

Georgia State University

ScholarWorks @ Georgia State University

---

Philosophy Theses

Department of Philosophy

---

Summer 8-12-2014

## Why Immoral Art Cannot Morally Harm Us

Maria Caruso

Follow this and additional works at: [https://scholarworks.gsu.edu/philosophy\\_theses](https://scholarworks.gsu.edu/philosophy_theses)

---

### Recommended Citation

Caruso, Maria, "Why Immoral Art Cannot Morally Harm Us." Thesis, Georgia State University, 2014.  
doi: <https://doi.org/10.57709/5670223>

This Thesis is brought to you for free and open access by the Department of Philosophy at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact [scholarworks@gsu.edu](mailto:scholarworks@gsu.edu).

# WHY IMMORAL ART CANNOT MORALLY HARM US

by

MARIA CARUSO

Under the Direction of Daniel Weiskopf

## ABSTRACT

Both philosophers and literary critics have championed artworks as necessary to moral education. As a result many of these critics believe that art that is bad or immoral can causally affect our character, resulting in moral harm. Moral harm is the idea that artworks possess a strong disposition to affect our moral beliefs such that we are less able to distinguish between what is good and what is bad. I examine this concept of moral harm and argue that immoral artworks do not have this kind of causal power over our moral beliefs. Proponents of the moral harm thesis are in error to attribute such a power to artworks. Additionally, I propose a definition of immoral artworks consistent with moral harm, as well as discuss the distinction between immoral artworks and artworks that are merely elicit disgust or offense.

INDEX WORDS: Aesthetics, Art and morality, Ethicism, Moral harm, Moral disgust, Normative offense

WHY IMMORAL ART CANNOT MORALLY HARM US

by

MARIA CARUSO

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2014

Copyright by  
Maria Karin Caruso  
2014

WHY IMMORAL ART CANNOT MORALLY HARM US

by

MARIA CARUSO

Committee Chair: Daniel Weiskopf

Committee: Jessica Berry

Andrew I. Cohen

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2014

## ACKNOWLEDGEMENTS

I am grateful to my father, mother, and brothers for their encouragement and love. I would like to thank Dr. Andrew I. Cohen for being great mentor and teacher. He has graciously read countless pages of my work and I would not have succeeded without his support. I am grateful to Dr. Daniel Weiskopf for working with me on such an ambitious project and for his keen philosophical eye. I am also grateful to Dr. Jessica Berry for her help in putting together my graduate application, as well as her help on this thesis. Additionally, I want to thank Dr. Sebastian Rand, Dr. George Rainbolt, Dr. Sandra Dwyer, and Dr. Tim O'Keefe for being a source of encouragement and wisdom during and after my undergraduate years. I am also indebted to Claire Kooy, Felicia Thomas, and Ellen Logan for their administrative support. I am thankful to my friends Chetan Cetty, Sean Huff, and Cami Koepke for laughing at my jokes, listening to my rantings, and generally being all-around good human beings. Finally, I am grateful to Pierce Randall, my partner, for being the best.

## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS.....</b>		<b>iv</b>
<b>LIST OF FIGURES.....</b>		<b>vi</b>
<b>1 INTRODUCTION.....</b>		<b>1</b>
<b>1.1 The History of the Moral Harm Thesis.....</b>		<b>3</b>
<b>1.2 The Structure of the Argument.....</b>		<b>7</b>
<b>2 THE POSSIBILITY OF IMMORAL ART WORKS.....</b>		<b>9</b>
<b>2.1 Moral Harm, Artworks, and Agency.....</b>		<b>9</b>
<b>2.2 An Overview of Ethicism.....</b>		<b>12</b>
<b>2.3 The Ethicist Thesis and the Causal Claim.....</b>		<b>15</b>
<b>3 DISTINGUISHING DISGUST AND OFFENSE FROM MORALITY.....</b>		<b>18</b>
<b>3.1 Disgust and Morality.....</b>		<b>20</b>
<b>3.2 Offense and Morality.....</b>		<b>26</b>
<b>4 AGAINST MORAL HARM.....</b>		<b>31</b>
<b>4.1 Case Examples of Immoral Art.....</b>		<b>31</b>
<b>4.2 Challenging the Moral Harm Thesis.....</b>		<b>36</b>
<i>4.2.1 Imagination's Influence on Belief.....</i>		<i>37</i>
<i>4.2.2 Emotion's Influence on Belief.....</i>		<i>45</i>
<b>5 CONCLUSION.....</b>		<b>50</b>
<b>REFERENCES.....</b>		<b>53</b>

## LIST OF FIGURES

Figure 1 Dread Scott, <i>What is the Proper Way to Display a US Flag?</i> .....	2
Figure 2 Cover illustration for <i>The Story of Little Black Sambo</i> .....	14
Figure 3 Robert Mapplethorpe, <i>Helmut</i> .....	18
Figure 4 Andres Serrano, <i>Immersion (Piss Christ)</i> .....	19
Figure 5 Nan Goldin, <i>Nan One Month After Being Battered, 1984</i> .....	24
Figure 6 Vanessa Beecroft, <i>VB46</i> .....	32
Figure 7 Allen Jones, <i>Table, Chair, and Hatstand</i> .....	34
Figure 8 Mizuno To, Woodblock print depicting the Battle of P'yongyang during the first Sino-Japanese War.....	35

## 1 INTRODUCTION

Arthur Danto, in “Dangerous Art,” writes that the rise of aesthetic theory has paralleled the rise of expressive freedoms in the state. For Danto, this is problematic. Contrary to appearances, expressive freedoms are not so free. Theory, according to Danto, relegates these newly freed expressive works to background noise. Transgressive art crosses no real boundaries on Danto's analysis; theory strips potentially dangerous, political, inflammatory or immoral artwork of its power. Furthermore, the viewer is told that he or she must approach art only from a theoretical standpoint. The aesthetic point of view separates the viewer from the message of the art: “in aesthetization we have the supreme disenfranchisement, for the work of art is reduced by its means to something that exists for gratification” (Danto 1992, 186). Through the disinterested gaze of the viewer, art loses its gravity; aesthetic distance “[allows] the artist perfect freedom, but at the cost of total and logically guaranteed harmlessness” (Danto 188). In effect, theory declaws art.

Viewers and critics appeal to aesthetic theory when they are struck by works that appear to be immoral, distasteful or foreign. For instance, in the defense of Robert Mapplethorpe's photographs during a 1990 obscenity trial, the curator's lawyers focused not on justifying Mapplethorpe's or the curator's rights to depict “extreme sexual acts” but instead on the refined and beautiful staging of the scenes (Julias 2002, 50). We use theory to explain, remove, or shift the focus away from artistic content that we would rather not confront directly.

Many people feel that though theory may try to redefine the audacious features of art—be it sexual explicitness, carnage, or religious desecration—it cannot completely banish them. The danger remains in spite of the theory, and it comes in many forms. Some identify the danger as obscenity; others, as offense. Art can cause us to feel deeply, to cry or laugh. Sometimes, if the

artwork instills in us the right beliefs and desires, art can move us to action—even to virtue or vice.



Fig. 1: Dread Scott, *What is the Proper Way to Display a US Flag?*, 1989.

