from counting as bad faith towards others, since, as explained above, fictional or purely imaginary events and characters still refer to some aspect of the non-fictional world and disclose it, i.e. present it in a certain way. Even if characters are not actual people and so aren’t beings-for-themselves, they are still depictions of persons and so present an image of humanity through which an author can disclose some aspect of humanity. Characters that are determinable or that exist in the manner of the in-itself will present an image of humanity as being similarly determinable and reducible to factual elements, where this is what is in bad faith.

It isn’t only a work’s characters towards whom engaged literature, as Sartre understands it, will be in bad faith. By assuming

---

20 A similar call for the reduction of characters to social types can be seen in his remarks on Maupassant (WL: 104).



that readers will fall into certain demographics based on factors like class or race, and that how they will experience and respond to a work—what they can understand of the perspectives it discloses, how it will appeal to them, etc.—will be limited based on such factors, Sartre risks identifying readers with these elements of their facticity and so being in bad faith towards them. When discussing the author Richard Wright, for example, Sartre claims that his novels address only specific types of readers, viz. “the cultivated negroes of the North and the white Americans of goodwill” (*WL*: 58), and that they will be understood by readers of each type in different ways: he writes that black readers will understand more immediately because they will share the same experience but that white readers will understand more abstractly, “only by an extreme stretch of the imagination and by relying upon analogies” (*Ibid.*). While it seems true that what individual readers will get out of the experience of reading a novel will be partly based on the hermeneutic ‘horizon’ each brings to their reading, where this includes their past experiences and where aspects of their facticity will factor into these experiences in some way and to some extent, Sartre risks bad faith by asserting not just that each reader will understand the work in the context of their own personality and past experience, but that there is *a* context in which black readers, *qua* black, will understand it, *a* context in which white readers will, *qua* white, etc. (*WL*: 59). And similar assumptions are made with regard to class when Sartre talks of works addressing a bourgeois or a proletarian audience (see, e.g., *WL*: 74, 92-93, 110).<sup>21</sup>

By holding that authors should explicitly address social problems and endorse progressive political solutions as a way of mo-

---

21 Sartre’s claim that African-American authors can only write literature that addresses racial politics similarly risks being in bad faith: viz. “if an American negro finds that he has a vocation as a writer, he discovers his subject at the same time. He is the man ... each of whose books will show the alienation of the black race within American society” (*WL*: 57). This denies black authors the imaginative freedom to write about other subject-matter, limiting them to expressing this one aspect of their facticity. However, while this might show bad faith on Sartre’s part, it is separate from the question of how engaged literature, understood in a certain way, might itself promote bad faith.

tivating concrete social change, Sartre risks reducing literature to propaganda—or what Collingwood calls ‘magic’, i.e. the calculated arousal of predetermined thoughts and feelings in order to influence action (Collingwood, 1938, p. 66-68)—despite claims of wanting to avoid this (see *WL*: 198, 220). Sartre assumes here that readers can be influenced in this way, and that writers can aim to bring about predetermined ends by using their writing as a means, but this is also bad faith insofar as it takes readers’ thoughts, feelings, and actions to be causally determinable such that one could find, in advance, the right means to cause certain desired changes in their behaviour. Rather than trying to appeal to readers by expressing why one thinks some social condition is problematic and leaving it up to them to agree or to take action, Sartre’s call for writers to *make* their readers adopt certain interpretations and judgments of social issues overlooks readers’ autonomy as disclosers of being for themselves.

This treatment of the reader as one whose consciousness can and should be determined is explicit in Sartre’s claim that in good literature the reader “will be *led by the hand* until he is *made to see* that, in effect, what he wants is to eliminate the exploitation of man by man” (*WL*: 204, my emphasis). This not only contradicts Sartre’s claims not to be calling for literature to be used as propaganda—cf. “in no case can I address myself to his passiveness, that is try to *affect* him, to communicate to him, from the very first, emotions of fear, desire, or anger” (*WL*: 34)—but it also runs up against Sartre’s insistence that it is contradictory to try to fight oppression by means that limit or deny human freedom (*WL*: 213). Treating readers as beings who can be *led* to think or want certain pre-specified things, or as members of demographics or social groups that can be appealed to and who will think, feel, and react *as* members of these groups, furthers the kind of thinking that makes social oppression possible in the first place, viz. regarding others as objects rather than persons. To the extent that what Sartre calls committed literature does this, it promotes bad faith and so is ‘inauthentic’, where this is a problem insofar as denying or failing to acknowledge readers’ transcendence is inconsistent with Sartre’s claim that engaged

literature must be the act of “a free man addressing free men” and his insistence that “any attempt to enslave his readers” will threaten a writer’s commitment to promoting freedom (*WL*: 46).

This tension between remarks that promote or are in bad faith and those that characterize the writer’s commitment to freedom in terms of a recognition of and appeal to the transcendence of readers and characters runs throughout *What is Literature?* In addition to the claims noted above that run counter to the bad faith understanding of engaged literature, Sartre also writes that engaged literature needs to work *against* positing psychological and social laws for people conceived in terms of types (e.g. *WL*: 107-08), pronounces against fictional characters being “transformed into objects ... and states of soul”, whereby what characters *are*—i.e. as identified with some element of their facticity—takes precedence over what they *do* (*WL*: 174),<sup>22</sup> and complains that “[t]he determinism of the naturalistic novel crushed out life and replaced human actions by *one-way mechanisms*” (*WL*: 98, my emphasis).

As with the particular inconsistencies discussed above, this general inconsistency is likely due in part to the tension between Sartre’s existentialist and Marxist commitments.<sup>23</sup> Because of the problems that Sartre’s main account of engaged literature runs into,

---

22 Cf. *WL*: 216—“Let us never envisage their [i.e. the characters’] situation as *factual data*” (my emphasis).

23 Sapiro speculates that Sartre may have formed his view of the writer’s vocation and responsibility in reaction to his grandfather’s influence. His grandfather was a minister who, as Sartre reports in his autobiography *The Words* (1964), conveyed to him the impression that “the world was prey to Evil” and that “there was only one way of salvation: to die to oneself and to the World and contemplate impossible Ideas from the depths of a shipwreck” (Sartre, 1964, p. 178). While this may have helped shape Sartre’s view of writers needing to be distanced from the world in order to reflect upon it *and* having a responsibility to use this position to reveal and speak out against ‘Evil’, it is also plausible that Sartre may have insisted on the writer’s need to be engaged out of a desire that they avoid the bad faith of denying facticity and identifying as pure transcendence that his grandfather’s position exhibits. (See also *WL*: 96 on Sartre’s opposition to an ‘art for art’s sake’ view of artistic autonomy.) Since bad faith is hard to avoid, with any attempt to get out of one kind of bad faith easily falling into the other, it would not be surprising if, wanting authors to avoid the one form of bad faith, Sartre fell into the other.

as outlined in these last two sections, and because his remarks that are in tension with this account suggest a more promising way to conceive of literature's connection with human freedom and how an author can be committed to this freedom in her or his works, in the rest of this essay I propose an alternate account that is in keeping with these remarks of how literature—and by extension, other arts—can be 'engaged' and have positive social value in a way that is existentially authentic, i.e. not in bad faith. This alternate account develops Simone de Beauvoir's ideas about literature and its value, where these ideas complement and help to flesh out the claims of Sartre's that I take to be preferable.

## 7. Another Take: de Beauvoir on Literature as Discovery

Both accounts of what it is for literature to be engaged see it in terms of a commitment to promote human freedom; the difference lies in what they take promoting freedom to involve. As explained, Sartre sees this as a matter of authors being committed to progressive, socialist, and democratic political perspective, using their fiction to endorse this perspective and speak out against social problems through how characters are represented, what the dramatic conflicts are, the theme or message 'behind' the story, etc.<sup>24</sup> In contrast, de Beauvoir sees it in terms of writers *exemplifying* and *expressing* human freedom or transcendence through *how* their work discloses its subject-matter, and through the experiences the work offers readers. In other words, where Sartre calls for writers to be committed to writing *about* freedom, de Beauvoir calls for writers to practice it themselves in their writing and in how they relate to readers through their works.

De Beauvoir does not advance a theory of political engagement in literature *per se*, but does offer one of how a work can appeal to

---

<sup>24</sup> As noted in the introduction, I take this position not only to be found in Sartre but also to be a view of what it is for artworks to be politically engaged that is common among artists, audiences, and teachers and students of the arts today, to the point where it likely has come to seem like 'common sense'. Assuming I'm not off-base here, readers should at least recognize this view as familiar, and perhaps will have absorbed a version of it themselves.

its readers' freedom and of how this is tied to its value as literature, which I will argue has a political element. In two essays, "Literature and Metaphysics" (1946; hereafter *LM*), published in *Les Temps modernes* two years before Sartre's articles on literature, and her contribution to a panel discussion on the question "What Can Literature Do?" (1965; hereafter *QPL*), she defines literature as a form of disclosure. In the latter she writes that it is "an activity carried out by human beings, for human beings, with the aim of unveiling the world for them, and this unveiling is an action" (*QPL*: 73; cf. Moi, 2009, p. 191),<sup>25</sup> and in the former she concludes that a novel, when "honestly read, and honestly written, provides a disclosure of existence in a way unequalled by any other mode of expression" (*LM*: 276). Like other arts, literature reveals aspects of the world, but how it communicates these to readers is not by conveying information or propositional truths. Rather, a literary work presents the world as "a detotalized totality" (*QPL*: 76; cf. Moi, 2009, p. 192), i.e. as a whole that is only ever apprehended from a particular situated perspective but that always exceeds this perspective.

This distinguishes literature from other forms of communication, such as the essay.<sup>26</sup> Whereas essays are suited to convey abstract general ideas by telling them to the reader, i.e. presenting propositions that the reader is meant to accept (though is of course free to reject), a novel puts readers "through [an] experience of things and events in imagination" in which "they exercise their freedom to judge, interpret, and react" to a greater degree, which Mary Sirridge takes to be "a more radical appeal to [the reader's] freedom" (Sirridge, 2003, p. 132). Through the creation of imaginative experiences for readers to undergo, de Beauvoir sees literature

---

25 If 'unveiling' or disclosing is itself an action, and if every human action is always carried out within a social, intersubjective context, then *how* it is carried out falls within the scope of the political and the ethical, and so is open to political and ethical evaluation independently of *what* is unveiled or disclosed.

26 Note the difference between this and Sartre's conflation of all forms of communicative writing under the label 'prose', and the greater nuance of de Beauvoir's position. Sartre's classification doesn't allow him to distinguish between how fictional works and, say, journalism communicate, where this may explain his idea of engagement as reporting and commenting on social issues.

as uniquely positioned to present lived experience in its thickness and ambiguity and so to disclose what could be called the concrete meaning of such experience and its objects,<sup>27</sup> which isn't generalizable or reducible to abstract concepts, and so can't fully be captured in any finite set of propositions or expressed in any other way. To realize this potential, a writer must present her fictional world, including her characters, their actions, etc., with the same ambiguity that actual existence and experience have, as "a thick and substantial world open to alternative interpretations" (Sirridge, 2003, p. 147) rather than a 'thin' world of definite yet general meanings and of characters with fixed essences. As de Beauvoir puts it, the writer "must attempt to present [reality] in its integrity, as it is disclosed in the living relation that is action and feeling before making itself thought" by "evok[ing] the original upspringing of existence in its complete, singular, and temporal truth" (*LM*: 274-75).<sup>28</sup>

De Beauvoir takes this potential for disclosing the world as it is experienced—as an ambiguous tangle of relations that we are required to make sense of, rather than something that comes to us already interpreted and explained—and for presenting objects in their concrete meaningfulness to be literature's primary value. This value could be called 'cognitive' in a broad sense that doesn't only pertain to reason and propositional knowledge but which can include our capacities as feeling, perceiving, and interpreting beings.<sup>29</sup> By presenting readers with thick experiences to respond to

---

27 See *LM*: 270—"In the real world, the meaning of an object is not a concept graspable by pure understanding. Its meaning is the object as it is disclosed to us in the overall relation we sustain with it, and which is action, emotion, and feeling."

28 De Beauvoir is drawing at least partly on Henri Bergson here; see Bergson, 1889, p. 164, 185-89. Cf. de Beauvoir, 1945, p. 212 on the authentic human subject as a dynamic process of becoming, "a constantly renewed upspringing that is opposed to the fixed reality of things".

29 The discussion of the cognitive value of art and literature, at least within the analytic philosophical tradition, has been dominated by the narrower conception of 'cognitive' as pertaining to knowledge—and with 'knowledge' mainly being understood in terms of propositional knowledge, or 'knowledge-that'. See Stolnitz, 1992 for the paper most responsible for setting the current agenda for this discussion.

and make sense of, where this can be done only by staying in and working through the experience and not by stepping outside it to apply a theory to explain it,<sup>30</sup> literary works model a way of engaging with experience and show how things can be meaningful other than by signifying ideas. As de Beauvoir says, with good literature “[t]he reader ponders, doubts, and takes sides; and this hesitant development of his thought *enriches him in a way that no teaching of doctrine could*” (*LM*: 270, my emphasis).

She insists that the kind of thick imaginative experience that can realize this value only emerges when the author, while writing, goes through this same process of developing his thinking: “the novelist himself [must] participate in the same search he has invited his readers on” (*LM*: 270). For both reader and author this process must be “an adventure of the mind” (*LM*: 272), i.e. one of discovering new directions of thought and new ways of seeing, feeling, and understanding, rather than finding a vehicle to convey a pre-established idea or perspective or communicate something already known, where this is necessary for a work to disclose the world in the way that gives literature its value. To disclose *new* meanings through his writing, “the author must constantly confront his sketches [i.e. pre-existing plans and intentions] with their realization”—with the potential for these ‘sketches’ to be modified as the result of this confrontation—and “as the story unfolds, he [will see] truths appear that were previously unknown to him, questions whose solutions he does not possess” (*Ibid.*). This suggests a discovery-oriented model of artistic creation and experience, according to which the meaning expressed in or disclosed by an artwork isn’t available prior to its expression or disclosure but emerges through these processes. This is because concrete meanings aren’t separable from the objects and experiences that bear them and are always particular and occasional; as de Beauvoir writes, “a smile is indistinguishable from a smiling face, and the meaning of an event in-

---

30 For a similar perspective see Sontag, 1966. The sort of bottom-up understanding this calls for, as opposed to the top-down application of a theoretical framework, is in the spirit of what Kant (1793) calls reflective judgment.

distinguishable from the event itself” (*LM*: 275).<sup>31</sup>

Accordingly, any literary work that is used as a vehicle to illustrate or deliver an already worked-out idea will fail to realize the distinct communicative potential of literature, operating in a mode closer to that of the essay, with the writer functioning more like what de Beauvoir calls a ‘theoretician’ than a novelist proper. This clearly contrasts with Sartre’s call for engaged literature to convey a political perspective that is already worked out and to which the writer is already committed, and to promote preconceived solutions to already diagnosed social problems. As de Beauvoir writes, literature’s distinct mode of disclosure “is not a matter of exploiting on a literary plane truths established beforehand on the philosophical plane, but, rather, of manifesting an aspect of ... experience that cannot otherwise be manifested” (*LM*: 274-75), where one could just as well read ‘political’ for ‘philosophical’. Likewise, she argues that fictional characters should not “be fashioned, a priori, out of a heavy reliance on theories, formulas, and labels,” which, notably, she links to “a certain measure of bad faith” (*LM*: 271).

Not only is it inauthentic to use characters and events as vehicles for illustrating theoretical positions, but it works against the reader’s freedom. “[I]f in advance [a writer] predicts the conclusions to which his reader must come,” she writes, “if he ... pressures the reader into adhering to pre-established theses, if he allows him *only an illusion of freedom*, then the work of fiction is only an incongruous mystification” (*LM*: 271, my emphasis). Instead, a genuine appeal to readers’ freedom will treat them as other autonomous disclosers of being, capable of working through the ‘thick’ or ‘ambiguous’ experience of a complex narrative event and making sense

---

31 De Beauvoir ties concrete meaning to the metaphysical view that “appearance is reality, and existence is the support of essence”, in contrast to views that, “separating essence from existence, disdain appearance in favor of the hidden reality” (*LM*: 275). The former view could be called broadly pragmatist, and includes existentialism with its insistence on existence preceding essence, while the latter could be called broadly Platonist. See Carney, 2000 for a similar distinction between what he calls pragmatist and idealist accounts of meaning in fiction, specifically in relation to cinema.



of it themselves, and not as dependent on being given an explanation or interpretation in order to know what to think or how to feel about what they are being given to imagine.<sup>32</sup>

This is not, of course, to say that appealing to the reader's freedom is a matter of the work being a hermeneutic free-for-all with artworks being equivalent to Rorschach-test ink blots. This also falsifies how people and events are, insofar as the objects of lived experience, while 'ambiguous' and multivalent in what they mean, don't support *just any* interpretation. The idea that a writer should interpret the work for its readers and give them its meaning—e.g. tell them its political significance or moral status—and the idea that the meaning of a work is something each reader chooses for him- or herself, are *both* in bad faith. The former takes the reader's potential for transcendence and disclosure to be limited, while the latter treats readers as if they were purely transcendent and unconstrained by the particular factual aspects of a work in their understanding of it. Instead, a genuine appeal to readers' freedom—one that treats them authentically—will present them with an experience from one or more perspectives that disclose characters and events as meaningful in some way, but with these meanings not being presented as definitive to be accepted and held onto, but rather as to be 'entertained', i.e. for the reader to experience, reflect on, and make sense of how one might think or feel them to be true from within that perspective; that is, not to *tell* them *what* things mean, but to *show* them *how* things can be found meaningful.

The way literature can be most cognitively valuable, then, is not a matter of the 'content' it discloses but lies in the experience of taking up and 'entertaining' them, which exercises one's freedom to transcend any given perspective and always see things from another angle. This could also be considered ethically valuable insofar as authentic being-with-others is at the core of any existentialist

---

32 Cf. *LM*: 270—"The theoretician wants to compel us to adhere to the idea that the thing and the event suggested to him. Many minds find such intellectual docility repugnant. They want to retain their *freedom of thought*; they like instead a story that imitates life's opacity, ambiguity, and impartiality."

ethics (see de Beauvoir, 1948; Grene, 1952), and insofar as authors and readers must relate to one another and to a work's characters authentically in order to write and read in a way that allows for this value to be realized. An author must not be in bad faith towards readers or characters in the ways discussed above, and readers must approach characters as individuals to understand, if not empathise with, and must not expect the work to tell them what to think and how to feel about the experiences it gives them, while also refraining from approaching the work as something onto which they can project any meaning they want.

This understanding of how literature can appeal to readers' freedom gives a clearer and more plausible account of how literature can promote freedom and reduce alienation—e.g. by helping readers avoid bad faith towards others or towards themselves—than Sartre does. But while literature, written and read authentically, can be cognitively and ethically valuable on this account, one might wonder how this counts as literature being politically engaged or committed, since it prohibits good (i.e. authentic) literature from being used to advance a pre-established political agenda<sup>33</sup> or to convince the reader to agree with a predetermined perspective, rather than presenting perspectives on political matters to be 'entertained' but not necessarily endorsed.

## 8. Towards an Authentic Understanding of Engaged Art

As already noted, de Beauvoir was interested in literature's potential cognitive and ethical value—though these aren't her terms—and didn't intend to offer a theory of politically engaged literature as did Sartre; however, there is still a political dimension to be found in her account. Specifically, the aspects of literature that are cognitively and ethically valuable on her account have to do with the sorts of relations that can be formed between authors and

---

<sup>33</sup> Cf. *LM*: 272, where de Beauvoir writes that one can't write authentic literature "if one limits oneself to disguising a *preconstructed ideological framework* in a fictional ... garment" (my emphasis).

readers, and with an author's or reader's ways of relating to characters as persons: viz. by regarding others as beings-for-themselves who are always both factual and transcendent, both situated and free, and so avoiding bad faith towards them (the ethically valuable part), and by imaginatively adopting other perspectives and experiencing how things can be meaningful or valuable from another standpoint, thereby exercising one's ability to transcend one's own fixed perspective in thought, feeling, and imagination (the cognitively valuable part). Because they intrinsically involve relations with others, these same aspects of literature can also be considered to have *social* or *political* value to the extent that they exemplify socially or politically positive interpersonal relations. Moreover, these values are realized together: one can't transcend one's current standpoint to take up and entertain another's perspective without (i) being a factual-and-transcendent being, (ii) recognizing the other whose perspective is entertained as also a factual-and-transcendent being, *and* (iii) relating to this other through an authentic form of interpersonal relation that is socially and politically positive insofar as it is based on a mutual recognition of each person as a being-for-itself.

This third way in which literature can be valuable can be understood in terms of the idea of solidarity, which highlights its social-political dimension. 'Solidarity' here shouldn't be taken to imply an endorsement of the positions or support for the goals of those with whom one is in solidarity, but should be understood more broadly in terms of seeing others as beings who are fundamentally like oneself in some relevant way. In what Richard Rorty calls the concept's traditional philosophical understanding, it's a recognition "that there is something within each of us ... which resonates to the presence of this same thing in other human beings" (Rorty, 1989, p. 189). For de Beauvoir, literature not only presents readers with new perspectives but presents them as tied to another consciousness, differently situated than us, with which we are nevertheless able to identify by taking up these perspectives ourselves.<sup>34</sup>

---

34 As with 'solidarity', talk of 'identifying' shouldn't be taken to imply approv-

Through reading, such a perspective “becomes mine without ceasing to be other. I give up my ‘I’ in favor of the ‘I’ of the person who is speaking and nonetheless I remain myself” (*QLP*: 82). As Sirridge puts it, “[w]hen I read literature ... I remain perfectly aware that I am not Kafka or Balzac; yet, I adopt the novelist’s situation, so to speak, from the inside out” (Sirridge, 2003, p. 131).

This co-inhabiting of a perspective that we recognize as at once both our own and another’s makes us aware that we share a mode of being or ‘form of life’ with anyone whose perspective we can adopt in this way. As such, it helps us overcome the gap that separates our consciousness from those of others and connects us in a more fundamental way than what is often referred to as ‘empathy’, which is more likely to involve having a feeling that is qualitatively similar to but numerically distinct from another’s feeling than co-inhabiting a perspective. And solidarity of this sort is a politically valuable way of relating to others insofar as the recognition that they are beings like oneself is a minimally necessary condition for any other politically or socially valuable relation or activity,<sup>35</sup> and insofar as the absence of this recognition in the objectification or dehumanization of others is at the root of most or all forms of oppression. Moreover, while we can become aware of our solidarity with others in an abstract, intellectual way, e.g. by reading philosophical arguments for the claim *that* all humans share a form of life, what literature and other arts are uniquely positioned to do is allow us to grasp this concretely and feel it, letting us know *what* it’s like to share a mode of being with others.

On this account, literature that is written from an authentic commitment to human freedom will realize this in the way the author relates to readers through the work, and in the way he or she

---

al or sympathy, only understanding. Arguably, such understanding is a necessary basis for any judgment like approval or sympathy—or disapproval or condemnation—to be warranted and made responsibly. In other words, *ethical* critique must start from a position of ‘good faith’ understanding.

<sup>35</sup> This is analogous to the way that de Beauvoir, in her ethics, takes freedom to be minimally necessary for any other ethical choice or action; see de Beauvoir, 1948, p. 29-34.

conceives and presents characters and their actions, with both of these being done in good faith and with writing in this manner *itself* counting as a socially or politically valuable act. Unlike Sartre's account, this doesn't require works to address overtly political matters—class conflicts, social prejudices, etc.—or to be about the political situations or problems of the societies and times in which their authors are writing in order to be good *qua* literature. This requirement is most likely tied to Sartre's understanding of prose writing as 'transparent', as connecting readers to the world by directing their awareness to it through the written word. Sartre's focus on the ideas or objects that words refer or point may have led him to assume that literature's engagement must also involve the work's content, with a work that promotes freedom needing to be *about* freedom as its subject-matter. However, instead of taking the author's commitment to be a commitment to write about freedom, de Beauvoir's account takes it to be a commitment to write in a way that itself realizes and promotes freedom, which is what gives it its social value.

Literature that is authentically engaged, then, will be written in a spirit of discovery and exploration without the author selecting in advance a message for the work to convey or deciding how the subject matter will be disclosed and what it will mean, but instead having these things emerge during the development of the author's thinking, feeling, and understanding as she goes beyond her preconceived notions and positions in the process of writing to discover a new perspective on her subject-matter. By writing in this way, an author will avoid imposing a meaning or interpretation on the narrative's events, its characters, or the reader's experience and understanding, and so can disclose for both herself and her readers genuinely new ways in which things can be understood or valued, and moreover will do so in a way that is ethically and socially positive in that it recognizes readers' freedom and doesn't reduce them, or the work's characters, to their facticity or treat them in the manner of the in-itself.

Furthermore, on this account it will be inauthentic for authors to use their writing to advance pre-existing agendas or to illustrate

---

pre-established theories, whether philosophical, psychological or political. Doing so risks being in bad faith toward one's characters and readers and limits the ways in which one's work can be valuable. Treating the reader as one who is expected to passively receive and accept the ideas presented by a text, or as a means to social change by trying to motivate her or him to take certain kinds of actions, is an ethically flawed and politically problematic way of relating to another person regardless of how positive the ideals are in the name of which this is done; propaganda for a good cause is still propaganda, and so fails to respect autonomy to some extent. What is more, a work will be artistically worse on this account insofar as its presentation of characters, events, and experience relies on clichés or received ideas and is thereby artificial, lacking the thickness and particularity of actual persons, events, and lived experiences and presenting them as abstractly rather than concretely meaningful.

This doesn't mean that authors can't authentically write about explicitly political issues, including presenting perspectives on social phenomena that disclose them as problematic; the political dimension of life is as much a part of human existence as any other and so can't be overlooked without risking bad faith by denying a factual element of many, if not all, experiences. But what might an *authentic* politically-themed work look like on this account?

Situations with overtly political aspects, e.g. those that exemplify recognized social problems, will be presented first as concrete experiences as they are felt and lived through by their participants, and only secondarily as 'political' in a theoretical sense. In other words, existentially authentic political art will deal in the 'raw materials' out of which political issues are theorized, so to speak, as they are prior to their 'processing' in this theorizing. Characters won't be presented as representatives of general types or mouthpieces for ideological positions, and situations, events and actions won't be presented to illustrate general or abstract theories. Instead, characters who hold or act on political beliefs will be presented as fully realized, multi-dimensional human beings who are more than just their political stances or social groups, and situations or events with an explicit political dimension will be occasions for the writer

to express something of what these situations or events are like for their participants, giving the reader a perspective on them from within rather than a diagnosis of them from on high, where this will involve undergoing an imaginative experience that is always partly ambiguous or uncertain with what the reader thinks and feels developing through the experience rather than being predetermined and dictated.

Likewise, and most importantly, authentic political art will be treated as an occasion for the artist to explore and develop her or his thinking and feeling about the political aspects of the 'content' presented in the work, e.g. getting clearer on, and disclosing something new about, why exactly some issue or situation is or is not problematic, rather than as an occasion to present her or his existing perspective on the issues. In other words, authentically political art will be part of the *development* of the political thinking of both the artist and the audience—where the direction in which this development will lead can't fully be foreseen in advance—instead of merely reporting on, reflecting, or appealing to the current state of their political understanding.

The account just outlined has a number of advantages over Sartre's. Most obviously, it avoids the problems of bad faith and internal inconsistency discussed in Sections 5 and 6, and it does so while being compatible with, and even helping to make sense of, the remarks that Sartre meant to be part of his view of engaged literature but which are in tension with what comes across most clearly as his main account. These remarks include, for example, his claim that "a writer is committed when he tries to achieve the most lucid and the most complete consciousness of being embarked" (WL: 56), with 'embarked' meaning situated and responsible. This takes the engaged writer to be committed to expressing, authentically, the perspective taken on that situation, i.e. what the situation is like for the one in it and conscious of it, rather than being committed to endorsing a political position or ideology.

Sartre also writes that engaged literature is a matter of "discovering new countries of the mind" rather than repeating "common-places" (WL: 68) —which fits with de Beauvoir's emphasis on writ-

ing as discovery but is in tension with a theory of engaged writing as endorsing pre-established positions—and that “[i]nsincerity [i.e. bad faith] begins when the artist wants to ascribe a meaning to his misfortunes, a kind of immanent finality” (WL: 233), which again fits more with de Beauvoir’s account than with most of Sartre’s other remarks. Moreover, he writes explicitly against bad faith in the depiction of a work’s characters when he writes that people, including fictional people, must be understood as “absolutes, inimitable and incomparable” rather than as representatives of general types (WL: 234). And his claim that literature is “alienated” when it loses consciousness of its autonomy by “submit[ting] to temporal powers or to an ideology ... when it considers itself as a means and not as an unconditioned end” (WL: 113, my emphasis), is in line with de Beauvoir’s view that works of art lose much of their value when used to illustrate a preconceived theory or ideology, but is again in tension with his main account. Furthermore, much of what is contained in the appendix to *What is Literature?*, “Writing for One’s Age”, when it is read in tandem with de Beauvoir’s two essays on literature, can be seen to be working towards the kind of account of authentically engaged literature that I have argued for, and away from the position he takes earlier in the book, as can his later views on commitment which are seen, for instance, in his changed position on Flaubert (for more on which, see Goldthorpe, 1992).

Finally, it is a virtue of the de Beauvoirian account that it connects literature’s artistic value with its social-political, moral, and cognitive values so that the questions of why the latter values will count in favour of a work’s literary quality, and why the failure to realize these values will count against its quality *qua* art, no longer arise. The connecting link here is authenticity in the existentialist sense, which is held both to be a criterion of artistic success and to be a necessary condition for the ways that literature can be cognitively, ethically, and socially or politically beneficial. A work that presents events and experience with the thickness and ambiguity of actual human actions and of lived experiences as seen and felt from within, that presents its characters as both factual/situated and transcendent/free, and that presents the meaningfulness of



events and experiences as concrete and qualitative rather than abstract and conceptual, will be better artistically insofar as it is more complex, nuanced and concrete, and insofar as it gives the reader a vital imaginative experience to undergo rather than a series of ideas to grasp. For the same reasons, such a work will be more cognitively valuable, giving the reader a deeper understanding of genuinely new perspectives that disclose the world in new ways and that make available new ways in which things can be meaningful or valuable, while also putting the reader and author into an ethically and socially positive relation of solidarity that involves the mutual recognition of each other's freedom.<sup>36</sup>

## 9. Conclusion

This essay has argued for an alternative to Sartre's account of engaged or committed writing, using de Beauvoir's understanding of literature as able to disclose new perspectives and meanings that go beyond the author or reader's familiar horizon and habits of thinking and feeling, at least when writing and reading are pursued as processes of discovery that allow for the development of thoughts, feelings, and understanding.<sup>37</sup> For Sartre, literature is engaged when it addresses political issues and advocates in favour of social conditions that promote human freedom and against conditions of oppression that alienate people from their inherent freedom, and he equates engaged literature with *good* literature and sees literature that is not engaged in this way as artistically flawed. In contrast, on the alternative account, literature best promotes human freedom when it *doesn't* illustrate pre-established theories or make points that an author is already committed to but when

---

36 See also Baugh, 1988 on the existentialist idea of authenticity as connected to both the artistic and political values of artworks.