Our response to these “dangerous” artworks is as variable as the works themselves. For an example, consider Dread Scott's *What is the Proper Way to Display a US Flag?* (fig.1). Steven C. Dubin catalogs the explosive responses to this artwork. The exhibit incited daily demonstrations by Vietnam veterans; a state senator on separate occasions folded the flag, then attempted to mail it, and finally, mounted it on a pole; white supremacist groups (taking issue with both the “unpatriotic” display of the flag on the ground and Scott's skin color) chanted “the

flag and the artist: hang them both high” (Dubin 1992, 113). It even spawned a reactionary artwork by Gary Mann called, *How to Display “Dread” Scott Tyler*, that included a flag on the wall, a book for comments, and a tape outline of a homicide victim on the floor.

Considering the reactions inspired by provocative works such as this, we might think that art is, contra Danto's claims, often harmful. But on what is our intuition founded? An answer to this question is complicated. Many people believe that immoral art can make us worse people, even when the artwork in question does not immediately motivate us to act in immoral ways. In this thesis I will argue that, although many people share the intuition that certain immoral art can *morally harm* us, immoral art does not have this capacity.

### 1.1 The History of the Moral Harm Thesis

The idea of moral harm can be traced back to Plato, who notoriously wrote in *The Republic* not only that artists are superfluous in the ideal state, but also that certain art forms are harmful to the soul. Plato has Socrates argue:

We [must] look for craftsmen who are naturally capable of pursuing what is fine and graceful in their work, so that our young people will live in a healthy place and be benefited on all sides as the influence exerted by those fine works affects their eyes and ears...and imperceptibly guides them from earliest childhood into being similar to, friendly toward, and concordant with the beauty of reason. (Plato 1997, 401c4-d2)

More succinctly put, Plato believes that art shapes our character. This idea has persisted, and has been used to justify censorship in our own era. In *Miller vs. California*, Chief Justice Warren Burger posed a similarly-inspired question on the consequences of obscene media. When considering why we should create laws that outlaw access to pornographic materials for minors, Burger asked:

If we accept the unprovable assumption that a complete education requires the reading of certain books, and the well nigh universal belief that good books, plays, and art lift the spirit, improve the mind, enrich the human personality, and develop character, can we

then say that a state legislature may not act on the corollary assumption that commerce in obscene books, or public exhibitions focused on obscene conduct, have a tendency to exert a corrupting and debasing impact leading to antisocial behavior? (Koppelman 2005, 1641)

Here Burger articulates the view that art is dangerous not only because it can lead to corrupt actions, but also because it can corrupt one's character.

The idea that art can affect one's moral character has been used to establish the role of literary critic and literature teacher as moral educator. In defending this role, Wayne C. Booth argues that “when we are fully taken in by a story-world and feel ourselves loving and admiring or hating and detesting portrayed characters, our own aspirations and habits of thought are changed” (2006, 231). Likewise, Booth asserts that “nearly everyone concedes, and all imply, that *no matter what we do about the moral powers of art, those powers are real*. Even the most ardent opponents of censorship do not deny artworks can harm some who 'take them in'” (ibid., 240, my emphasis). It is this understanding of harm—harm as a “moral power” or capacity of art—that I will challenge.

This kind of harm has typically been called *moral harm* and is distinct from other types of harm. Plato focuses on harm to one's soul or character, but contemporary theorists describe the harm more precisely, as directed toward one's *faculty for moral understanding or perception*.<sup>1</sup>

Andrew Koppelman defines moral harm along these lines:

If A were to induce B to make such a [moral] mistake, this would harm B even if B did not complain about it. It would impair B's ability to discern the morally better from the morally worse. This is a kind of moral harm. B has an interest in having good moral capacities whether he knows it or not. (Koppelman 2005, 1643)

Koppelman develops his definition out of Charles Taylor's work on the self in modern society.

<sup>1</sup>On this reading, having a *bad character* means that one is unable to grasp correctly which actions, aims or things are morally desirable. This is compatible with the idea that a person might not be correctly discerning which aims are rational, if rational aims are synonymous with good aims generally. I will take this faculty to be a fundamentally *cognitive one*, i.e., belief-based, an interpretation which aligns with a broad number of moral theories.

Specifically, he regards one's separation from human nature and rational aims as a kind of self-harm. Koppelman summarizes that if “a person could be mistaken about his fundamental purposes,” then this person is being harmed (ibid., 1643). That person will want or believe things that undermine his true ends.

This definition need not be so specific. Parents do not worry that their children are undermining their own rationality when viewing questionable artworks. Instead, they worry about the child's moral capacities generally. With this in mind we can give the following account of moral harm: *if A impairs B's faculty for moral understanding or perception such that B has diminished capacity to discern the morally better from the morally worse, then A has morally harmed B*. Moreover, B need not complain about this moral harm—there can be moral harm regardless of a victim's awareness of it. Here the harm done is one in which the harmed person's interest in having a sound (or well-functioning) faculty for moral understanding is “set back.” The most plausible way in which this faculty can become impaired will be that one's beliefs or desires are changed in such a way that the person has diminished ability to distinguish good from bad. I call advocates of this theory, such as Booth or Plato, *moral harm theorists*.

Using this definition, my thesis will focus on whether or not certain kinds of art, specifically *immoral* art, can *morally harm* us. This is a separate question from whether or not art can slander people or physically harm them. Consider that there are three ways that artworks can be immoral: (1) the means by which the artwork came about were immoral, (2) the artwork depicts immoral acts or has immoral features, or (3) the artwork is the cause of harmful consequences.<sup>2</sup> If the moral harm theorist wants to trace the harm back to the artwork, then she

---

<sup>2</sup>I borrow these distinctions, in part, from A. W. Eaton's essay “Where Ethics and Aesthetics Meet: Titian's *Rape of Europa*.” Eaton includes a fourth category in which the artwork is produced by an immoral person. I do not include this one as I will not address it, and I see it as being almost indistinguishable from category (1).

must focus on (2).

Option (1) is easily dismissed. If the artwork was created through immoral means, it is not clear how this could have an affect on the *viewer* of the artwork. Likewise, if one believes that the immoral means by which the artwork came about leave traces in the artwork (i.e., morally taint it), then we should focus on how those traces appear in the artwork. Consequently, we should morally evaluate the artwork according to option (2).

One might wonder why option (3) is not consistent with the moral harm thesis. There are a number of reasons that (3) cannot explain moral harm theorist's worries. Firstly, option (3) does not cohere with the idea that morally harmful works are *objectively harmful*. The moral harm theorist hopes to show that these artworks pose a threat to people generally, and that their power over viewers is consistent. For instance, if we have a series of artworks which are all similar, the moral harm theorist will be unable to say of any of them that they are immoral until some morally harmful consequence has occurred. Moreover, if such a consequence occurs, the moral harm theorist may only cite the individual artwork as being immoral. She will be unable to attribute immorality to the other artworks, no matter how similar they are. In this way, the goodness of an artwork may be claimed as lucky happenstance: it did not produce a moral harm, therefore it remains either amoral, or moral. Therefore, option (3) is likely to involve an objectionable degree of moral luck.

Secondly, this view of immorality may lead to unintuitive or contradictory consequences. Consider a hypothetical painting of a Parisian café. If this painting causes the viewer to seek out a lover who once spurned her, and attack him, the moral harm theorist would have to say that the painting is immoral for this. Moreover, if the painting causes someone to realize their love for another, then the moral harm theorist must endorse both claims, leading to a contradiction.

Finally, option (3) simply collapses the distinction between immorality and moral harm. Since it is plausible that we could have immoral art without supposing that it causes moral harm, the two ideas should be kept separate. For these reasons, if the moral harm theorist wants to argue that certain artworks have the capacity to morally harm us, then she will first have to establish that the artwork has certain immoral features that cause the harm; thus she will have to argue for (2).