37 While the focus has been on literature, this account can be extended to other artforms insofar as they also can make available genuinely new ways of perceiving, feeling, understanding, and experiencing and thereby can help their audiences transcend their perceptual, emotional, cognitive and evaluative habits.

it presents thick, ambiguous experiences for the reader to work through and make sense of, since this itself both instantiates and exemplifies a positive form of interpersonal relation, viz. authentic being-with-others. On this account, the use of an artwork to illustrate ideas or deliver messages, however positive the ideas or noble the messages, counts as an artistic flaw in the work *and* makes it less potential ethically and politically valuable.

Unlike Sartre's account, the de Beauvoirian alternative locates the political dimension of an artwork not in its content but in its form, and specifically in the relation it sets up between artist and audience. This is echoed by Marcuse's claim that truly politically valuable art won't be art that has overtly political content or that is used as a means of making statements or pushing an agenda,<sup>38</sup> but will be art that involves what he calls the 'restructuring' of consciousness and the transcendence of the status quo (*AD*: 44-53, 72-73). There are other points of similarity between the position that Marcuse advocates and the account I have argued for here that are unfortunately beyond the scope of this essay to examine. Even more unfortunately, it would extend the length of this essay farther than I already have for me to consider examples of what I see as a tendency in current discourses about art and fiction, both academic and popular, to assume something like Sartre's account and to consider artworks as good when they present the 'right' political perspectives, and as flawed *qua* art when their content is deemed to be politically problematic. Since I don't have space for this, I want to end by noting why, if there is such a tendency, the above account of art's value and of how art can be authentically politically engaged is important.

To the extent that literature and other arts are often presented in educational contexts in a way that teaches students to expect—or even demand—that artworks, especially fictional narratives, will endorse political stances and illustrate theories, they are being

---

38 See *AD*: xii-xiii—"The more immediately political the work of art, the more it reduces the power of estrangement [i.e. the ability to disclose things in genuinely *new* ways] and the radical, transcendent goals of change."

kept from learning to engage with art in the ways that realize the kinds of values that de Beauvoir insists art can have, and, as I have argued, from the way in which artworks can truly have political value. Moreover, this way of presenting art may be promoting bad faith to the extent that readers and viewers are taught to understand fictional characters and their actions in ways that reduce them to elements of their facticity and treat them as representing a social type. As Sartre says, it is a bad work of art that “aims to please by flattering” (*WL*: 45), where one of the ways a work can flatter is by endorsing a political position that is in line with and seems to confirm its audience’s existing beliefs, values, and commitments. Not only does this risk encouraging the epistemic vice of confirmation bias, but by affirming what a reader or viewer is assumed to already think and by pandering to the ways its audience members are assumed to already feel, such an approach denies them the chance to go beyond these fixed perspectives and develop further in their thinking, feeling, and understanding, where this won’t necessarily result in an overturning or rejection of their political position but can add to and thereby deepen their understanding of it, making their commitment to it more authentic.

As de Beauvoir was aware, works that flatter their audience by giving them what they already know and like are easier to consume and so are likely to be more popular than works that challenge audiences by presenting them with something genuinely new from which they can develop not only as thinkers and feelers, but as moral and political agents. Near the end of “Literature and Metaphysics” she writes that “the reader quite often refuses to participate sincerely in the experiment into which the author tries to lead him; he does not read as he demands that one write; he is afraid to take risks, to venture. [...] But the reader must not try to elude this uncertainty and his share of the adventure. He should not forget that his collaboration is necessary, since the novel’s distinctive feature is, precisely, to appeal to his freedom” (*LM*: 276). In other words, the value that literature and other arts can have, including their social or political value, depends not only on how they are created but also on how they are engaged with by their recipients: readers,

---

viewers, listeners, etc. If the account I have outlined and argued for is right, education in literature and other arts should teach people how to engage with works that challenge them and that present new and genuinely other perspectives, and moreover should work against their desires to be flattered and their expectations that artworks will cater to what they already think, know, and agree with, or their expectations that artworks will tell them what to think and feel about what they present them with rather than leaving them with experiential ambiguities to work through for themselves. This is especially true for educators who are concerned with increasing their students' political awareness and promoting positive social relations, since teaching people to want to be flattered and to want others to hand them an interpretation of their own experiences can itself alienate them from their inherent freedom *qua* conscious beings-for-themselves, and so will help to make them exactly the sort of political subjects that such educators, if they are truly 'committed', should want most to avoid.<sup>39</sup>

---

<sup>39</sup> Thanks to Michal Brienza, Sophie Fernier, and Jordan Shaw for their helpful comments on the penultimate draft of this paper, and to Iris Vidmar Jovanović for her support, and for allowing the paper to be as long as it needed to be.

## Bibliography

- Baugh, B. (1988). "Authenticity Revisited." *The Journal of Aesthetics and Art Criticism* 46(4): 477-87.
- Bergson, H. (1889). *Time and Free Will: An Essay on the Immediate Data of Consciousness*, trans. F. L. Pogson. Mineola, NY: Dover, 2001.
- Carney, R. (2000). "Two Forms of Cinematic Modernism: Notes toward a Pragmatist Aesthetic." In T. Luddington, ed., *A Modern Mosaic: Art and Modernism in the United States*, 357-415. Chapel Hill, NC: University of North Carolina Press.
- Caute, D. (1967). "Introduction." In Sartre 1948, vii-xvii. London: Methuen & Co Ltd.
- Collingwood, R. G. (1931). "Aesthetic Theory and Artistic Practice." In D. Boucher, W. James, & P. Smallwood, eds., *The Philosophy of Enchantment*, 81-112. Oxford: Oxford University Press, 2005.
- Collingwood, R. G. (1938). *The Principles of Art*. Oxford: Clarendon Press.
- de Beauvoir, S. (1945). "Existentialism and Popular Wisdom," trans. M. Timmermann. In M. Simons, ed., *Simone de Beauvoir: Philosophical Writings*, 203-20. Chicago: University of Illinois Press, 2004.
- de Beauvoir, S. (1946). "Literature and Metaphysics," trans. V. Zaytzeff & F. M. Morrison. In M. Simons, ed., *Simone de Beauvoir: Philosophical Writings*, 269-77. Chicago: University of Illinois Press, 2004.
- de Beauvoir, S. (1948). *The Ethics of Ambiguity*, trans. B. Frechtman. New York: Citadel Press, 1976.
- de Beauvoir, S. (1965). Contribution to "What Can Literature Do?" In Y. Buin, ed., *Que peut la littérature?*, 72-92. Paris: L'herne.
- Emerson, R. W. (1844). "The Poet." In B. Atkinson, ed., *The Essential Writings of Ralph Waldo Emerson*, 287-306. New York: The Modern Library, 2000.
- Grene, M. (1952). "Authenticity: An Existential Virtue." *Ethics* 62: 266-74.
- Goldthorpe, R. (1992). "Understanding the Committed Writer." In C. Howells, ed., *The Cambridge Companion to Sartre*, 140-177. Cambridge: Cambridge University Press.
- Guignon, C. and Pereboom, D., eds. (2001). *Existentialism: Basic Writings*. Indianapolis: Hackett.
- Howells, C. (1978). "Sartre and the Commitment of Pure Art." *British Journal of Aesthetics* 18(2): 172-82.
- Hume, D. (1740). *A Treatise on Human Nature*, D. F. Norton & M. J. Norton, eds. Oxford: Oxford University Press, 2000.
- Kant, I. (1793). *Critique of the Power of Judgment*, trans. P. Guyer & E. Matthews. Cambridge: Cambridge University Press, 2000.
- Marcuse, H. (1978). *The Aesthetic Dimension: Toward a Critique of Marxist Aesthetics*, trans. H. Marcuse & E. Sherover. Boston: Beacon Press.
- Marx, K. and Engels, F. (1973). *Marx and Engels on Literature and Art: a Selection of Writings*, L. Baxandall & S. Morawski, eds. St. Louis, MO: Telos Press.

- Matravers, D. (2014). *Fiction and Narrative*. Oxford: Oxford University Press.
- Moi, T. (2009). "What Can Literature Do? Simone de Beauvoir as a Literary Theorist." *PLMA* 124(1): 189-98.
- Murdoch, I. (1953). *Sartre: Romantic Rationalist*. London and Glasgow: Fontana Library.
- Orwell, G. (1948). "Writers and Leviathan." In *The Penguin Essays of George Orwell*, 459-65. London: Penguin.
- Rorty, R. (1989). "Solidarity." In *Contingency, Irony, and Solidarity*, 189-98. Cambridge: Cambridge University Press.
- Sander, H.-D. (1970). *Marxistische Ideologie und allgemeine Kunsttheorie*. Tübingen: Mohr.
- Sapiro, G. (2006). "Responsibility and Freedom: the Foundations of Sartre's Concept of Intellectual Engagement," trans. S. Emanuel & G. Townsend. *Journal of Romance Studies* 6(1-2): 31-48.
- Sartre, J.-P. (1943). *Being and Nothingness*, trans. H. Barnes. New York: Washington Square Press, 1984.
- Sartre, J.-P. (1946). *Existentialism is a Humanism*, trans. C. Macomber. New Haven and London: Yale University Press, 2007.
- Sartre, J.-P. (1948). *What is Literature?*, trans. B. Frechtman. London: Methuen & Co Ltd., 1978.
- Sartre, J.-P. (1957). *Search for a Method*, trans. H. Barnes. New York: Random House, 1968.
- Sartre, J.-P. (1960). *Critique of Dialectical Reason*, vol. 1, trans. A. Sheridan-Smith. Atlantic Highlands, NJ: Humanities Press, 1976.
- Sartre, J.-P. (1964). *The Words: The Autobiography of Jean-Paul Sartre*, trans. B. Frechtman. New York: George Braziller.
- Sartre, J.-P. (1971-72). *The Family Idiot*, 5 vols., trans. C. Cosman. Chicago: University of Chicago Press, 1981-1993.
- Sirridge, M. (2003). "Philosophy in Beauvoir's Fiction." In C. Card, ed., *The Cambridge Companion to Simone de Beauvoir*, 129-48. Cambridge: Cambridge University Press.
- Sontag, S. (1966). "Against Interpretation." In *Against Interpretation and Other Essays*, 3-14. New York: Picador.
- Stolnitz, J. (1992). "On the Cognitive Triviality of Art." *British Journal of Aesthetics* 32(3): 191-200.
- Vercors, (1941). *The Silence of the Sea*, trans. C. Connolly. London: Bloomsbury Academic, 2002.
- Walton, K. (1990). *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press.
- Whiting, C. G. (1948). "The Case for 'Engaged' Literature." *Yale French Studies* 1, Special Issue on Existentialism: 84-89.



# Introducing Cinematic Humanism: A Solution to the Problem of Cinematic Cognitivism

*Britt Harrison*

---

## 1. Introduction<sup>1</sup>

Cinematic Humanism is both an example of philosophy of film *without theory* and a commitment to a particular set of tenets about film. These tenets include, but are not limited to, the following:

- i) Some fiction films illuminate the human condition and thereby enrich our understanding of ourselves, each other and our world;
- ii) Such understanding requires our sensitive, reflective, and critical engagement;
- iii) Such sensitive, reflective, and critical engagement requires appreciating the relations between a film's aesthetic and non-aesthetic features;
- iv) Fiction films are a medium that can be used in and for philosophical investigation.

With such tenets, Cinematic Humanism looks to characterise a fundamentally cognitivist approach to the content and value of film, where cognitivism about film is the view that film can be a source of knowledge. As such, Cinematic Humanism might reasonably be called 'Cinematic Cognitivism'. Furthermore, given the third tenet, which points to an important relation between a film's cognitive value and its cinematic value, Cinematic Humanism appears to offer the kind of full-blooded cognitivism found in

---

<sup>1</sup> This paper was originally published in the *Croatian Journal of Philosophy* (XIX, 56, 2019).



the works of, say, Matthew Kieran (2004) on art, or James O. Young (2001) on literature.

Cinematic Humanism is not, however, an example of Cinematic Cognitivism, rather it is an alternative – indeed, a *corrective* – to it. The need for a corrective is motivated by scepticism about the very notion of the cognitive. For the terms ‘cognition’ and ‘cognitive’ are, in fact, theory-laden terms of art, and questionable ones at that. Appreciating this immediately generates two specific problems in the philosophy of film. In the first instance, if justified, scepticism about matters cognitive generates potential worries for the leading methodology of anglophone analytic philosophy of film: *cognitive film theorising*. Secondly, it raises questions about the fundamental assumptions that shape and direct debates about *cognitivism* in film (and beyond).

In this paper, I explore two scepticism-provoking ambiguities relating to the notion of the cognitive and diagnose their source in a pair of stipulative definitions made by Noam Chomsky. These are, I reveal, responsible for changing the meaning of the word ‘cognitive’ into a questionable piece of philosophical jargon. Having identified and articulated these concerns, I introduce Cinematic Humanism as an alternative to Cinematic Cognitivism. I also propose that the methodology of Cinematic Humanism – which I call an example of philosophy of film ‘*without theory*’ – offers a viable way to resist the problems attendant on much of Cognitive Film Theorising, without being driven (back) into the arms of its methodological rival, Film Theory.

## 2. The Cognitive Compromised

Contemporary philosophy currently brims over with things *cognitive*: cognitive processes, cognitive abilities, cognitive mechanisms, cognitive agents, cognitive responsibility, cognitive virtues, cognitive gains, cognitive bloat, cognitive ooze, cognitive bleed, cognitive angst, cognitive dissonance, cognitive sandwiches and

so on.<sup>2</sup> But just what is it to characterize something as cognitive? At first blush it looks like ‘cognitive’ is an adjective used to mean *of or pertaining to* knowledge, as ‘hedonic’ means *of or pertaining to* pleasure. Things are not, however, quite so simple. For there are two key ambiguities at play in the contemporary philosophical use of the notion of the cognitive: a Scope Ambiguity and a Level Ambiguity. With the Scope Ambiguity there are inconsistencies as to what kind of knowledge is supposedly cognitive; with the Level Ambiguity there are obfuscations and equivocations as to whether or not the notions of cognition and the cognitive pick out *person-level* features, properties, or activities, or *sub-personal* ones. Moreover, such Level Ambiguities further compound the various ambiguities of scope. Before diagnosing the source of these difficulties, I take a look at each, in turn.

## 2.1. The Scope Ambiguity

The philosophical scope of the cognitive is, it would seem, as narrow or generous as the scope of knowledge itself. If one has a narrow philosophical conception of knowledge – say one limited to non-Gettierized justified true belief – this engenders a comparably narrow use of ‘cognitive’. On such a view only that which is, or relates to, propositional knowledge can be correctly characterized as cognitive. According to Jukka Mikkonen it is just such a narrow scope of the cognitive that is the default position in Literary Cognitivism. “The traditional cognitive line of thought maintains that literature conveys propositional knowledge.” (Mikkonen, 2013, p. 9)

Yet Cognitive Pluralists, such as Dorothy Walsh (1969), Catherine Wilson (1983), Eileen John (1998), Gordon Graham (2005), and Iris Vidmar (2013), have a broader, more diverse appreciation of what counts as knowledge. On their, and others’, views knowledge is by no means limited to the merely propositional. Rather

---

2 At the 7<sup>th</sup> Dubrovnik Philosophy of Art Conference (2018) James O. Young gave us cognitive *toxicity*, Dustin Stokes championed cognitive *penetration* and there was repeated reference to cognitive *gaps*.

knowledge is also one or more of knowledge-what (something's like), non-propositional know-how, acquaintance knowledge, conceptual knowledge, understanding and, indeed, almost anything that is thought- or ability-enriching. The very elasticity of the potential scope of the cognitive makes it possible for some, more liberal, Literary Cognitivists to champion literature for its capacity to do any or all of the following:

educate emotionally, train one's ethical understanding, call into question moral views, cultivate or stimulate imaginative skills and/or cognitive skills, 'enhance' or 'enrich' the reader's knowledge, 'deepen' or 'clarify' her understanding of things she already knows, 'fulfil' her knowledge or help her 'acknowledge' things, give significance to things, provide her knowledge of what it is like to be in a certain situation, that is, offer her a 'virtual experience', often of situations she could not or would not like to encounter in her real life, and so on... (Mikkonen, 2013, p. 9-10)

Sympatico to such a view is Peter Lamarque:

Who would deny that art is often involved with "exploring aspects of experience," "providing visual images," "broadening horizons," "imagining possibilities," "exploring and elaborating human ideas"? If this is cognitivism, then I too am a cognitivist. (Lamarque, 2006, p. 128-129)

Yet this cognitive largesse is short-lived as Lamarque maintains his debate-shaping anti-cognitivist position by continuing, "But I don't think this has anything essentially to do with truth or knowledge or learning" (Lamarque, 2006, p. 128-129). In so doing, he shuts the door on any hoped-for pluralism: the scope of the cognitive shrinks back once again to its default propositional borders.

If one looks to contemporary epistemologists for clarity on the topic, their philosophical focus on knowledge is almost exclusively on propositional knowledge. As a result, it is practically impossible to ascertain whether or not non-propositional knowledge is or may be deemed cognitive. Recent forays into the area of know-how by Jason Stanley & Timothy Williamson (2001) and Stanley alone (2011) argue resoundingly that knowledge-how is but a particular mode of presentation of what is fundamentally propositional knowledge. This so-called 'intellectualist' view of know-how is increasingly dominant, obscuring the extent to which non-prop-

ositional know-how might *also* be, characterised as cognitive. This difficulty continues in the work of leading virtue epistemologists, such as John Greco and Ernest Sosa, who characterise a virtuous knower as one whose *propositional* knowledge and belief-forming mechanisms are reliable. In Duncan Pritchard & Sven Bernecker's 2011 *Routledge Companion to Epistemology*, there are 900 pages containing sixty so-called 'state of the art' articles, every one of which is dedicated to the consideration of propositional knowledge. If, as is claimed, this book displays contemporary epistemology at its most comprehensive then there is no questioning, let alone avoiding, the hegemony of what, elsewhere (2013, p. 140ff.) I call "the propositional presumption" of epistemology. Unsurprisingly, in practice the notions of the cognitive and the propositional are regularly used interchangeably.

This need not, of course, prevent a philosopher of art who wishes to characterise both propositional and non-propositional knowledge as 'cognitive' from doing just that, and indeed a number of leading analytic aestheticians do so. Support though, for any such 'cognitive pluralism' is not to be found in contemporary epistemology. Indeed, for pluralists about knowledge who work in the philosophy of art it now looks like epistemology is not so much a possible resource for pluralist perspectives, but rather a philosophical area in potential need of them. The valuable direction of travel is perhaps *from* the philosophy of art *to* epistemology, and not vice versa. Were this Scope Ambiguity to be the only ambiguity at play with the cognitive, then I, for one, would willingly take up the cognitivist cause in the hopes of bringing to bear insights offered by so-called 'cognitive pluralists about art' on epistemology. Unfortunately, the second ambiguity – the Level Ambiguity – makes this tempting option not just problematic, but intractably so.

## 2.2. The Level Ambiguity

Level ambiguities about the relation between knowledge and the cognitive turn on confusions as to whether or not knowing and cognizing both occur at the personal level or one occurs at the per-

sonal level and the other occurs at the sub-personal level. As the demarcation between epistemologists and philosophers of mind blurs – as a consequence of the naturalizing ambitions of contemporary analytic philosophy – many philosophers in both areas work with a notion of cognition that is less a way of *characterising* our knowledge, and instead something that, supposedly, *explains* it. Instead of knowledge and cognition both being potential philosophical explananda, cognition is offered as an explanans for the explanandum that is knowledge. Moreover cognition, qua explanans, is conceived of as wholly sub-personal: cognitive sub-personal processes, mechanisms and states are theoretical constituents of a particular view of what the mind is, and how it works. One of the key commitments of this view is that to be minded is to engage in sub-personal information-processing over representational states. In other words, however (potentially) pluralist you might be, *au fond* such niceties disappear as knowledge bottoms out in sub-personal propositional knowledge. Appreciating this shift, helps to explain Stanley's insistence on the propositionality of all knowledge, including know-how, thereby showing that the level ambiguity and the scope ambiguities are internally connected. If cognition is now a sub-personal theoretical posit designed to explain person-level knowledge, then it is not, and cannot be, synonymous with knowledge. When and where did all this happen?

### 3. Just Say 'Yes' to the History of Philosophy

There are four people whose historical confluence is crucial to turning 'cognitive' into, at best, a theory-laden term of art, and, at worst, a misdirecting piece of jargon. The four are Alan Turing, Warren McCulloch & Walter Pitts, and Noam Chomsky. Their work in, respectively, computing, neuroscience & A.I., and linguistics, is crucial to the creation, and location, of the perfect storm that changed the meaning of 'cognitive' and in so doing put the cognitive into cognitive science.

### 3.1. Going Cognitive

The early clouds of this perfect storm gather with the analogy Turing draws between humans and machines, “We may compare a man in the process of computing a real number to a machine which is only capable of a finite number of conditions” (Turing, 1937, p. 231). In other words, in considering ourselves as thinkers, as computers, we can think of ourselves as computing machines. In Turing’s wake comes neurophysiologist and soon-to-be Head of MIT Cybernetics, Warren McCulloch who, together with colleague Walter Pitts, runs with Turing’s suggestion in the provocatively titled paper ‘A Logical Calculus of the Ideas Immanent in Nervous Activity’. Here McCulloch and Pitts argue for an in-principle marriage between the firing of neurons and propositional representation.

The “all-or none” law of nervous activity is sufficient to insure that the activity of any neuron may be represented as a proposition. Physiological relations existing among nervous activities correspond, of course, to relations among the propositions... (McCulloch and Pitts, 1943, p. 117)

McCulloch & Pitts’ paper ends with a powerful vision of the potential of their proposal. “Thus both the formal and the final aspects of that activity which we are wont to call *mental* are rigorously deducible from present neurophysiology...” (McCulloch and Pitts, 1943, p. 132). That is to say, personal-level thoughts are (according to this theoretical proposal) *inferable* from sub-personal propositionally construed neuronal firings. Confirming this radical suggestion and thence exploiting such a claim is cognitive science’s *raison d’être*. It is the Holy Grail cognitive science has been chasing ever since its inception as a discipline born of a view of the mind as a localizable intercranial proposition-encapsulating neuron-firing computer. Indeed, by 1950 Turing is confident that computers can be made to “mimic the actions of a human computer very closely” (Turing, 1950, p. 438). He suggests one way to bring this about:

Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child’s? ... Our hope is that there is so little mechanism in the child-brain that something like it can be easily programmed... We have thus divided our

problem into two parts. The child-programme and the education process. (Turing, 1950, p. 456)

The temptations of such a research project are clear: In the “child-machine... one might have a complete system of logical inference ‘built in’” (Turing, 1950, p. 457) And there’s a footnote here: “Or rather ‘programmed in’ (Turing, 1950, p. 457, fn1.) This, then, is Chomsky’s cue, his springboard. For throughout the 1950s Chomsky synthesizes the ideas of Turing and McCulloch & Watts to develop his own claims that what it is to know how to speak a language *just is* to have such an innate sub-personal propositional-based language-constituting programme or mechanism. By the time he unleashes his castigating review of Skinner’s ‘Verbal Behaviorism’ in 1957, Chomsky is not simply engaging in methodological criticism he is simultaneously unveiling a brand new approach, and set of theoretical presumptions, applicable not just to language, but to all of our intelligent and intentional behavior:

One would naturally expect that the prediction of the behavior of a complex *organism (or machine)* would require in addition to information about external stimulation, *knowledge of the internal structures of the organism*, the way in which it *processes input information* and organizes its own behaviour. (Chomsky, 1957, p. 27, emphases added)

Chomsky presents his Universal Grammar as the first of these innate information-processing internal structures, proposing that we are born with a so-called ‘Universal Grammar’, whose individual ‘initial state’ incorporates a postulated fundamental structure of *all* languages. This language faculty or organ then grows into its mature ‘steady state’. Both the initial and the mature steady states are mental states represented in the mind/brain that are constitutive of the information-bearing, propositional representations and rules that we process, or compute. All this happens at the sub-personal level, “far beyond the level of actual or even potential consciousness” (Chomsky, 1965, p. 8).

So we arrive at the critical move that spawns, and still shapes, today’s ambiguity-ridden notion of the cognitive, namely: Chomsky’s stipulative theoretical definition:

I have been speaking of “knowing English” as a mental state (or a

stable component of mental states), or a property of a person in a certain mental state, but... What is it that is known? Ordinary usage would say: a language – and I have so far been keeping to this usage, speaking of knowledge and learning a language, eg. English. But... this way of talking can be misleading... To avoid terminological confusion, *let me introduce a technical term devised for the purpose, namely “cognize” with the following properties...* The particular things we know, we also “cognize”... Furthermore, *we cognize the system of mentally-represented rules from which the facts follow. That is we cognize the grammar that constitutes the current state of our language faculty and the rules of this system as well as the principles that govern their operation. And finally, we cognize the innate schematism, along with its rules, principles and conditions.*

In fact, I don't think that “cognize” is very far from “know”... *If the person who cognized the grammar and its rules could miraculously become conscious of them, we would not hesitate to say that he knows the grammar and its rules, and this conscious knowledge is what constitutes his knowledge of language. Thus cognizing is tacit or implicit knowledge, a concept that seem to me unobjectionable... cognizing has the structure and character of knowledge... but may be and is in the interesting cases inaccessible to consciousness. I will return to the terms “know” and “knowledge”, but now using them in the sense of “cognize”... The fundamental cognitive relation is knowing a grammar.* (Chomsky, 1980, p. 69-70, emphases added)

With this strategic announcement Chomsky separates knowledge and cognizing, making the latter a theoretical notion that is a constitutive part of a (naturalised) theory about what it is to know, or to know how to speak, one's first language. Moreover, by announcing his intention to return to using the terms ‘know’ and ‘knowledge’ *in ways that now mean* (or are synonymous with) this theory-laden notion of cognize, Chomsky and his heirs in the philosophy of mind, linguistics and cognitive science do not just equip themselves with their key theoretical posit, they commit to a practice that cannot but generate and embed the kinds of level and scope ambiguities that are constitutive of today's philosophical and cognitive science ‘research’. By the time Chomsky's gives the 1969 John Locke Lectures at Oxford, Universal Grammar's central notion of cognition as unconscious, sub-personal propositional tacit knowing, is now the model on which most, if not all, scien-



tific and naturalized philosophical attempts to understand not just language, but human intelligence and mindedness *tout court*. Chomsky successfully baits his hook with the familiar (person-level) concept *knowing*, then switches its meaning to a new (sub-personal-level) theoretical concept *cognising*, before reverting to the original nomenclature of knowledge to exploit person-level intuitions and conceptual connections relating to our more familiar notions of knowledge, language and mindedness.

One might think, however, that the concept *know-how* would be excluded from, or immune to, such deliberate theoretical repurposing. One might think it reasonable to characterise what it is we know, when we know how to speak our first language, as a kind of non-propositional know-how, an ability, and thus it is in some way untouched by theoretical proposals that reconceive person-level *propositional* knowledge as sub-personal propositional cognising. But non-propositional know-how offers no escape from Chomsky's 'bait-and-switch' manoeuvre. For it turns out that there is no such thing as the non-propositional know-how of language.

### 3.2. Reconceiving Competence

Having turned accessible personal-level knowing into inaccessible sub-personal cognising, Chomsky makes a second, related stipulation that does not simply consolidate, it exacerbates, the dual-level ambiguity inherent in the notion of the cognitive. He first separates the notions of *competence* and *performance*. "We thus make a fundamental distinction between *competence* (the speaker-hearer's knowledge of his language and *performance* (the actual use of language in concrete situations)" (Chomsky, 1965, p. 4, emphasis added). Chomsky then drives a theoretical wedge between such competence and performance, announcing: "...one might have the cognitive structure that we call "knowledge of English" fully developed, *with no capacity to use this structure*" (Chomsky, 1975, p. 23, emphasis added). That is to say, that what it is to know English no longer means, entails or is constitutive of being able to speak English and understand other English-speakers. Instead,

Chomsky proposes, or better he theoretically stipulates, that:

...it is possible in principle for a person to have a full grammatical competence and no pragmatic competence, *hence no ability to use a language appropriately, though its syntax and semantics are intact.* (Chomsky, 1980, p. 59, emphasis added)

With these stipulations Chomsky confirms his philosophico-theoretical claim that one can be linguistically competent in English, in other words you can *be* in a sub-personal cognitive state, yet unable to actually speak a language. To know-how to speak and understand English is no longer one and the same as having the ability to speak and understand English. Just as theory-laden *cognising* usurps (propositional) knowledge, *competence* usurps know-how. Only grammatical not pragmatic competence (a newly minted theoretical distinction) is required to know (or know-how) to speak or to understand a language. Moreover, grammatical competence is, unsurprising, sub-personal, propositional and – by Chomsky’s own lights, cognitive. Sub-personal cognition now supposedly *explains* personal level knowledge, understanding and ability. Yet, at the same time, the use of these notions and terms trades on our non-theoretical associations and assumptions about knowledge, understanding and ability.