## 1.2 The Structure of the Argument

In chapter 2, I propose a definition for immoral artworks. In chapter 3, I explain cases in which my definition for immoral art will not apply. I claim that while many people consider disgusting and offensive artworks to be immoral, disgust and offense are not moral judgments. Chapter 4 is devoted to moral harm. I present case studies of immoral artworks and argue that there are two ways in which they could morally harm the viewer: either they could (a) affect a person's beliefs through her imaginative engagement with the artwork, or they could (b) incite emotions which then lead to new beliefs. I argue that both possibilities are inconclusive. Specifically, I argue that in the case of (a) the imaginative experience involved in viewing art is not parallel to typical cases in which imaginings result in beliefs. In the case of (b) I argue that the ways in which emotions can affect our beliefs, likewise, will not be applicable in the case of immoral artworks.

The implications of this thesis—that immoral art cannot morally harm us—will be useful in three ways. First, it will be useful to aesthetics by providing an account of immoral artwork based on moral features of that artwork. Previous considerations of immoral artworks have limited their analyses primarily to how moral defects affect the overall *aesthetic* value of the artwork. My theory instead focuses on moral evaluation of artworks without qualification. As

such, it is applicable in a wider variety of cases than competing theories. Next, little has been written about the connection between disgust, offense and aesthetics. This thesis will add to the discussion and suggest that features of an artwork that elicit disgust or offense are not moral features. Finally, this thesis will provide a basis for ethical considerations in censorship cases. Many obscenity cases rely on the possibility that media can morally harm us. I will argue otherwise.

## 2 THE POSSIBILITY OF IMMORAL ARTWORKS

In this chapter I argue that an expanded *ethicist* approach is the most plausible theory for understanding the moral harm thesis. *Ethicists*, or proponents of the aesthetic theory “ethicism,” state that an immoral artwork is one that prescribes a positive attitude toward some immoral action or state of affairs. In what follows I present an approach to artworks which takes them to be expressions of attitudes. This understanding of artworks serves as the groundwork for the ethicist's thesis. Next I present an expanded version of ethicism that can account for implied (versus explicit) immoral content in artworks. Finally, I argue that although the ethicist commonly denies the “causal thesis”—that artworks cause their viewers to act or believe certain things—they need not deny it; this view is compatible with ethicism.

### 2.1 Moral Harm, Artworks, and Agency

Before presenting the ethicist's view, we should distinguish between two approaches one might take toward immorality as it appears in artworks: approaching an artwork with a moral schema versus judging the artwork *itself* as moral or immoral. Regarding the first approach, Noël Carroll notes that “[u]nderstanding a narrative also requires mobilizing the emotions that are appropriate to the story and its characters. One does not understand Trilby unless one finds Svengali repugnant” (1996, 228). We bring moral attitudes to the table when we study or enjoy works such as *King Lear* and *Madame Butterfly*.

This approach is not the one that the moral harm theorist endorses. In line with the second approach, the moral harm theorist is concerned with the individual's engagement with the work *as a whole*. Mary Devereaux makes such a distinction between one's judging works as a whole versus judging the individual features of an artwork: “Just as we can evaluate the act of breaking a confidence, independent of whether my perfidy is discovered and independent of its harmful

effects, so too we can evaluate the morality of a literary work *in and of itself*" (2004, 5, emphasis added). According to Devereaux, an artwork can be immoral because of qualities inherent in the work.

These moral qualities, taken together, are expressive of a certain viewpoint or moral attitude. To make sense of this expressive cohesion Devereaux and Bruce Vermazen argue that our judgment of such qualities typically involves positing some hypothetical agent to whom these moral beliefs or attitudes belong. For instance, Devereaux compares one's reading a novel to interacting with a kind of agent: "to see a text as organized in this way is to see it as purposive. *The text allows us to ask certain questions*" (ibid., 6, emphasis added). Because the work has been written in a certain way, with a certain perspective, when we engage with the work we thereby engage with a "posited author," who according to Devereaux is a persona that expresses the overall attitudes of the work (ibid., 6).

Bruce Vermazen argues for a similar point. He claims that when we speak of expression in artworks "we are attributing mental states, events and processes to an imagined utterer of the work. We are not, in making these attributions, saying that the imagined utterer expresses such and such thought or emotion, but that the work expresses it" (1986, 196). The work itself is a spontaneous and expressive utterance; the imagined utterer (who need not be the author) is the one to whom properties such as "sad" or "joyous" are attributed. Vermazen provides a definition for cases of expression: "An object expresses a mental property if and only if (subject to certain constraints) attributing that property to an utterer of the object would explain the object's having the features it has, and the property is not one of those presupposed in an attempt to interpret" (ibid., 207). I can only say of an artwork that it expresses "joy" if attributing this mental state to a hypothetical utterer (of the artwork) would explain why the artwork has the features it has.

Additionally, we can exclude properties that are needed to interpret the work, such as “written in English.”

If we understand artworks as *expressive evidence* for some hypothetical agent's mental state then a few consequences follow from this reading. The first consequence is that our engagement with artworks will be akin to an interaction with an agent. We will want to decipher the true intentions or attitudes of this hypothetical agent, to clarify the meaning behind their expression, and to do our best to interpret the expressions themselves correctly. According to Vermazen, the viewer “imagines that the object has been uttered by someone, [...] and then asks himself what mental economy would be behind such an utterance, what properties of an utterer would make it appropriate to utter just such an object as this” (ibid., 208-9). The work of art is an utterance, and as such, is open to imaginative investigation.

The second consequence is that any moral judgments we make must rest on this interaction with the work. We must not look only at aspects of the work, but instead we must consider the artwork's expressed attitude overall. In the same way that listening to a person's words and ignoring their body language may result in a misunderstanding of the expressive meaning of their words, our interest is not merely in the expressive elements of the work, but in what the work expresses as a whole.

This way of understanding artworks also supports the moral harm theorist's claims. Just as we might worry that individuals may be negatively influenced by engaging with their morally corrupt friends, the moral harm theorist worries that engagement with immoral art will affect our judgments. In the next section I give an account of ethicism, a view which similarly claims that artworks express attitudes on their own subject matter. I argue that the ethicist's definition of immoral artworks is the one which the moral harm theorist should endorse.

## 2.2 An Overview of Ethicism

Proponents of ethicism, such as Berys Gaut and A. W. Eaton, focus primarily on the question of how moral value affects the aesthetic value of a work. For our purposes, the important part of the ethicist's thesis is her discussion of expressive “attitudes” in artworks, and their connection to the immorality of the artwork. On this topic, Berys Gaut believes that artworks can endorse moral attitudes or claims without “explicitly stating an opinion” (1998, 183). They can do this in a variety of ways including symbolically, metaphorically, or verbally.<sup>3</sup>

When an artwork “manifests” a certain attitude, it does so by displaying a “pro or con attitudes towards some state of affairs or things” (ibid., 183). In displaying this attitude the artwork is also “prescribing” a response (ibid., 192). Eaton asserts that *prescription* “refers to the various rhetorical devices through which works aim to shape the audience’s sentiments regarding the characters and events they depict” (2003, 165-6). Therefore, artworks which manifest attitudes take a positive or negative stance on a particular state of affairs or thing, and prescribe that the viewer do so as well. A. W. Eaton provides an example of manifestation: an artwork such as Titian's *Rape of Europa* “calls upon viewers to adopt sentiments toward rape that play a role in the reproduction of gender inequality” (ibid., 167). Thus, on the ethicist's view, what makes an artwork *immoral* is that the work prescribes a response or attitude that is “unmerited” (referring to the ethical permissibility of the response) (Gaut 1998, 195). Therefore, if Leni Riefenstahl's *Triumph of the Will* prescribes that we respond to Nazis, Fascism, and Hitler with a pro attitude, it is reprehensible in this prescription, and we can properly call the artwork immoral.

---

<sup>3</sup>Gaut does not explicitly state all features that contribute to a work's attitude; however he does argue that the ethicist thesis is applicable to a wide array of artworks. Gaut explains, “The claim that works prescribe certain responses to the events described is widely applicable. [Ethicism can be applied to] to paintings, films, and other representational arts. Music without a text is also subject to ethical criticism if we can properly ascribe to the music a presented situation and a prescribed response to it” (1998, 193).

On the ethicist's view, the immorality of a painting such as Titian's *The Rape of Europa*, lies in its attempt to “produce strongly positive feelings toward a despicable act that currently plays an important role in perpetuating social injustice” (ibid., 167). The art then tempts us, in a manner of speaking: it aims to produce in us sentiments that are “wrong on ethical grounds for us to enter into” (ibid., 167). Because the artwork aims to produce *in the viewer* some response, there is a interpretative difficulty in determining that actual prescribed response. Gaut notes that “sometimes a work may explicitly claim to condemn something, while on a closer reading it actually expresses a positive attitude toward that thing” (ibid., 184). Because of this difficulty, Gaut introduces higher-order prescriptive attitudes as a way of drawing a distinction between initial prescriptions and a complete set of prescriptions an artwork gives. Likewise, while we might believe that an attitude of humor has been prescribed toward some demeaning scenario, the work might actually prescribe shame at our own response of humor.

Before we accept the ethicist's definition, we must consider one objection: what about artworks that do not *explicitly* show actions that are immoral, but rely on immoral beliefs in their viewers? Consider a paradigmatic case of an immoral artwork: *Triumph of the Will*. There is little dispute that this film expresses admiration for Hitler and the Third Reich. However, there are no immoral acts, *per se*, that are shown in the film. We can sidestep this worry by expanding the proposed definition to include not just acts, but sets of beliefs. There is a sub-text to the film: the glorification a regime and its particular ideology through the overt glorification of its symbols. Thus the work's expressed attitude is not toward what is explicitly shown, but rather, it is toward set of beliefs that motivate the shown state of affairs, e.g, the belief that Hitler is a great leader because he wants to purify Germany. This, I argue, is what occurs when one watches *Triumph of the Will*. A hypothetical set of background beliefs are a necessary part of engaging

fully with the film.

For this reason, it may be that works are immoral also when they reinforce certain oppressive stereotypes, or at least, elements of those stereotypes. The film is immoral for expressing an attitude toward these sets of beliefs. As a consequence, when the right sorts of people (Nazis) watch the film and utilize their background beliefs about the images in the film, the film reinforces those beliefs. We can find the concerns for background beliefs and sub-text in certain racist texts and illustrations like *Little Black Sambo* (fig. 2). The sub-text of the story is something like “black children get themselves into trouble because they are clumsy, rash or stupid.” This sub-text is not, in itself, immoral—even if the sub-text were true, there is nothing morally wrong with being clumsy. However because the artwork and its sub-text is part of a larger stereotype that is used to oppress others, the artwork becomes immoral. Thus, the reproduction of the image of Sambo and its use in other contexts is an instance of immoral art.

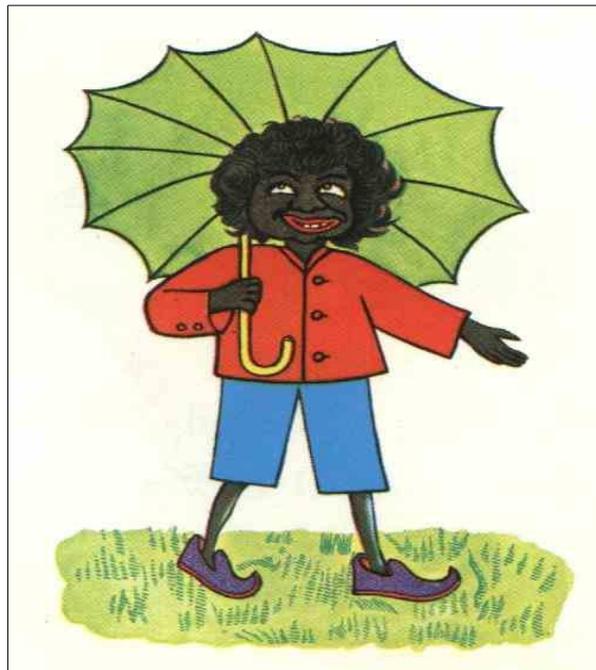


Fig.2: Cover illustration for *The Story of Little Black Sambo*, 1900, 1st ed.

Taking this expansion into account I propose the following definition of immoral art: *an*

*artwork is immoral when it expresses a positive attitude toward some immoral act or set of immoral background beliefs.* Here the representation of the immoral act or beliefs can be implicit or explicit in the content of the artwork. Determining the actual represented content will require some interpretative skill on the part of the viewer, but such a consequence is not an insurmountable problem.

There still remains one concern about our definition's relation to the moral harm thesis. Eaton, like Gaut, asserts that her version of ethicism does not support the causal claim that “the painting does in fact engender...attitudes in its audience” (ibid., 167). Because Eaton and Gaut deny the causal thesis, we might worry that any definition based on ethicism should lead one to conclude that these artworks are causally inert; thus moral harm is impossible. Against this concern, in the following section I propose that artworks have a disposition to produce in us their prescribed attitudes.

### **2.3 The Ethicist Thesis and the Causal Claim**

Eaton and Gaut argue that the ethicist thesis is independent of the causal claim that artworks have any effect on their viewer's attitudes or beliefs. If we take this to be true, we might then dismiss the possibility of moral harm outright: artworks may be immoral, but they are inert. This I argue need not be the case. If ethicism is taken as a heuristic to assessing the goodness or badness of an artwork, it is compatible with causal claim. For conscientious and engaged viewers imaginative interaction with an artwork will lead such viewers to entertain the attitudes expressed by the works.

In other words, while it is possible that the disinterested viewer can adopt the ethicist's theory, the typical viewer is not disinterested. If ethicism is not just a theory, but a practical apparatus for assessing the moral standing of certain artworks, then we should amend it to

acknowledge this context. Subsequently, we can understand artworks as having a *disposition*—one that is the result of the attitude it prescribes toward its subject—to produce correlate imaginings or attitudes in their viewers. By disposition I mean the tendency of the artwork to produce certain kinds of effects in certain conditions. In this way, artworks have a certain kind of causal power. Therefore, an artwork can be, given the right context, the sort of thing that presents attitudes and causes viewers to adopt such prescribed attitudes (at least temporarily).

That artworks have such a disposition does not necessarily lead to the conclusion that moral harm exists. It is one thing to say that artworks given certain circumstances will influence the viewer's beliefs, and another thing to say that they do this consistently and often. In other words, the moral harm theorist does not only believe that artworks have a disposition to influence their viewers, but that this a *strong* disposition: the moral theorist believes that artworks' disposition to affect the viewer's beliefs requires few background conditions. For instance, the moral harm theorist may argue that the only conditions that need to hold for a person to be morally harmed by an artwork is that the viewer simply think about the artwork, and imaginatively engage with it. Because immoral artworks have a disposition to cause us to imagine their content, the moral harm theorist will argue, immoral artworks also have a disposition to morally harm us.