#### 4. The Double Irony of So-called ‘Cognitive Competence’

Unperturbed by the implausibility (and dubious coherence) of this, Chomsky offers a further justification for the value of his newly minted, theory-laden terms:

...my concept ‘knowledge of a language’ is directly related to the concept ‘internalization’ of the rules of grammar’... [ and I have] tried to avoid, or perhaps evade the problem of explication of the notion ‘knowledge of language by using an *invented technical term, namely the term ‘competence’* in place of ‘knowledge’. However, the term ‘competence’ suggests ‘ability’, ‘skill’ and so on, through a chain of associations that leads directly to much new confusion. *I do not think the concepts of ordinary language sufficient for the purpose at hand; they must either be sharpened, perhaps somewhat arbitrarily, or replaced by a new technical terminology.* (Chomsky, 1975, p. 315, emphasis added)

With ‘competence’ joining ‘cognising’ as the twin pillars of Chomsky’s new technical terminology, matters are poised for a third theoretical posit: ‘cognitive competence’. Cognitive competence supposedly picks out sub-personal propositional knowledge whilst making no commitments to any person-level propositional knowledge, know-how or abilities. Not only that, but this product of Chomsky’s double bait-and-switch is now tied to the denigration of our standard vocabulary, newly reconceived as ‘folk psychological talk’ and, as such, inadequate. ‘Cognitive’, ‘competence’ and ‘cognitive competence’ become key theoretical terms: tools of choice for naturalizing philosophers eager to ‘improve’ upon our ordinary language which has now been shown, supposedly, to be incapable of rising to the latest philosophical demands. But if any contemporary use of the term ‘cognitive’ and ‘competence’ cannot but consolidate theory-laden views where does this leave philosophers of art, or film? And what of cognitive film theorists? Are they unaware of the metaphysics of mind and language that are constitutive (thanks to Chomsky) of these notions or do they deliberately embrace it? And for those philosophers of art and film who might be cautious of making such commitments in the metaphysics of mind – what to do?

## 5. A Trilemma

Do philosophers of art use the term ‘cognitive’ with all the ambiguities and attendant sub-personal commitments exploited by philosophers of mind or metaphysics-first epistemologists? If not, must they? Can a notion of the cognitive that is not theory-laden in the way outlined in the previous section be identified and/or maintained? Do the silos of specialism in philosophical academia perpetuate nomenclature confusions or offer ways to transcend such worries, and if so, how? To what extent are, or might, these *intradisciplinary* conundrums be ramified by *interdisciplinary* engagement? Philosophers of art, including film, are, I suggest, facing a trilemma as to how best to respond to, and engage with, these theory-laden notions. Should the terms ‘cognising’, ‘cognitive’, and

‘competence’ (i) be embraced; (ii) be used but in only tandem with caveats and clarifications that modify and/or mollify concerns relating to scope or level ambiguities; or (iii) be eschewed altogether?

The first option – to continue unruffled, undaunted – can be seen in the standard practices of the majority of contemporary anglophone analytic philosophers whose work involves or overlaps with the philosophy of mind and naturalized epistemology. It is also the preferred approach of so-called ‘cognitive film theorists.’ For cognitive film theorists these theory-laden notions are key to their methodological *modus operandi*. Perhaps they have found a way to diffuse the scepticism that I propose compromises the very notion of the cognitive. To evaluate the merits of this diffusion, I first consider why cognitivism has been, and continues to be, so important to the creation and maintenance of cognitive film theorising.

### 5.1. 1996 and All That

In 1996, Noel Carroll and David Bordwell’s edited collection of articles, *Post-Theory: Reconstructing Film Studies* brought together a range of critical challenges directed at the then dominant methodology of film studies – Theory. The editors’ own contributions to the volume led the attack: the claims of Theory were not simply false (where coherent), but the Theoretical methodology was, itself, inadequate. Carroll invited the purveyors of Theory to justify their approach and rise to the scholarly responsibility of engaging in dialectic debate about their *modus operandi* and its products. The invitation has remained unanswered; the gauntlet unrun.

*Post-Theory: Reconstructing Film Studies* opens with the editors’ individual articles articulating and cataloguing the limitations of Theory, or as they sometimes call it ‘Grand Theory’. At the same time, both David Bordwell and Noel Carroll champion their insistence on high standards of clarity, rigour, and rationality to which cognitive film theorising is to be accountable. Bordwell contrasts the cognitivists’ own, “middle-level research programmes... based in evidence” (Bordwell and Carroll, 1996, p. 29) with the

‘ethereal speculations’ (Bordwell and Carroll, 1996, xiii) and “sedimented dogma” (Bordwell and Carroll, 1996, xvii) of Theory. He characterises the various manifestations of Grand Theory – be they Marxist, psychoanalytic, semiotic, structuralist, poststructuralist, postmodern, or feminist – as resulting from an “esoteric merger of antirationalist philosophy, unorthodox psychoanalysis and the frequently changing views of an official philosopher of the French Communist Party” (Bordwell and Carroll, 1996, p. 14). The purveyors of Theory traffic in ideas that meet “no canons of reasonable inference” (Bordwell and Carroll, 1996, p. 23) and their theories are little more than “a bricolage of other theories” (Bordwell and Carroll, 1996, p. 25). Bordwell’s historical reconstruction of Theory’s highways and byways, from subject-position theory through to cultural studies, charts the “deep continuities of doctrine and practice” (Bordwell and Carroll, 1996, p. 13) that began in the 1970s and continue unchallenged up to this Post-Theory confrontation. As well as cataloguing the failures, follies and inadequacies of the results of Theory, Carroll identifies methodological “impediments to film theorizing” (Carroll, 1996, p. 38). These range from the misconceived overextension of psychoanalytic theory (overextended because the standard clinical use of psychoanalysis is limited to explaining just those deviations that are recalcitrant to ‘normal’ understanding) to engaging in *ad hominem* attacks on any critic who refuses to acknowledge the supposedly ever-present politico-ideological dimension of a film; from using a notion of interpretation in such a way as to transform distinct films into the homogeneous products of a “standard-issue sausage machine” churning out (readings of) films that look and smell the same (Carroll, 1996, p. 43), to inventing concepts of questionable use, such as “the male gaze” (Carroll, 1996, p. 45); and from incorrectly insisting that content-free formalism is the inevitable consequence of any attempt at political or ideological neutrality, to offering “arguments for suspecting science [that] are as feckless as those for suspecting truth” (Carroll, 1996, p. 59).

Carroll announces his hopes of engendering a “methodologically robust pluralism” (Carroll, 1996, p. 63); one that would en-

courage and enable cognitivists and Theoreticians to engage with each other, sharing agreed standards and protocols of reasoning; together facing the tribunal of empirical evidence. Such academic engagement fails to come to pass. Critical challenge as a route to pan-theoretical corrective is not, and was not, to be. Instead, Carroll's vision of robust pluralism gave way to the very thing he had hoped to avoid: "coexistence pluralism" (Carroll, 1996, p. 63). The result was – and indeed continues to be – not so much a live-and-let-live mutual respect, but a live-and-rarely-if-ever-mention disparagement.

The lack of any serious reaction from Theoreticians was perhaps unsurprising given Bordwell and Carroll's choice of language was not designed to cushion their critical onslaught. Calls to scholarly engagement are problematic when paired with declarations that the leading Theoretical emperors are not wearing any clothes. The dust jacket illustration of *Post-Theory* displayed a photograph of Laurel & Hardy 'teaching': surely little more than a pointed accusation of the clown-like hopelessness of Theory, and a motivating invitation for real, rather than buffoon, teacher-scholars to step up to the academic plate.

In extolling the virtues of cognitive film theorising Carroll announced that the new methodology would deliver rigorous argument and clarity where Theory was awash with impenetrable, obscure prose. It would offer the authority of legitimate scientific investigation, where Theory just stumbled around, committing every sin in the Analytic Philosophers' Handbook. Carroll didn't hesitate to name and shame those whom he took to be the key culprits of Theory: Louis Althusser, Jacques Lacan, and Roland Barthes. Nor did Carroll's condemnation stop there: he attacked profit-hungry over-productive university presses that pandered to the 'arcane peregrinations of Theory' by publishing anyone who had the audacity to draw not just from the well of their founding fathers, but from the writings of Michel Foucault, Julia Kristeva, Pierre Bourdieu, Gilles Deleuze, 'maybe sometimes' Jacques Derrida, and the list - like the 'juggernaut of Theory' - went on. (Carroll, 1996, p. 37-40).

Since then, the division between these two camps has deepened: cognitivist film theorising blossoms in the soil of the naturalized analytic philosophy which is now the default paradigm of the contemporary analytic philosophical academia. Theory carries on unabashed and unabated, for the most part ignoring challenges to its ideological cornerstones, seemingly unperturbed by the fact that its prose style is incomprehensible to the uninitiated. The actual ongoing philosophical battle, as Carroll anticipated in 1996, is not, however, *between* these mutually exclusive methodologies, but *for* the undecided readership who have yet to make up their philosophical mind and/or who are still to be inculcated into the practices and norms of one or other of these camps.

Yet although Bordwell and Carroll target the trio of Althusser, Lacan and Barthes as the miscreant source of Theory's problems, they too have their own equivalent Triumvirate in Chomsky, Fodor, and Quine whose philosophical commitments - of method, substance and nomenclature - they embrace. For in rejecting Theory (with a capital 'T') as unscientific gibberish, cognitive film theorists turn to the representational and computational theories of mind that are constructed out of sub-personal semantic theories of content, 'cognition', 'competence' and intentionality. Furthermore, even when their work seems not to require any such commitments to such philosophies of mind, they are now exploiting the conceptual-theoretical resources and vocabulary sourced in, and constitutive of the metaphysical underpinnings of their methodological orientation. In other words, cognitive film theory is no less dependent on its own fundamental theoretical commitments as Grand Theory was, back in 1996. Yet for many, the very idea of sub-personal propositional knowledge, sub-personal notions of cognition and content is at best wrong, and at worst incoherent.

## 5.2. Myths, Broken Dreams & Cul De Sacs

The catalogue of unresolved charges filed against the various presumptions that shape the cognitivist metaphysics of mind includes the Chinese Room Argument against the very idea of

sub-personal semantic content<sup>3</sup>, the category mistake constitutive of attempts to *localise* powers<sup>4</sup>, the Homunuclus and Merelogical Fallacies that mistakenly predicate of brains what can only be predicated of people<sup>5</sup>; the unfathomable challenge of showing how moods, skills and understanding might be sub-personally represented<sup>6</sup>; the impotence of sub-personal ‘competence’ to be, to replace or to explain public standards of correctness; the frame problem; accusations of scientism, etc., the list goes on. Yet cognitive film theorists such as Greg Currie, David Bordwell and Carl Plantinga not only embrace but readily acknowledge the importance of the very same theory-laden notions of cognition and competence laid out above together with the very representational theories of mind they enable and nourish<sup>7</sup>. They are undeterred by those, like Norman Malcolm, who regard the idea of understanding or explaining mindedness and intelligence using so-called ‘cognitive processes’ as nothing but a case of “replacing the stimulus-response mythology with a mythology of inner guidance systems” (Malcolm, 1971, p. 392). They are undaunted by those, like Herman Philipse, who describe cognitivism as yet another misguided attempt to turn philosophy into science, the history of which he characterises as a “boulevard of broken dreams” (Philipse, 2009, p. 163). They are uninterested in the pronouncements of leading cognitive science apostates, such as Rodney Brooks (whose 1970s MIT team built one of the first robots to move around an ‘ordinary’ environment) who now acknowledges that the computer “intelligence” is a primarily a matter of computational brute force rather than anything that involves meaning or is, in any way, comparable to understanding. Brooks recently announced:

---

3 John Searle, 1980.

4 Ryle, 1949; Kenny, 1989; and Kenny, 2009.

5 Kenny, 1989; Bennett & Hacker, 2003, p. 68-108.

6 Haugeland, 1978, p. 22.

7 See Currie’s continued commitment to a Chomksy-informed understanding of matters cognitive. “Our speech-production runs... much slower than the cognitive processes that enable us to think and draw inferences from our thoughts.” (Currie, 2010, p. 15).



I believe that we are in an intellectual cul-de-sac, in which we model brains and computers on each other, and so prevent ourselves from having deep insights that would come with new models... The brain has become a digital computer; yet we are still trying to make our machines intelligent... When you are stuck, you are stuck. We will get out of this cul-de-sac, but it will take some brave and bright souls to break out of our circular confusion of models. (Brooks, 2012, p. 462)

They are entirely undeterred by Rorty's observation that, "[f]rom a Wittgensteinian perspective, the approach taken by Chomsky and his fellow cognitive scientists look like that taken by the man who searches for his missing keys under the lamp-post, not because he dropped them near there but because the light is better" (2004, p. 221).

That said, there has been a move by Greg Currie to step away from the (potentially problematic) nomenclature of the 'cognitive'. Unlike his fellow cognitive film theorists, Greg Currie has declared the label 'cognitivism' to be "of limited usefulness", even "burdensome". He suggests a better name would be "rationalism" (Currie, 2004, p. 170). Though his preferred approach still welcomes "help from the empirical sciences" the crucial idea captured by rationalism is that it maintains a "commitment to reasoned and reasonable ways of thinking" whilst avoiding the requirement of maintaining allegiance to any specific theory of mind (Currie, 2004, p. 170). This seems like a promising suggestion, perhaps one ready to acknowledge if not all the list of above-mentioned challenges, then at least some of the Scope and Level Ambiguities involved in the notion of the cognitive, and attendant assumptions of Cognitivism. Yet let it be remembered that Chomsky regards his Universal Grammar to be a case of what he calls Cartesian Linguistics; a 'Chapter in the History of *Rationalist* Thought' (2009). Currie's suggestion is perhaps more accurately appreciated as an attempt to re-brand Cognitive Film Theory, whilst holding on to its fundamental commitment - ie. the principle that cognition is subpersonal information-processing. In the preface to his 1995 'Image & Mind, Film and Cognitive Science', Currie acknowledges that his book "owes much, in spirit at least, to the linguistics of Chomsky" (Currie, 1995, xxiii).

Nothing has changed; or is likely to.

### 5.3. Back to the Trilemma

What to do, then, if one does not want to use the notion of the cognitive, or any related cognitivist methodology; if one wants to avoid the pitfalls of Scope and Level Ambiguity, and wishes to ‘opt out’ of the problematic cognitive-informed picture of the metaphysics of mind? The second option of the trilemma is to continue to use these notions, but suitably accompanied by the appropriate caveats, clarification and disambiguations. This is, indeed, a viable option. It does, however, come with its own difficulties: how best to engage with colleagues, interlocutors and philosophical adversaries who are neither interested in, versed in, nor see the need for, such clarificatory preliminaries? Is it practically possible to regularly and repeatedly rehearse questions about the meaning and implications of what, to many, are seemingly innocuous terms?

That leaves the third option: eschewing the cognitive. Perhaps it is time to acknowledge the merits of Gilbert Ryle’s prescient advice. The “proper policy” when faced with the question *Is imagining a cognitive or non-cognitive activity?* is to “ignore it. ‘Cognitive’ belongs to the vocabulary of examination papers” (Ryle, 1949, p. 244). But can we do without the term and its associated notions?

## 6. Doing Without and Doing Away with the Cognitive

Resisting the use of questionable theory-laden notions such as cognition, the cognitive and cognitivism, is not easy. These notions pervade almost all of the various philosophical sub-disciplines that make up today’s naturalized analytic philosophy. They are also part of the currency of contemporary cognitive science and so would appear to be prerequisites for any interdisciplinary engagement. Furthermore, just as evidence shows that fMRI imagery is taken, by non-specialists, to be more explanatory powerful<sup>8</sup>, so too, there

<sup>8</sup> See Weisberg et al (2008) on the so-called ‘seductive allure of neuroscience explanations.’

seems to be a rhetorical authority that comes with the terminology of cognition. In the financial marketplace of contemporary academia it is all too easy, even for the sceptic, to embrace the rhetoric power of terms like ‘cognition’ which project a seemingly scientific robustness attractive to those non-specialists who often, and increasingly, hold the purse strings of ‘research’ grants. This may well be a sociological aspect of the slippery slope that goes some way to explaining the appeal of scientism. Nonetheless, individual philosophers of film, of art, and beyond, must still decide whether or not they wish to use the terms ‘cognition’, ‘cognising’, ‘cognitive’ and ‘competence’ and take responsibility for their role in maintaining and contributing to what these terms have come to mean. Cognitive film theorists once recoiled from the ‘arcane peregrinations’ that is the language of ‘Theory’, yet their own cognitivist picture of the mind is no less a product of a highly specialised practice of talking and writing into which its adherents have been inculcated. This is confirmed, unwittingly, by Stephen Stich, the cognitivist philosopher who originally articulated the theory-laden notion of the (supposed) sub-doxastic mental state.