I grant that artworks can causally affect the viewer's imagination and emotions, but I argue that this causal disposition is not a *strong* one, in the sense the moral harm theorist wishes it to be. In other words, I deny that a disposition of an artwork to cause the viewer to imaginatively engage with it, implies a disposition of the artwork to morally harm. The instances in which art artwork does affect changes in one's moral beliefs will be extremely few in number and rely heavily on idiosyncratic contextual factors.(I will present arguments for this position in

Chapter 4). Because this is the case, their disposition to cause such belief-adoption is weak at best; therefore, to generalize from these particular cases and argue that any particular artwork is morally harmful, is unjustified.

To summarize, in this chapter I argued for an expanded ethicist definition of immoral art. This theory can account for key claims that moral harm theorists make in regard to art. Second, I argued that we can understand immoral artworks as having a disposition to influence a viewer's imaginings and emotions, without thereby granting that the artwork has a disposition to morally harm (i.e., a disposition to cause the viewer to adopt various immoral beliefs). In the next chapter I will consider the cases that do not satisfy this definition.

### 3 DISTINGUISHING DISGUST AND OFFENSE FROM MORALITY

Given our definition we might be inclined to think that the category of immoral artworks is broad. Consider two artworks that are commonly cited as immoral and consequently, as causing moral harm: Robert Mapplethorpe's *Helmut* which is part of his *X Portfolio*, a series of 10 photographs referencing sadomasochistic and homosexual acts (fig. 3); and Andres Serrano's *Piss Christ* (fig. 4). Both of these works elicit responses of disgust and offense. For example, according to the art critic Lisa Jardine, Mapplethorpe's *X Portfolio*, is one that "simultaneously attracts and disgusts" (1996, 70). The most obvious element of these photographs that might trigger this reaction is their suggestion of homosexual and sadomasochistic acts.



Fig. 3: Robert Mapplethorpe, *Helmut*. 1979. Copyright the Estate of Robert Mapplethorpe.

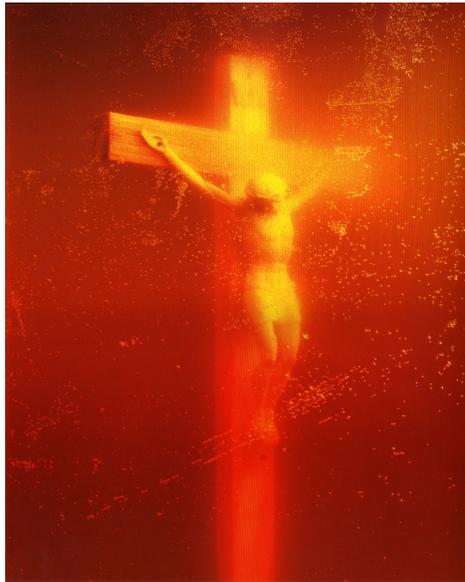


Fig. 4: Andres Serrano, *Immersion (Piss Christ)*. 1987. Copyright Andres Serrano.

Likewise, Andres Serrano's *Piss Christ* does not disgust and offend its viewers because it shows the image of a crucifix *submerged*, but rather, because the crucifix is submerged in *urine*. It is important to note that it is not only the desecration of the crucifix that is at work here, but also the sight of urine touching an object that is considered "pure," thus contaminating the symbol. *Piss Christ* would not necessarily be disgusting if Serrano had depicted the crucifix as burning. Therefore, the urine itself is the important trigger of the disgust and offense reactions.

I argue that our definition will not include these artworks and others like them. Specifically, our definition will not apply to artworks like these because the critic wrongly identifies objects which elicit disgust and offense as immoral. Therefore, I argue that these reactions are not good evidence for the immorality of an artwork, and disgust and offense are conceptually distinct from immorality. For this reason, unless the critic can cite additional reasons why artworks such as these are immoral beyond the reactions the artworks elicit, our definition will not encompass them. In the next two sections I establish the conceptual distinction between disgust and offense reactions, and moral judgments.

### 3.1 Disgust and Morality

Danto writes in, “Beauty and the Beastly,” that disgust is like beauty “in the mind of the beholder, but it is one of the mechanisms of acculturation, and there is remarkably little variation in our schedules of what disgusts” (2001, 25). Drawing on this shared response, some argue that disgust can alert us to moral violations. Against this claim, I argue that having a disgust reaction to an artwork is not sufficient for moral condemnation of that artwork; insofar as moral harm stems from immoral artworks, these works can not be said to produce moral harm. I begin my discussion with an evolutionary account of disgust, proceed to case examples, and end with counter-arguments against those who claim disgust properly grounds moral claims.

Charles Darwin first wrote about disgust as an evolutionary adaptation. He describes disgust as “a sensation” that “refers to something revolting, primarily in relation to the sense of taste, as actually perceived or vividly imagined; and secondarily to anything which causes a similar feeling, through the sense of smell, touch, and even of eyesight” (Darwin 2006, 265). Contemporary psychological findings suggest that disgust reactions occur toward a much broader range of stimuli than Darwin suspected. For example, people are disgusted by objects and other people that have come into contact with contaminants. Daniel Kelly argues that one feature of disgust is that it tends to cause cognitive biases. Kelly labels these biases as *core disgust*. Core disgust evolved to respond to disease and parasites (ibid., 45). One cognitive bias that makes up core disgust is *contamination sensitivity*. Contamination sensitivity refers to the bias that when we mark something as disgusting, we perceive it as being able to infect or contaminate other objects that were once pure.

To explain feeling of disgust at moral transgressions, Kelly proposes that the cognitive aspect of the disgust system, or core disgust, was “co-opted” for this purpose. It was “recruited to

also play a number of roles in regulating social interactions. More specifically, it became systematically involved in the cognition of social norms and group boundary markers” (ibid., 144). Because the regulation of social interactions had no definite social signal, the disgust system with its clear signaling was recruited. Metaphorically speaking, out-groups and deviant behavior were treated as a kind of pathogen and thus elicited disgust.

With this understanding of disgust, and the kinds of potential contaminants or pathogens it is supposed to target, one can consider the various universal elicitors for disgust. These include: “feces, vomit, blood, urine, and sexual fluids”; “corpses and signs of organic decay”; “bodily orifices”; “artificial orifices or breaches of physical bodies such as cuts, gashes, lesions, or open sores”; “items and substances once within those boundaries, which were once inside or part of the body, but then exit or are detached from the body” (e.g., vomit); and finally, “marks of disease and parasitic infection” (Kelly, 2011, 28-29).<sup>4</sup> This list, and the evolutionary theory that supports it, will be helpful in understanding why certain artworks can elicit disgust.

Having presented the origins of disgust, we can now turn to the question of its normative standing. There are various ways one can argue that disgust and morality intersect. One way of answering this question is to prove that people often feel disgust at moral transgressions. The other way of answering this question is to argue that people’s feelings of disgust provide reliable evidence for the immorality of the act in question. The former is a descriptive claim about people's self-reported reactions; the latter a normative claim about the relation of those reactions to morality. Both claims, I argue, are problematic; therefore, disgust is conceptually distinct from morality.

There is a good deal of evidence supporting the first claim—that people report feelings of

---

<sup>4</sup>Kelly compiles this list from a number of sources; in particular, Rozin, et al., 1992, 1995, and 2008. Additionally, he cites Curtis et al. 2004, and Haidt et al. 1994. See bibliography for full citations.

disgust toward moral transgressions. For instance, In a meta-study, done by Hanah A. Chapman and Adam K. Anderson, the authors found that many studies suggested that participants experienced disgust at “pure moral transgressions,” i.e., moral transgressions that “[do] not reference physical disgust” (2013, 303). They conclude that “[f]ive out of seven studies of verbal self-report are consistent with the idea that ‘pure’ moral transgressions can evoke disgust” (Chapman and Anderson 2013, 309).<sup>5</sup> Some psychologists have gone further claiming that there are three “realms” to moral disgust, two of which are relevant to the discussion at hand: *autonomy violations*, in which “an individual harms another or infringes on the other’s rights (e.g., a husband who abuses his wife)” and *divinity violations* in which “an individual disrespects the sacredness of god or behaves in an impure manner (e.g., eating rotten meat)” (ibid., 306). Each realm corresponds to a different role a person can fill: autonomy violations correspond to the role of person as “an individual preference structure”; while divinity violations correspond to the role of person “as a divine creature” (ibid., 575). Divinity violations include things like “eating rotten meat,” due to the the association of food preparation with religious purity. Note that the autonomy violations do not reference physical disgust, but divinity violations do.

In addition to this descriptive claim, critics, such as Leon Kass, have claimed that disgust is a source of reliable evidence for immoral acts. Kelly calls people who believe moral disgust is legitimate evidence for a moral transgression “disgust advocates” (DAs henceforth); their position he calls, “Deep Wisdom Theory” (DWT henceforth) (Kelly 2011, 138). Leon Kass has coined the phrase “the wisdom of repugnance,” where repugnance is interchangeable with

---

<sup>5</sup>It is worth noting that the five studies that Chapman and Anderson refer to are sufficiently unclear as to arouse suspicion about whether truly “pure” moral transgressions were presented to participants. For instance, it is unclear how photographs showing “moral transgressions (e.g., ethnic cleansing, child abuse),” could do so without referencing any universal elicitors of disgust (207).

disgust.<sup>6</sup> Kass claims that the disgust people feel toward cloning is “the emotional expression of deep wisdom, beyond reason’s power fully to articulate it” (Kass 1997, 20). He says that disgust “warn[s] us not to transgress what is unspeakably profound,” and that it “defend[s] the central core of our humanity” (Kass 1997, 20). In other words, repugnance, i.e., disgust, serves as evidence of a moral transgression.

We can understand the DWT argument as follows: Historically, disgust is a good indicator of things that can harm us or indirectly lead to harm. There is some correlation between things that are harmful and things that are immoral. Therefore, one can often trust disgust to indicate immorality, even if he or she cannot articulate the basis of his or her judgments. That is, even though we cannot explain the connection between disgust and morality, we can trust that it gives us moral insight.

Both the descriptive and normative claims are epistemically problematic. I will deal with each in turn. Even if the psychological studies are sound, we have reason to be skeptical about the presumed distinction between *moral disgust*, and *physical or pathogen disgust*. I argue that *clear* cases of immoral actions fall within the scope of “autonomy violations”: murder, abuse, theft and so on. Unclear cases are those in the “divinity violations” category. If there is, at the very least, agreement about whether autonomy violations are immoral, it is hard to imagine how actions such as these could be fully represented in artworks *and* elicit disgust, without reference to physical disgust. The critic could cite a few examples that do depict autonomy violations, but only by relying heavily on imaginative interaction between viewer and artwork. For instance, Nan Goldin’s photograph *Nan One Month after Being Battered, 1984* (Fig. 5).

---

<sup>6</sup>Kass identifies repugnance with a long list of features including: “Offensive... Grotesque... Revolting... Repugnant... Repulsive” (Kass 1997, 19). Additionally, Nussbaum has taken Kass as referring to disgust when he speaks of repugnance (Nussbaum 2004, 79).



Fig. 5: Nan Goldin, *Nan One Month After Being Battered*, 1984. 1984. Copyright Nan Goldin.

Here, the violation is one of autonomy—we imagine the situation that led to this abuse and to bear witness to its effects. Nevertheless, there are still remnants of physical disgust here—the blood in her eye, the bruises now turning yellow (suggestive of illness), and scratches. I argue that one cannot successfully elicit disgust in artworks depicting pure autonomy violations. One will inevitably still reference the traces of these violations—blood, gore, death—all of which are universal elicitors of disgust. Likewise, a photograph of, say, a theft, will not produce any disgust reactions in its viewers.

We can also describe this worry as an epistemic one: we cannot know if one's report of disgust is contingent upon amoral features of an artwork (e.g., the presence of blood, sexual content, or references to death and decay) as opposed to the moral content of the artworks. Likewise, people are poor judges in distinguishing between negatively valenced emotions. When they report disgust at such autonomy violations, they may instead be feeling anger. Rosa and Mir write that “there is evidence suggesting that disgust and anger are elicited by different cues of

moral situations: whereas anger is associated with the perception of harm and intentionality, disgust is typically triggered by bodily norm violations” (2013, 224).<sup>7</sup> Not only do we face difficulties in identifying immorality as a primary trigger of disgust, but also in identifying disgust itself.

There are also problems with the DA's normative claim, i.e., DWT. Firstly, unless the DA can explain why disgust sometimes falters and why it sometimes issues correct verdicts, she will be advocating disgust as evidence of morality ad hoc. In other words, the reliability of disgust is questionable. Disgust may occasionally correlate with immoral properties (and thus immoral actions), but unless one can know the underlying mechanism of disgust that tracks immoral features, the feeling of disgust is not a reliable guide to this judgment. For instance, one might have an intuition regarding whether or not it will rain. If one cannot explain how this intuition is connected with the probability that it will rain, then this intuition should not count as *reliable evidence* for knowing that it will rain.

This leads us to the major problem with DWT. The DA incorrectly views her emotional response as evidence for immorality. The problem with this appeal to emotion is that if the act in question is immoral, it is not because the act elicits an emotional response. It is immoral because of certain features about it or for its consequences, not for the emotions it elicits. An appeal to sentimentalism does not help the DA here. Sentimentalism only describes and explains our moral concepts, it makes no claim to objective morality. To say something is immoral under sentimentalism just means that I judge it to be immoral based on my particular sentiments. The DA wants more. The DA wants to say something about the act itself, not about his judgments. He wants his feelings to count as evidence for immorality, *not to be in what the immorality*

---

<sup>7</sup>They cite Russell and Giner-Sorolla (2011). See bibliography for full citation.

*consists.*

For these reasons and the reasons given earlier, disgust reactions are insufficient evidence for the immorality of artworks. As Kelly puts it, disgust is “a fairly blunt instrument. The mere fact that something is disgusting is a far from fail-safe indicator that something is poisonous or infectious, let alone immoral” (Kelly 2011, 147). In the next section, I discuss offense and its connection to disgust. I argue that offense, like disgust, does not track moral properties. Consequently, artworks that are offensive will not always be artworks that satisfy the proposed definition for immoral artworks.

### **3.2 Offense and Morality**

In this section I will provide an overview of offense as it appears in legal literature. Offense as it appears in this literature is commonly seen as a parent-category of disgust; that is, what disgusts us will also frequently offend us. For this reason, my discussion will build upon the previous presentation of disgust. I will however present a new category of *belief-based offense* and discuss its potential relation to immorality.