Though talk of [sub-personal] states representing facts is difficult to explicate in a philosophically tolerable way, it is surprisingly easy to master intuitively. *Even the barest introduction to work in artificial intelligence and cognitive simulation quickly leaves one comfortable with attributions of content or representational status to the states of an information processing theory* (Stich, 1978, p. 510, emphasis added).

Scientism comes, I suggest, with its own arcane peregrinations.

The historical reconstruction and arguments above are sufficiently worrisome, I believe, to justify why it is important to remain uncomfortable with what is ultimately a misguided picture of philosophising about knowledge, know-how and understanding, and to resist using the language that engenders it. For as Peter Hacker reminds us, “According to Chomsky, someone who cognizes cannot tell one what he cognizes, cannot display the object of his cognizing, does not recognize what he cognizes when told, never forgets what he cognizes (but never remembers it either) has never

learnt it, and could not teach it. Apart from that, cognizing is just like knowing! Does *this* commend itself as a model for an intelligible extension of a term?" (Hacker et al, 2007, p. 138). I think not.

In resisting the language of the cognitive and its sister notion of competence, one is not merely turning away from scientific jargon, but opening the door to the possibility of rehabilitating the value of our ordinary, rich, person-level vocabulary and concepts: knowledge, know-how, experience, understanding, insight, judgement, explanation, appreciation, wisdom, reflection, consideration, taste, exploration, practice, imagination, etc. These are not the impoverished notions of some primitive folk psychology in urgent need of philosophical overhaul. They are the tools of our human trade and traffic, the raw material of some of our finest art, and the wherewithal with which we live our lives. Just saying no to the use of all things cognitive is not only a *solution* to the trilemma posed but an *opportunity* for the philosophy of film, and art, to find a different way forward in the 21<sup>st</sup> century: an opportunity I characterise as *humanist*.

Cinematic Humanism offers an alternative to the methodology of cognitive film theorising without being forced back into the no less questionable theoretical claims of (Grand) Theory and its heirs<sup>9</sup>. Cinematic Humanism is, instead, an example of a non-cognitive-involving way of doing the philosophy of film *without theory*, as well as a commitment to a set of tenets about the non-trivial value of fiction films. As a methodology it resists employing naturalized theories in the philosophy of mind, avoids the associated theory-laden vocabulary and jargon, and refuses to participate in the downgrading of the philosophical value of our ordinary language. The challenge Cinematic Humanists face is to discern and articulate the similarities, distinctions and reticulations that constitute that understanding of ourselves, each other and the world achieved in and through our sensitive, reflective and critical engagement with films. I would hope that supporters of what might be

---

9 I take 'film-philosophy' to be one such heir: an iteration of Theory triggered (in part) by the cognitive film theorists' original 1996 criticisms.

termed ‘Cinematic Cognitivism’ find much to support in the tenets of Cinematic Humanism, for – representational and computational theories of mind apart – there is a not insubstantial set of shared commitments. I further hope that by encouraging scepticism about the very notion of the cognitive Cinematic Humanist approaches offer ways for debate about ‘cognitive’ value to move beyond current stalemates. Cinematic Humanism is, and will continue to be, a solution to the constitutive problem of Cinematic Cognitivism by reminding us that it is at the personal and interpersonal levels, and not the sub-personal level, where our philosophical understanding of what it is to be human is to be found. It is at the personal and interpersonal level where the meaning, insight and value of our cinematic achievements are to be recognised, appreciated, and cherished.<sup>10</sup>

---

<sup>10</sup> I am particularly grateful to the organisers of the 7<sup>th</sup> Dubrovnik Philosophy of Art Conference for inviting me to present an earlier version of this paper. My thanks also to Peter Lamarque, Daniele Moyal-Sharrock, and audiences and colleagues in York, Hatfield, Tampere, and Dubrovnik for helpful comments and discussion.

## Bibliography

- Bennett, M.R. and Hacker, P.M.S. (2003). *Philosophical Foundations of Neuroscience*. Oxford: Blackwell.
- Bennett, M., Dennett, D., Hacker, P. and Searle, J. (2007). *Neuroscience and Philosophy: Brain, Mind and Language*. New York: Columbia University Press.
- Bernecker, S. and Pritchard, D. (2011). *Routledge Companion to Epistemology*. London: Routledge.
- Bordwell, D. (1989). "A Case for Cognitivism." *Iris* 9(Spring): 11-40.
- Bordwell, D. and Carroll, N. (1996). *Post-Theory: Reconstructing Film Studies*. Madison, Wisconsin: University of Wisconsin Press.
- Brooks, R. A. (2012). "Is the Brain a Good Model for Machine Intelligence." *Nature* 482: 462-63.
- Carroll, N. (1996). "Prospects for Film Theory: A Personal Assessment." In D. Bordwell & N. Carroll, eds., *Post-Theory: Reconstructing Film Studies*, 37-68. Madison, Wisconsin: University of Wisconsin Press.
- Chomsky, N. (1959). "Review of Skinner's Verbal Behaviour." *Language* 35: 26-58.
- Chomsky, N. (1965). *Aspects of a Theory of Syntax*. Cambridge, MA.: MIT Press.
- Chomsky, N. (1975a). "Knowledge of Language." In K. Gunderson, ed., *Language, Mind and Knowledge*, 295-320. Minneapolis: University of Minnesota Press.
- Chomsky, N. (1975b). *Reflections on Language*. New York: Random House.
- Chomsky, N. (1980). *Rules and Representations*. New York: Columbia University Press.
- Chomsky, N. (2000). "Linguistics and Brain Science." In A. Marantz, Y. Miyashita & W. O'Neil, eds., *Image, language brain: papers from the First Mind Articulation Project Symposium*, 1-28. Cambridge, MA: MIT Press.
- Chomsky, N. (2007). "Language and Thought: Descartes and Some Reflections on Venerable Themes." In A. Brook, ed., *The Pre-history of Cognition*, 38-66. Basingstoke: Palgrave Macmillan.
- Chomsky, N. (2009). *Cartesian Linguistics: A Chapter in the History of Rational Thought*. Cambridge, UK: Cambridge University Press, 3<sup>rd</sup> edition.
- Currie, G. (1995). *Image and Mind: Film, Philosophy and Cognitive Science*. Cambridge, UK, Cambridge University Press.
- Currie, G. (1999). "Cognitivism." In T. Miller & R. Stam, eds., *A Companion to Film Theory*, 105-122. Oxford, UK: Blackwell Publishers.
- Currie, G. (2004). *Arts and Minds*. Oxford: Oxford University Press.
- Currie, G. (2010). *Narratives & Narrators: A Philosophy of Stories*. Oxford: Oxford University Press.
- Currie, G. (2016). "Methods in the Philosophy of Literature and Film." In H. Cappelen, T. S. Gendler & J. Hawthorne, eds., *The Oxford Handbook of Philosophical Methodology*, 641-656. Oxford: Oxford University Press.

- Graham, G. (2005). *Philosophy of the Arts: An Introduction to Aesthetics*. Oxford: Routledge, 3<sup>rd</sup> edition.
- Harrison, B. (2012). *The Epistemology of Know-How*. Doctoral Thesis. University of Hertfordshire. Available at: <https://core.ac.uk/download/pdf/9840538.pdf>
- Haugeland, J. (1978). "The Nature and Plausibility of Cognitivism." *The Behavior and Brain Sciences* 2: 215-60.
- Kenny, A. (1989). *The Metaphysics of Mind*. Oxford: Oxford University Press.
- Kenny, A. (2009). "Cognitive Scientism." In H. Glock & J. Hyman, eds., *Wittgenstein and Analytic Philosophy: Essays for P.M.S. Hacker*, 250-262. Oxford: Oxford University Press.
- Kieran, M. (2005). *Revealing Art*. London: Routledge.
- Lamarque, P. (2006). "Cognitive Values in the Arts: Marking the Boundaries." In M. Kieran, ed., *Contemporary Debates in Aesthetics and the Philosophy of Art*, 127-142. Oxford: Blackwell.
- Malcolm, N. (1971). "The myth of cognitive processes and structures." In T. Mischel, ed., *Cognitive Development and Epistemology*, 385-392. Academic Press.
- McCulloch, W. S. and Pitts, W. (1943). "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics* 5: 115-33.
- Mikkonen, J. (2013). *The Cognitive Value of Philosophical Fiction*. London: Bloomsbury.
- Mikkonen, J. (2015). "On Studying the Cognitive Value of Literature." *The Journal of Aesthetics and Art Criticism* 73(3): 273-82.
- Philipse, H. (2009). "Can Philosophy Be a Rigorous Science?" *Royal Institute of Philosophy Supplement* 65: 155-70.
- Quine, W. V. (1969). "Epistemology Naturalized." In W. V. Quine, ed., *Ontological Relativity and Other Essays*. New York: Columbia University Press, 69-90.
- Ryle, G. (1949). *The Concept of Mind*. London: Penguin.
- Stanley, J. and Williamson, T. (2001). "Knowing How." *The Journal of Philosophy* 98(8): 411-44.
- Stanley, J. (2011). *Knowing How*. Oxford: Oxford University Press.
- Stich, S. (1978). "Belief and Sub-doxastic States." *Philosophy of Science* 45: 499-518.
- Stokes, D. 2009. "Aesthetics and Cognitive Science." *Philosophy Compass* 4(5): 715-33.
- Turing, A. M. (1937). "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of London Mathematical Society* S2-42: 230-65.
- Turing, A. M. (1950). "Computing Machinery and Intelligence." *Mind* 59: 433-60.
- Vidmar, I. (2013). *The Nature of Fictional Testimony and its Role in Reaching, Fulfilling and Promoting our Epistemic Aims and Values*. Doctoral Thesis. University of Rijeka. Available at: <https://dr.nsk.hr/islandora/object/svkri:2706/preview>.
- Walsh, D. (1969). *Literature and Knowledge*. Middletown: University of Connecticut Press.
- Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., and Gray, J. R. (2008). "The Seductive Allure of Neuroscience Explanations." *Journal of Cognitive Neuroscience* 20(3): 470-77.

- Wilson, C. (1983). "Literature and Knowledge." *Philosophy* 58: 489-96.
- Young, J. (2001). *Art and Knowledge*. London: Routledge.





# Arts & Ents

*James Hamilton*

---

A quick word about the title: it is common among newspapers in England, especially (apparently) among the tabloid presses, to name the pages in which one finds arts reviews and art-critical essays along with parallel essays on the local entertainments on offer “Arts&Ents.” I chose this as a title not because I thought many people would actually know this fact but because it is a colloquialism and, as such, fit in with many of the themes of this essay.

## 1. Introduction: Some Caveats and the Plan<sup>1</sup>

In this essay, I will not be talking about objects, events, activities, genres, or kinds that have cult followings and are presented as items in any sort of *avant-garde* movement. I want us to think about things that are non-controversially both entertaining and popular.

Moreover, for the purposes of this discussion I could have picked popular entertainments consumed by relatively large groups of people “live”—such as classical, rock, or hip hop concerts, or roller derby, or stand-up comedy, or shows by *Cirque du Soleil*. Just as easily I could have picked mass produced entertainments that are widely popular but usually consumed by individuals on their own schedules: magazines, video games, novels, and so on. Neither the form nor the distinction among various modes of consumption is relevant to my topic. Although the same problems can readily be described using other examples, I use certain movies because they are both entertaining and were popular, at least when they first came out.

Here is the plan. In the next three sections I will present and

---

<sup>1</sup> This paper was originally published in the *Popular Inquiry: The Journal of the Aesthetics of Kitsch, Camp and Mass Culture* (2/2018).

explain the first three problems. In the last section I draw out an uncomfortable implication of the conjunction of the second and third problems. The individuals who are made most uncomfortable by this implication are our colleagues who conduct cultural studies scholarship. I do not single out cultural studies scholars for invidious reasons, nor do all of them face this implication in their work. In fact, I hold that if we can see the nature of this implication by focusing on the activities of one particular group that wishes to take the contents of popular culture seriously, *in a certain way*, we can also see that the problem is quite general. For it besets that way of thinking about popular entertainments; and that way of thinking, I believe, is quite common.

## 2. The “High Art *Versus* Low Art” Problem

Many people have made crude equations among the items on the following list of allegedly deficient practices, or at least among some subsets of this list: popular activities, popular art, popular culture, entertainments, mass art, mass culture, low art, low culture. Terms like “popular” and “low” are terms of contrast with terms like “unpopular” (or “elite”) and “high.” What those contrasts consist of in each case has proved difficult to make out, even impossible in some cases. That is why I have characterized such equations as “crude.” And that, I believe, is about the nicest thing one can say about them. To be sure we can make contrasting lists in particular cases: *Alien* in contrast to *The Seventh Seal*, *The Matrix* in contrast to *Blow Up*, and *Do the Right Thing* in contrast to *Night and Fog* perhaps. But that is a very different thing from explaining the contrast among such kinds, or classes, as low and high culture or popular culture and elite culture.

Here are some explanations that have been proposed, in the form of systematic differences between the up and down classes. It has been said that the high *versus* low distinction mirrors the distinction between art *versus* craft, or the distinction between what is sophisticated and difficult *versus* what is plain and easy, or that between what is mass marketed *versus* what is non-commer-

cial, or that between what is familiarly conventional *versus* what is challengingly unconventional, or the distinction between things aimed at producing passive reaction *versus* those demanding active response, or the distinction between those things produced by an individual for her own goals and satisfaction *versus* those things aimed at satisfying others and reinforcing whatever other's goals already happen to be, or the distinction between politically manipulative art *versus* art that serves autonomous ends and respects autonomy in its audiences.<sup>2</sup> These proposals have been made by people with quite varied agendas – from High Modernists to Marxist critics – and by others with less worked out agendas.<sup>3</sup>

It is noteworthy that none of the foregoing proposals is successful. One reason is that each of the suggested explanatory proposals is susceptible to fairly obvious counter-examples.<sup>4</sup> To be sure, these counter-examples themselves are thought to be contentious; and, for sake of argument, let us agree that they might be. Instead, then, turn away from offering counter-examples and turn toward examining arguments.

Actually, two arguments are required if we are to make good on the contrasts just mentioned. One must argue that anything produced for entertainment is, for that reason, likely to be popular. And one must argue that there is some sort of plausible link between the *descriptive fact* that an entertainment is popular and the *normative claim* that it is morally, socially, or politically deficient, for the very reason that it is popular. But arguments for both a descriptively adequate equation and for a decisive normative judgment turn out to be very weak. Consider two examples, each of which begins from a widely accepted premise.

Consider the fact that many of us associate entertainment with what is produced or performed for an audience. Indeed this association is often stated as part of the standard meanings of “enter-

---

2 Fisher, 2013.

3 Horkheimer and Adorno, 2007; Adorno, 1975. Adorno is merely the most salient of the figures who have held this view, and not only because he was both a modernist and a Marxist.