Offense was first distinguished from harm by John Stuart Mill, who asserts “[t]here are many who consider as an injury to themselves any conduct which they have a distaste for, and resent it as an outrage to their feelings” (1978, 81-2). Legal theorists such as Joel Feinberg have expanded upon this understanding of offense as displeasure. Feinberg argues that the word “offense” typically covers a broad category of “universally disliked mental states” (1985, 1). Feinberg distinguishes a general understanding of offense from a normative one in which those people offended are ones to whom some wrong or rights-violation has been done (*ibid.*, 2).<sup>8</sup>

---

<sup>8</sup>In this distinction lies the connection to disgust. Disgust falls within the general category of offense. It is a displeasurable mental state. In spite of this, it is not always the case that people who suffer this displeasure can properly say they've been wronged in any way, or attribute this displeasure to another agent.

*Normative offense*, as Feinberg describes it, is our focus here.<sup>9</sup> Offense has the following features: it includes a “disliked mental state”; this mental state is attributed to the “wrongful conduct of another”; and we “resent” this conduct (ibid., 2).<sup>10</sup>

Feinberg gives six categories of offense: (1) “affronts to the senses”; (2) “disgust and revulsion”; (3) “shock to moral, religious, or patriotic sensibilities”; (4) “shame, embarrassment, and anxiety”; (5) “annoyance, boredom, and frustration”; and (6) “fear, resentment, humiliation, anger” (ibid., 10-13). Some of these categories will, according to the critic, serve as evidence for immorality. That is, the critic will argue that feeling offense toward some artwork or action (that falls within one of these categories) is an indicator that the eliciting object or act is immoral.

In order to pick out the categories in which, arguably, some wrong has been done, we need to discard categories that rely upon disgust. As I've suggested, it is only incidental that disgust reactions occur in cases of autonomy violations. We are not disgusted by the particular violation, but rather, by unnecessary features of that violation (gore, decay, etc.). One's feeling disgust at some action (or artwork) will never be sufficient to establish arguments for immorality. Any offense founded on disgust will, for the same reasons, not be a reliable indicator of immorality. Likewise, other cases of offense like category 6 will only partially target moral transgressions: just because something scares me, and thereby offends me, does not mean that it is immoral. A plausible case in which offense stemming from fear or resentment might be the result of a moral transgression, is one involving “empty threats.” This however, does not seem to be a problem that will affect viewers of artworks, since artworks do not normally threaten their

---

<sup>9</sup>Unless otherwise specified, from this point forward when I refer to offense I will mean *normative offense*.

<sup>10</sup>The critic might object that there is no secondary agent when considering artworks, therefore, Feinberg's conception of normative offense is misplaced. However, under the conception of interaction with artworks given in chapter 2, we can see the artwork as being, at the very least, an expressive act of some hypothetical agent. In this way, artworks like *Piss Christ* can be understood as an *offensive* expressive act, and our resentment can be directed at the hypothetical agent.

viewers in this way.<sup>11</sup> If they do, it will be through visual or verbal depictions that somehow target the person's standing in society, e.g., an artwork that calls lauds the oppression of homosexuals.

Therefore with minor reservations, I propose that if the critic wants to argue that cases of offense are evidence for some moral violation, he or she will have to cite ones from categories 3, 4 (in part), and 6 (in part). Consequently, cases of offense include violations of moral, religious, or political norms; general nudity (though not sexually explicit nudity); and fear, resentment, humiliation or anger resulting from “threats” or “insults” (ibid., 13). More generally, the critic can be taken to argue the following: *cases in which offense is elicited by the violation of certain social, political, moral, and religious norms may count as evidence for the immorality of that violation.*

Given this narrowed class of things and actions which elicit offense reactions, I propose that what unites them is that they are elicited by an insult to one's moral beliefs. In cases where some norm has been violated, we can still cast the norm violation in terms of a belief: the viewer is offended because they hold a moral belief that X should not be done (because it is immoral), and the artwork does X. For instance, in Dread Scott's installation (Fig. 1), people responded angrily because they felt that their deeply held beliefs had been disrespected; moreover, that these beliefs were held for normative reasons. Thus an insult to them was a kind of moral transgression. Likewise, they believed that certain aspects about the artwork were violating moral principles, such as the proper treatment of a country's flag. Consequently, I will focus my

<sup>11</sup>There are exceptions to this. For instance, Eric Orr's “Colt .45” and “Saturday Night Special.” These are interesting but not defeating examples for two reasons. First, it is debatable how threatening they actually are—when on display the guns were unloaded. Thus, they only had the illusion of being threatening (even if it was a strong illusion). Second, they are a small minority in the history of artworks. Most artworks do not directly confront the viewer in this way. Thanks to Dan Weiskopf for providing these examples.

discussion on the connection between offense and the beliefs it depends upon.

Peter Jones discusses three explanatory categories of offense: *sensory offense*, *convention-based offense*, and *belief-based offense* (2011, 81-3). Belief-based offense is defined as offense that “presupposes and stems from the holding of a belief,” and “the fact that an act is perceived as wrongful by the offendees is crucial to their finding it offensive” (ibid., 82). Jones argues that it is not the displeasurable mental state that concerns us, but rather what is behind it: the particular *acts* or *ideas* that offend. Jones notes that “[i]f ‘offensive’ retains its meaning of inducing a negative mental state, *reporting* that I find something offensive is not very different from reporting that I do not like it or that I find it objectionable” (ibid., 84). Nevertheless, this is not good evidence for whether or not a particular act or expression is immoral. If acts are immoral, it is not because they elicit the response of offense. For instance, we may find that certain things that are immoral do not offend us. When reading about genocide, people are not necessarily offended by the information (though they may be angered or saddened); but most of them would readily say they hold the belief that genocide is wrong.

If the critic objects that the offense is itself a type of insult, and therefore a harm, we can respond that having one's beliefs challenged does not necessarily amount to having those beliefs insulted. Most artworks will express only generalized attitudes: the person claiming that his beliefs have been insulted cannot legitimately claim that his particular beliefs are being insulted. For this reason, it will be difficult for any offended party to claim direct insult. Likewise, the mental displeasure that comes from experiencing offense is so slight, that it is implausible to categorize it as a real harm. For these reasons, we should not take offense as a necessary or sufficient condition for immoral judgments. Offense may turn out to be good evidence for other things like core beliefs, but it should not serve as evidence that the offensive act in question is

immoral.

In this chapter I have argued that, while there is an overwhelming amount of shocking art that incites moral outrage, reactions of moral condemnation are not appropriate in cases of offensive or disgusting subject matter. For this reason, our definition of an immoral artwork can be narrowed to exclude artworks that express a positive attitude toward merely disgusting or merely offensive actions. The onus will be upon the critic to cite further evidence that such artworks are immoral and fall under the given definition. It is worth nothing that even if one were able to argue that such artworks were immoral, this would not lead to the conclusion that they are morally harmful. In the next chapter I will argue that while immoral artworks may exist, moral harm does not.

## 4 AGAINST MORAL HARM

Now that I have separated disgust and offense from immorality, we can turn to cases of purely immoral art. Recall the definition given in Chapter 1: an artwork (given an audience with sufficient and proper moral knowledge) is *immoral* when it expresses a positive attitude toward some immoral act or set of immoral background beliefs. With this definition in mind, we can consider potential cases of immoral artwork and the possible complications these cases present.