4 Novitz, 2005, pp. 734-735, 738-739; Fisher, 2013, 476-478.

tainment” in our dictionaries.<sup>5</sup> But here is an argument so bad that you’d have to think no one would credit it:

P1: If something is an “entertainment” it is aimed at being performed for an audience.

P2: If it is aimed at being performed for an audience, it falls within the scope of the popular.

----

C: Entertainments are aimed at falling within the scope of the popular.

The first premise is the widely-held belief just mentioned and the conclusion is the descriptive claim that is needed for us to go on (see below). The second premise fills in the necessary argument step between the widely-held belief and that conclusion. And the argument is valid.

But the second premise is false, as can be shown by considering only a few theatrical cases. Historically many theatrical entertainments that were aimed at audiences have not been aimed at large audiences nor considered fit or appropriate for popular consumption. “Closet dramas,” for example, are meant to either be un-staged and read privately or in small groups or are meant to be staged but only for small groups. The Kings Men was a theatre company that played only, or at least mostly, for the royal court of England from 1603 to 1608, when they began performing in the Blackfriars Theatre which probably had room for only several hundred spectators. In contrast the Globe, where they did not initially perform, held upwards of 2500-3000 spectators.

I am inclined to think, without further argument, that the connection between what is entertaining and what is popular is purely contingent. And I assume this contingent, non-conceptual, relationship in the rest of this essay. I do not deny that many enter-

---

5 For examples, the Merriam Webster Online Dictionary: <https://www.merriam-webster.com/dictionary/entertainment>, and Wikipedia’s entry on “Entertainment,” <https://en.wikipedia.org/wiki/Entertainment>.

tainments are also popular – indeed it is those I wish to talk about in what follows. And this connection, when it *has* occurred, is an important historical fact. There is plenty of room for informative discussions of when and how entertainments have become widely popular. I just do not think the connection is conceptual; hence it is not more than a matter of historical contingency.

Consider next the normative argument. Why do many people think that there is something “low” or at least unseemly about *popular* entertainment? One reason is that they think if something is a popular entertainment, it must be aiming at producing or inducing specific effects which render whatever is performed easy for that audience to grasp. But now look at this as an actual argument.

P1: If something is a popular entertainment, it must be aiming at producing or inducing specific *effects* that render it easy to grasp.

P2: In producing or inducing *effects* that render it easy to grasp, its producers are *not* presenting a product for autonomous agents to consider accepting but rather at mere things occupying places in a causal chain.

----

C: If something is a popular entertainment, then it is *not* aimed at presenting a product for autonomous agents to consider accepting but rather at mere things occupying places in a causal chain.

Again, the first premise is just the widely-held belief and the conclusion is what we are supposed to be led to by that widely-held belief. Presumably, there is something “low” about causing an audience to respond in various ways in contrast to asking them, as autonomous agents, to consider accepting whatever is presented to them. And once again, the second premise is the necessary step to take us from the widely-held belief to the desired conclusion.

But this is a bad argument. It too is valid, but its second premise is false, and any variation on that premise that preserves the argument’s validity is also false. The problem with the second premise stems from thinking that either by aiming at inducing effects or by making the entertainment easy to grasp, the autonomy of each

audience member is somehow subverted. This is a mistake.<sup>6</sup> The fact that some responses in audiences for art or entertainment are caused is not grounds for thinking every such response is caused. Moreover, causation of an effect inducing a response is very likely to be well below the level of consciousness. While the critic of causal effects thinks of such causes as subverting autonomy *for this very reason*, those critics have not given much thought to precisely *how* that response is configured in an audience member's attempt to figure out the work of art or entertainment. In fact, once we ask how anything that is responded to sub-consciously gets into the conscious reasoning of audiences, the only plausible models are ones that treat the process as largely an exercise of critical reasoning, and hardly subversive of autonomy at all.

Thus, I conclude we should deny that, when the contingent connection between entertainment and popularity does obtain, by necessity it also has normative consequences.

Many think this conclusion is mistaken because they believe that audience passivity – and hence the purely causal story of entertainment effects – is simply part of what it is for something to be a popular entertainment.<sup>7</sup> This leads Stephen Bates and Anthony Ferri, for example, to offer the following tentative definition of “entertainment.”

We suggest that entertainment, defined in largely objective terms, entails communication via external stimuli, which reaches a generally passive audience and gives some portion of that audience pleasure.<sup>8</sup>

But this definition, because of its *a priori* inclusion of audience passivity, excludes more than it should. For one thing, there are many interactive entertainments. And, in any case, the relationship between popular entertainment and the passivity of audience perception, when it *is* passive, is as contingent a matter as that between entertainment and popularity itself. When or if there is a connection between the two is an empirical question; not one to be settled

---

6 Carroll, 1998, p. 15-109.

7 Brock and Livingston, 2000.

8 Bates and Ferri, 2010, p. 15.

by *a priori* reasoning.

Given this state of affairs, the very attempt to provide principled, *a priori* grounds for the “high art *versus* low entertainment” distinction seems less than helpful, at best. And it may be much worse. For, attempting this kind of fruitless classification project can divert attention from what is interesting and possibly valuable in the particular issues, instances, genres, or kinds of entertainment that should interest us.<sup>9</sup>

### 3. The Problem of Figuring out How These Cases Work

One of those issues that should interest us is this problem: How, precisely, do we learn from works of art and entertainments when they express a deep philosophical insight? Here are two claims that many of us are likely to believe true. First, attending to cases of works of art with a reflective eye can tell us something about ourselves, our lives, our politics, our economics, our conceptual schemes. Second, this is as true of popular entertainments as it is of things that have been collected under the concept “Art.”

Since the second claim is perhaps not as obvious to everyone as it seems to me, consider Stephen Mulhall’s analysis of *Alien*. Mulhall examines how this movie and its three sequels set forth themes of “the relation between human identity, integrity, and embodiment, as encountered in the field of our fantasies of sexual intercourse, pregnancy, and birth,” and thereby “evoke undismissable questions about what it is to be human.”<sup>10</sup> When reflecting on Mulhall’s description and analysis of the *Alien* quartet, the claim that we can learn about ourselves from popular entertainment as well from art may seem entirely obvious.

However, there are two contrasting ways to take these claims, first as holding that movies can *illustrate* such issues as Mulhall suggests, or second as holding that movies can *examine* such issues. This contrast can be framed as between a pair of questions: “can a

<sup>9</sup> Novitz, 2005, p. 740; Gould, 1999, p. 120.

<sup>10</sup> Mulhall, 2006; Mulhall, 2015.



movie *be* philosophical?” and “can a movie *do* philosophy?”<sup>11</sup> Or the contrast can be framed as this pair of questions, “can a movie occupy the same reflective space as philosophy?” and “can a movie occupy the same reflective space as philosophy *in the way that philosophy does*?”<sup>12</sup> I will refer to the second members of these pairs of questions as “The Philosophy Question.”<sup>13</sup>

One strategy for defending an affirmative answer to The Philosophy Question is to regard movies as “thought experiments.”<sup>14</sup> The point of both movies and thought experiments can be the same, some think, in that each can function as a device for teaching us about the application and limits of our concepts and, in particular, for teaching us how to get a clear view of some phenomena and expose inconsistencies or incoherencies in some alternative *conceptions* of those phenomena. In this way, a movie can stand in for an argument.<sup>15</sup>

In explaining the “thought experiment” strategy, Murray Smith references a thought experiment concerning the peasant and a king “switching places” that Bernard Williams offers as a way of examining mind-body dualism.<sup>16</sup> Smith comments that

Williams only needs to elaborate and extend the basic premise of the thought experiment over a few sentences in order to reveal the conceptual confusion on which (he argues) the dualistic conception of personal identity-of the self as a disembodied soul-rests.<sup>17</sup>

---

11 Shaw, 2006.

12 Smith, 2006.

13 The second way of framing the contrast is a bit more revealing than the first, I believe; but I am indifferent to that issue here.

14 Carroll, 2002.

15 Nor is this the only way that movies, and other thought experiments, might be said to stand in for arguments. Noël Carroll does argue for this way of defending the view; but Elke Brendel (in her 2004 essay, “Intuition Pumps and the Proper Use of Thought Experiments,” *Dialectica* 58/1: 89-108) holds that because thought experiments study “the functional dependencies of variables by planned and controlled data change...[and] depend on some background assumptions or background theories,” they are best seen as arguments with premises that can be directly challenged or supported. (Brendel, 91.)

16 Williams, 1973, p. 11-12.

17 Smith, 2006, p. 35.

However, when Smith employs the strategy for thinking about how the movie *All of Me* examines mind-body dualism, he concludes the proposed strategy is mistaken. And this is because *this* movie's "thought experiment," if we can call it that, does not aim at clarifying the limits of our concepts but at presenting such things as "complexity, ingenuity, inventiveness, density, ambiguity, . . . profundity," and above all "paradox."<sup>18</sup> To achieve that, one will have to extend the thought experiment well beyond only a very few sentences. Smith cautions that not all movies presenting thought experiments aim at presenting just that particular set of features. He argues that the aims of philosophical thought experiments, such as Williams' are to limn concepts whereas the aims of the thought experiments, when they occur in movies, have to be analyzed in terms of aiming at presenting *artistic* features.

However, others have claimed that some movies *do* aim at examining the limits of our concepts. So it is, some have claimed, with *The Matrix*.<sup>19</sup> It displays and examines the grounds for Cartesian skepticism and 'runs' a variant on Descartes' "evil demon" hypothesis to see how it plays out.

A second strategy for responding to The Philosophy Question is to regard movies as the special kind of thought experiment that Daniel Dennett has called "intuition pumps," that subset of thought experiments useful for (often illicitly) getting us to agree to judgments we might not have thought to agree to before and thus to be prepared to accept one sort of theory over another.<sup>20</sup> For example, up through *Alien3* (1992) the lines between humans and aliens have been transgressed, to be sure, but they have been restored and are in order; and so far as we were concerned that was the end of the matter. But then *Alien: Resurrection*, it might be said, prompts us to consider judgments about Ripley we would not have considered making had she been a hybrid in one of the first two movies. And that may get us to look more deeply at our theories of what it

---

<sup>18</sup> Smith, 2006, p. 36-40.

<sup>19</sup> Wartenburg, 2006.

<sup>20</sup> Dennett, 1995.

is to be human.<sup>21</sup> In this way we might think the movie acts like an intuition pump, not by standing in for arguments, but by clearing conceptual space and demonstrating the need for arguments.

Both of these strategies hold that the only real difference between thought experiments and intuition pumps with which philosophers are familiar and the things we engage when watching movies, reading novels, playing video games and the like has to do with the level of detail. Popular entertainments just contain more complex and sophisticated stories than those philosophers usually appeal to when they do philosophy. This observation holds that the difference, although real, is of minimal importance. Of course if Smith is right, that the aims of thought experiments in movies are to present artistic concepts and that to fulfill those aims movies must be replete with detail, then this is not a difference with minimal importance. But I do not know that Smith is right.

However, to assess these strategies, we should ask how thought experiments and intuition pumps actually work. In particular, I suggest we should examine the difference between the way *they* work and the way a parable works. This is because narratives come in kinds, or “genres,” and each genre sets up different expectations. And it could be that all that thought experiments and intuition pumps share with more complicated items of popular entertainment is the simple fact they are frequently structured as narratives. And this, I believe, is too simple, even if true.

Accordingly, let me remind you of a Biblical parable – the parable Nathan told to King David. David, you may recall, had sent a rival captain to war so that he could seduce his rival’s wife, Bathsheba. The prophet Nathan’s story, however, concerned a man stealing his neighbor’s sheep. At the end of the parable, having secured King David’s condemnation of the thief, Nathan exclaims, “Thou art the man.”

That’s part of how parables work. They end with one person saying to another something like “Thou art the man” with the one hoping to get the other to see *herself* in the story. This is something

---

21 Mulhall, 2007.

quite different from getting her to recognize her intuitions could be very different from what she might have thought they were or from getting her to recognize the limits of a favored conception of this or that sort of thing. In those cases what we hope to achieve has nothing in particular to do with *her*. But parables are personal even when they are broadcast.

So it is, I think, that Dan Flory shows us how Spike Lee's *Do the Right Thing* invites us to see ourselves, at least those of us who are white males of a certain age, as racists of the same kind and in the same way as Danny Aiello's character, Sal.<sup>22</sup> It is not a failing of the story or of its mode of presentation if I am not able to grasp the fact. It is, instead, more like the way the "seed sewn" in parables "falls upon stony ground," as the Bible has it.

Now for some, regarding the way we learn from movies as being like the way we learn about ourselves from parables will not be an attractive response to The Philosophy Question. For, if we think the parable analysis is right, we also have to accept the fact it does *not* show us that movies can reflect on our lives in the same way that philosophy does. Parables do not stand in for arguments, nor do they clear the ground for arguments; they convict us of our conditions.

In addition, parables frequently function by having us sympathize with characters. A familiar pattern is for us to sympathize with one character only to be confronted by the fact that in the story we are depicted in the role of the unsympathetic asshole. And this, or something like it, happens all the time, especially in popular narrative entertainment. In contrast, there are typically no appeals to sympathy for anyone in thought experiments or intuition pumps. In the famous "trolley cases," for example, we have as much reason to sympathize with unfortunate one caught on one track ahead as we do to sympathize with those on the other track who will surely die if you or I do not throw the switch and divert the trolley onto the first track. This is not to say one just couldn't come up with a thought experiment or an intuition pump that depended

---

22 Flory, 2006.

on appeals to sympathy; but such cases are likely to involve examinations of the concept referred to be the word “sympathy” itself.

These reflections might suggest to us the empirical guess that there are very few works of art or of popular movies that function as thought experiments or intuition pumps. And one option not considered so far is that, when movies tell their audiences what they have to tell them, it might be in the form of a parable rather than a thought experiment or an intuition pump. For such movies seem to be saying something like this: “Get out of here and do this or be like that.”

Notice, however, that if *any* of these analyses is on the right track, we will be pushed to reassess another familiar claim about popular entertainments, namely, that the ease of access so frequently characteristic of them makes them systematically *defective*. In contrast, one thing we know makes thought experiments, intuition pumps, *and* parables effective is their *perspicuity*. And any way we have of unpacking that thought suggests that ease of access may be, and in some cases and on some occasions must be, a *positive feature*. To be sure, there are differences between pleasures attending to ease of access and pleasures attending to difficulty of access. But, as Timothy Gould reminds us, it is not as though the relation between the ease of access and the easiness of the pleasures taken in that which is easily accessed is a straightforward matter.<sup>23</sup>

#### 4. The “Two-audiences” Problem

The “two-audiences” problem arises from the simple fact that there are cases of popular entertainment that are attended to both by people who normally attend only to traditional instances of art *and* by people who normally only attend to popular entertainments. When this simple fact is conjoined to a plausible back-story about how we determine what the audience for a particular art form or bit of entertainment is, we get the problem.

---

23 Gould, 1999, p. 121.

I rely on some work of Ted Cohen's for the back-story.<sup>24</sup> There are several parts to Cohen's story. First, when we respond to works of art we respond with others and form with others the audiences for those works. Second, what we respond to in the role of audience for a thing is largely determined by who we are, where we come from, what our ambitions are, and so on. Third, responding in the role of audience to one kind of thing as opposed to another also partially determines our own *sense* of who we are, of what audiences we belong to, and with. This combination of claims, Cohen thinks, partially explains why our taste in movies, music, TV programs, and in jokes matters so much to us.

Against that background, what are we to make of the simple fact that there are cases of popular entertainment that are attended to both by people who normally attend only to traditional instances of art *and* by people who normally only attend to popular entertainments? In a different article Cohen calls these "fancy" and "plain" audiences. To be sure, as Cohen reminds us, "fancy audiences often like both high and low movies, and ... at least some very high movies appeal to both fancy and plain audiences." But the fact of the existence of two audiences, he also observes, can lead to a problem, for it can "lead us to wonder (1) whether it is exactly the same auditor who likes Bach's unaccompanied cello music and Leon Redbone's blues, and (2) whether *North by Northwest* is the same work for the fellow who enjoys it as a nice example of Hollywood fluff and for the one who finds it a profound meditation on American identity."<sup>25</sup>

A caveat and an observation are in order. Resolving the second of Cohen's worries turns on defending some story about work-identity. I pass on that metaphysical project in this essay and focus only on the first of his concerns. I simply try to determine what Cohen means by asking "Is it the same auditor"?

The observation is that, although the question Cohen raises – "is it the same auditor...?" – is framed as a question about high

---

24 Cohen, 1983.

25 Cohen, 1999.

and low art, it is not dependent on a high *versus* low distinction nor on talk about art, *per se*. For the question could arise with exactly the same force, and for the same reasons, were his examples that of liking both Bach's unaccompanied cello suites and Schoenberg's *Pierrot Lunaire*, or were they that of liking both Leon Redbone's blues and Muddy Waters's, or were his examples that of liking both *WWE* and *Rollergirls*. Any pair of examples sufficiently different in aim, medium, style, function, or content will do.

So, am I the same auditor when I "like" Bach's unaccompanied cello suites and I also "like" Leon Redbone's blues – or any other pair of sufficiently disparate objects? Well, yes, of course. But it is also true that I have some work to do in order to maintain a coherent sense of my beliefs and attitudes. This challenge arises because I will have enjoyed the pair of disparate objects for different reasons and in different ways.

It is important that these are *considered* preferences. The pleasure I derive from each "is not," as Stephen Davies reminds us about the enjoyment of music, "some *frisson* to which the music...stands *merely* as the cause or occasion, for whereas such pleasure is indifferent to its cause, the pleasure of appreciating a [piece of music]... is not indifferent to the individuality of its object."<sup>26</sup> But precisely because I have deeply enjoyed some disparate objects for significantly different reasons, as I might well deeply enjoy different kinds of music, novels, movies, TV, and jokes, I *can be* torn between the audiences to which I belong. At the very least, in order to think that through, I will have to think a good deal harder about *myself* than I may have done hitherto. This challenge to think my way through to a coherent self when my *considered* preferences vary so widely from one another is the root of the two-audiences problem.

## 5. A "Perfect Storm" of Sorts

One variant of the two-audiences problem seems especially important for understanding the position of some scholars work-

---

<sup>26</sup> Davies, 1987, p. 316.

ing in cultural studies – including both philosophers and others more routinely associated with cultural studies programs.

Not a few papers and workshops presented in conferences I have attended have concerned social groups and the popular activities or institutions that bind them together, whether they know that or not. Some papers and workshops have been about the techniques of building works of art, popular or otherwise. But the majority of papers and panels have been about novels, plays, genres, or performances or performance kinds and how certain “readings,” or interpretations, of them show that those objects provide or prevent access to strategies “for resistance to dominant ideological constructions of...class, gender, [race] and family.”<sup>27</sup>

Often, the scholars working in those presentations have also taken up a certain kind of role as audience for those objects that is different from the role taken up for those same objects by the people who make those objects instances of popular entertainment. But, if Cohen is right, this should strike us as odd. Timothy Gould puts a related point this way:

...most investigators of popular culture do not belong to the classes of modern society for whom popular culture is the only form of culture. Investigating popular art requires certain sorts of decisions – including, frequently, career decisions. [And so,] we find ourselves having to justify, or at least extenuate, the forms in which we pursue our interests.<sup>28</sup>

And Gould suggests we should be cautious about these investigations because the academic investigator of popular entertainment must be *self-conscious* in her engagement with that entertainment in a way the popular audience for it never is *and* because we should have worries about the distortions of objects, activities, *and of ourselves* that can take place in the self-conscious engagement with objects or activities of any kind. Whether this latter conclusion is warranted or not, a particular combination of the parable account of how works of art and popular entertainments teach us and a plausible account of how we determine what audience we be-

27 Knight, 2005, p. 789.

28 Gould, 1999, p. 122.



long with *can* create an intellectual and practical “perfect storm.”

The teller of a parable is not in the same relation to the subject at hand as the person at whom the parable is directed: someone telling a parable will say “*thou* art the man” or “go *thou* and do likewise” but not “hey, that’s *me!*” In contrast, the person who provides an intuition pump or a thought experiment is explicitly aiming at saying “hey, look at this, this is *us*.” Although in different epistemic positions, the person providing a thought experiment or intuition pump is in the exact same relation of possession to the matter at hand as the persons for whom the experiment or pump is provided. Although the person for whom it is provided may not see clearly how some aspect of her concepts plays out, the concepts involved are still *her* concepts every bit as much as they are the concepts of the person who provides the case and asks her to examine how it plays out. Moreover, thought experiments and intuition pumps do not routinely issue in normative judgments; parables always do. Finally, one can just as easily offer a thought experiment or an intuition pump for oneself as for others; indeed that is usually the temporal order in which they *are* presented. None of this works well for parables.

These thoughts suggest that, if the parable analysis of how we learn about ourselves from art and entertainment is the right way to explicate how they inform us of important philosophical insights, there could be a deep question here about who is doing the learning and who is being taught when “*we*” focus on popular entertainments, and provide analyses or interpretations of any work of art or item of popular entertainment. Who, in these cases, is “the man”? *Who* could be the referent of “*thou*” in “go *thou* and do likewise”? What is the proper identity-establishing *audience* for such analyses?

Quasi-formally, the situation looks like this:

- (A) Artwork *A* functions as a parable convicting *S* and people like *S* of their condition.
- (B) I am the presenter of artwork *A* *and* I endorse the message of *A*

(hence, I am *not-S* or *not-S-like*).<sup>29</sup>

(C) I am the target of artwork *A* (hence, I *am S* or *S-like*).

Temporally considered, of course, this might be consistent, as follows:

(A), (then) $\Rightarrow$  (B), (then) $\Rightarrow$  (C).

But when considered *atemporally*, we get the inconsistent set,  $\{(A), (B), (C)\}$ .

To be sure, in analyzing what a movie can teach us about ourselves, the philosopher or cultural studies scholar need not be *telling* that parable even if she might also endorse it. Nor in telling it need she be *endorsing* it. That is, even if she is explaining how *Do the Right Thing* convicts some in its popular audience of its condition, or at least leads them to insights about themselves they had resisted, *she* need not be making a normative judgment of the “go thou and do likewise” form. Her work does have normative content, to be sure, but it can be descriptive of that content. And, of course, we may not be interested in holding a coherent sense of our beliefs and attitudes. It may also be the case that some works of art or of popular entertainment are, as Cohen himself suggests, “coded” to be given different interpretations by differently situated individuals. If either of these is true, then the problem I have been describing goes away.<sup>30</sup> But, if we do engage in telling the parable and endorsing its normative stance, if and to the degree we are interested in holding on to a coherent sense of self, and if the work of art or popular entertainment does not code for various interpretations in any obvious way, then anyone can find themselves stuck in the problem I have described.

The problem, in short, arises most acutely for those who wish not only to think through a work of art or a popular entertainment to a normative judgment but also to *endorse* a judgment of the kind

---

<sup>29</sup> An earlier version of this quasi-formal representation of the problem, as David Davies pointed out to me, had “or” where “and” should be (and now is) in statement (B).

<sup>30</sup> I am grateful to Ted Gracyk for reminding me of this aspect of Cohen’s own “solution.”

made through that entertainment, rendering the analysis and the parable one. It seems to arise especially for those who focus their analyses on why “we” are drawn to such entertainments. It is entirely unclear for whom or to whom they *could be* speaking. And although that just seems an odd result, I do not know how to avoid it.<sup>31</sup>

---

<sup>31</sup> This paper was originally written for and presented to the Cultural Studies Symposium at Kansas State University in 2007. A slightly revised version was presented at the Theatre Studies International conference at Leeds, United Kingdom, in 2013. Subsequently, I have presented later versions of this work at the Pacific Division of the American Society for Aesthetics in 2015, at Auburn University in 2016, and to the conference on the Aesthetics of Popular Culture in Warsaw, Poland, in 2018. The two early versions were crafted for presentations to cultural studies conferences. The later three versions were substantially re-written for philosophical audiences. I am grateful for the comments I received at all these venues. All of them have improved the paper. And I am especially grateful in this regard for the comments of Ted Gracyk made at the Pacific Division of the American Society for Aesthetics in 2015 and David Davies at the conference on the Aesthetics of Popular Culture in Warsaw, Poland, in 2018.

## Bibliography

- Adorno, T. W. (1975). "Culture Industry Reconsidered." *New German Critique* 6: 12-19.
- Bates, S. and Ferri, A. (2010). "What's Entertainment? Notes Toward a Definition." *Studies in Popular Culture* 33(1): 1-20.
- Brendel, E. (2004). "Intuition Pumps and the Proper Use of Thought Experiments." *Dialectica* 58(1): 89-108.
- Brock, T. and Livingston, S. (2000). "The Need for Entertainment Scale." In L. J. Shrum, ed., *The Psychology of Entertainment Media: Blurring the Lines between Entertainment and Persuasion*, 255-74. Mahwah, NJ: Erlbaum.
- Carroll, N. (1998). *A Philosophy of Mass Art*. Oxford: Clarendon Press.
- Carroll, N. (2002). "The Wheel of Virtue." *Journal of Aesthetics and Art Criticism* 60(1): 3-26.
- Cohen, T. (1983). "Jokes." In E. Schaper, ed., *Pleasure, Preference, and Values*, 120-36. Cambridge, England: Cambridge University Press.
- Cohen, T. (1999). "High and Low Art: High and Low Audiences." *Journal of Aesthetics and Art Criticism* 60(1): 137-43.
- Davies, S. (1987). "The Evolution of Music." In P. Alperson, ed., *What is Music?* 303-26. University Park, Pennsylvania: Pennsylvania State University Press.
- Dennett, D. (1995). "Intuition Pumps." In J. Brockman, ed., *The Third Culture: Beyond the Scientific Revolution*, 180-97. New York City: Simon and Schuster, Inc.
- Entertainment. (n.d.). In *the Merriam Webster Online Dictionary*. Retrieved October 11, 2017, from <https://www.merriam-webster.com/dictionary/entertainment>
- Entertainment. (n.d.). In *Wikipedia*. Retrieved October 11, 2017, from <https://en.wikipedia.org/wiki/Entertainment>
- Fisher, J. (2013). "High Art Versus Low Art." In B. Gaut & D. MacIver Lopes, eds., *Routledge Companion to Aesthetics*, 473-84. Abington, Oxon: Routledge.
- Flory, D. (2006). "Spike Lee and the Sympathetic Racist." *Journal of Aesthetics and Art Criticism* 64(1): 67-79.
- Gould, T. (1999). "Pursuing the Popular." *Journal of Aesthetics and Art Criticism* 57(2): 119-135.
- Horkheimer, M. and Adorno, T. W. (2007). "The Culture Industry: Enlightenment as Mass Deception," In G. S. Noerr, ed., *Dialectic of Enlightenment: Philosophical Fragments*, 94-136. trans. E. Jephcott, Stanford: Stanford University Press, 2002.
- Knight, D. (2005). "Aesthetics and Cultural Studies." In J. Levinson, ed., *The Oxford Handbook of Aesthetics*, 783-95. Oxford, England: Oxford University Press.
- Mulhall, S. (2006). "The Impersonation of Personality." *Journal of Aesthetics and Art Criticism* 64(1): 97-110.
- Mulhall, S. (2007). "Film as Philosophy: the very idea." *Proceedings of the Aristotelian Society* 107(3): 279-94.
- Mulhall, S. (2015). *On Film*. London: Routledge Publishers.

- Novitz, D. (2005). "The Aesthetics of Popular Art." In J. Levinson, ed., *Oxford Handbook of Aesthetics*, 732-47. Oxford, Oxford University Press.
- Shaw, D. (2006). "On Being Philosophical and 'Being John Malkovich.'" *Journal of Aesthetics and Art Criticism* 64(1): 111-118.
- Shusterman, R. (2009). "Popular art." In S. Davies, K. M. Higgins, R. Hopkins, R. Stecker & D. E. Cooper, eds., *Blackwell's Companion to Aesthetics*, 476-78. Oxford, UK and Malden, Massachusetts: Wiley-Blackwell.
- Smith, M. (2006). "Film Art, Argument, and Ambiguity." *Journal of Aesthetics and Art Criticism* 64(1): 33-42.
- Wartenburg, T. (2006). "Beyond Mere Illustration." *Journal of Aesthetics and Art Criticism* 64(1):19-32.
- Williams, B. (1973) "Personal Identity and Individuation." In B. Williams, ed., *Problems of the Self: Philosophical Papers 1956-72*, 11-12. Cambridge, England: Cambridge University Press.

## Index

Aesthetic anti-cognitivism	22, 74, 78
Aesthetic cognitivism	12, 13, 15, 18, 19, 21, 77
Aristotle	32, 35, 59, 87, 89
Art's social engagement and value; impact on prosocial behaviour	11–14, 17, 31, 32, 39–41, 45, 49, 55, 77, 90–104, 109–118, 120–126, 137–145, 147, 152, 153, 155, 162, 164, 165, 167, 170, 203
Audience, audience's engagement	16, 17–20, 23, 25, 27, 29, 30, 32, 37, 38, 45, 49, 53, 57, 60, 63, 118, 120, 132, 139, 143, 147, 152, 166, 169, 170, 203–206, 212–218
Cinematic cognitivism	15, 175, 176, 196
Cinematic humanism	15, 175–178, 195 – 196
Cognitive dimension / value of artwork	11, 12, 15–26, 28–31, 33–56, 58, 64, 66–72, 74–79, 124, 126, 157, 160, 161, 167, 168, 175
Cognitive film theory	176, 190, 192
Empathy, prosocial behavior	13, 33, 60, 87, 91–94, 96, 98–104,
Ethical dimension/value of art	11, 12, 15, 18, 22, 23, 54, 57, 63, 160, 161 – 165, 167, 169, 178
Humanism, direct and indirect	33, 40–43, 46, 48, 54, 66, 72
Knowledge/acknowledgment through art	12, 18, 21–22, 25–9, 43, 47, 50–53, 57, 59, 62, 66, 69–72, 88, 90, 157, 175, 177–180, 182–186, 188
Literary cognitivism	15, 41, 42, 47, 77, 177
Plato	13, 23, 24, 25, 29, 30, 32, 33, 54, 55, 56, 57, 58, 59, 60, 72, 73, 75, 76, 104, 115
Political dimension/value of art	14, 21, 23, 30, 54, 67, 68, 117, 119, 121, 123, 125, 138, 140–146, 148, 152, 155, 156, 159–171, 188, 203
Thought experiment	16, 24, 64–65, 208–212, 216
Understanding through art	13, 15, 17, 22, 25, 28, 32, 33, 35–41, 46–55, 60, 64, 69, 76, 88, 89, 91, 94, 95, 102–104, 116, 118, 130, 158, 160, 164, 168, 170, 175, 178, 185, 188, 194, 195, 196