### 4.1 Case Examples of Immoral Art

First, consider Vanessa Beecroft's performance art, which involves the use of live females and body painting. The process for one such performance, *VB46*, involves bleaching of the women's hair, body painting, a full-body waxing, and long hours of standing (fig.6). Her work is meant as an inversion of the painted nude. Dave Hickey describes it in the following way:

“Painting remains the touchstone of Beecroft’s work; one may usefully regard her ephemeral tableaux vivant as standing in the same relationship to painting that painting does to drawing. [...] They are not paintings of women, in other words, but they are painted women” (2000, 6).

How is this performance immoral? Julia Steinmetz, Heather Cassils, and Clover Leary identify a crucial difference between Beecroft's work and other feminist performance art: the causal process leading up to the final display is hidden. For instance, they note that “[n]one of the models was informed that [full body and pubic waxing] would be required until we were in the waxing room. We already had white fried hair and eyebrows, and we also knew that we would likely not receive any of the money if we did not go through with it. At this point no one walked away” (2006, 762). In addition, the heels that Beecroft describes as imbuing her models with power were actually physically debilitating. Beecroft claims, "I want women on heels

because that's powerful, that's not natural nudity or pureness."<sup>12</sup> Alternately, Steinmetz and Leary report that their "ankles and feet became swollen, blistered, and bruised as a result of continuous standing in the dramatically ill-fitting shoes. Due to the extreme swelling of our ankles, almost all of the elastic straps eventually had to be cut.... The straps were later altered in Photoshop to look unbroken" (ibid., 763).



Fig. 6: Vanessa Beecroft, *VB46*, 2001. Copyright Vanessa Beecroft.

Between these descriptions and the resulting images, there is a discrepancy. The pain of the models and the (supposed) imperfections of their bodies have been removed or concealed (e.g., body hair and skin-color variations). The images take a positive stance toward the female form as pure, white, thin, and unblemished. These images are, in effect, a tribute to male gaze. Women, not just their nude forms, become objects. The process by which they are objectified is hidden. While the final images are appealing and beautiful, they are false, and serve to (according to our definition of immoral artworks) enforce the idea that objectification is not only

<sup>12</sup>“Vanessa Beecroft,” <<http://www.designboom.com/portrait/beecroft.html>>. Accessed: March 29, 2014.

painless, but aesthetically pleasing. Steinmetz and Leary write:

Because Beecroft's performances conceal the literal and symbolic violence her models endure in their transformation into objects to be viewed, [and] *because members of the audience are not implicated in their desire to objectify and fetishize the models' bodies...*her work fails to produce the type of critique enacted in [classic works of feminist performance art]. (ibid., 767, emphasis added)

Here, the missing attitude, for the ethicist, is the condemnation of the audience's responses. We respond positively toward the images, and this positive response is not condemned. For this reason, any sexist attitudes are simply enforced. Beecroft's work under our definition is thus immoral.

Recall that there are three ways that an artwork might be immoral: it can be immoral because it was created through immoral means; immoral because it has immoral features (the definition of immorality I defend here); or immoral because it has immoral consequences. The objector might argue that I am conflating the immoral means by which Beecroft creates her work with the immoral features of the work.

One response is that Beecroft's work lies between performance art and photography. For this reason, the means by which the photography is created is a necessary part of the artwork. The performance is not just the final display; it must include the process behind that display. These women are not an installation by themselves—they are given performative instructions on how to behave. They are not naturally white-haired or waxed. All of these procedures are part of the performance. In Beecroft's work these procedures are consciously absent.

Another artwork that has been criticized by feminists is Allen Jones' sculpture series *Table, Chair and Hatstand* (fig. 7). Jones has claimed that he only wished for his pieces to be a commentary on “the female figure and the sexual charge that comes from it” (Gayford 2007). For our purposes however, Jones' intentions are irrelevant. The work itself expresses an attitude,

and this need not be the attitude of the author or artist. The mannequins are lifelike enough that dismissing the artwork's attitude as absurd or sardonic is implausible. In Jones' work, women are man's furniture, silent and subservient. The pieces express a positive attitude toward this objectification and, thus, are immoral.<sup>13</sup>



Fig. 7: Allen Jones, *Table, Chair and Hatstand*, 1960.

Finally, some immoral art will also take the form of propaganda. During the first Sino-Japanese war countless woodblock images were created to honor the Japanese' victories over the Chinese (fig. 8). Here the print takes a positive attitude toward Japan's imperialism. More importantly, it casts the Chinese as primitive and non-white, thus uncoordinated and lesser than the Japanese. As early as the 1850s, Japanese were introduced to minstrel shows and racist imagery. Atsushi Tajima and Michael Thornton write of a minstrel showed performed for a Japanese audience in 1854, that “the Japanese laughed at and not with the performers, and were

<sup>13</sup>It is interesting to note that on a different definition of morality, say a consequentialist one, Jones' work may end up being moral. It spawned intense debate as well as many feminist and critical artworks that illuminated the oppression and objectification of women. For instance, Jemima Stehli's pastiche of *Table I* employs her own body as the support for the glass tabletop.

thus partners with White Americans in derision directed at Africans. This initial partnership is symbolic of what Japan desired in a relationship with the West: superiority to others of color and equality to Whites” (2012, 350).



Fig. 8: Mizuno To, Woodblock print depicting the Battle of P'yongyang during the first Sino-Japanese War.

This racial line is apparent in the appearances of the soldiers. The Japanese soldiers have groomed mustaches and military uniforms. The Chinese have large lips, head-scarves, and tunics. Here the Japanese wish to be seen as equals to whites and cast the Chinese as counterparts to the stereotype of the racially inferior black man. John Dower notes that, invariably, these woodblocks depict the Chinese as “round faced, yellow-skinned figures—bearing considerable resemblance, in fact, to the stereotyped caricature of the Oriental that the Westerners loved to draw” (1986, 209). Because the woodblock print glorifies not only these racist attitudes and stereotypes, but also the battles being fought to conquer such people based on them, it is immoral.

With some examples of immoral artworks in hand, we can now turn to the possibility of

moral harm. One might wonder after acknowledging the immorality of these works whether it is wrong to engage with artworks that promote these immoral beliefs and acts. In the next section I argue that we cannot be morally harmed by such works. In arguing for this, I will suggest an error theory for moral harm: we attribute our beliefs about the connections between the artworks and the real world to the artworks themselves; we wrongly believe it is a property of those works that the attitudes they promote could have causal force. In the conclusion (Chapter 5) I will expand on this error theory, explaining that we can still condemn such artworks without falling to the suspect claim of moral harm.

#### **4.2 Challenging the Moral Harm Thesis**

Recall the definition of moral harm given in the introduction: some artwork A is morally harmful if A were to “impair B’s ability to discern the morally better from the morally worse” (Koppelman 1643). We expanded this to the following definition: artwork A is morally harmful *if A impairs B’s faculty for moral understanding or perception such that B has diminished capacity to discern the morally better from the morally worse*. In our case an immoral work—any artwork that expresses a positive attitude toward some immoral act or set of immoral background beliefs—morally harms the viewer when his or her capacity to discern the morally better from the worse is diminished.

How might the immoral artworks given in 4.1 do this? There are two possibilities for how immoral paintings can morally harm us. The first option is that through imagination, we adopt the immoral attitude, i.e., we endorse the attitude that the painting endorses. The second option is that we are emotionally driven to adopt the attitude the painting endorses, or we consider the emotion to be evidence for the immoral attitudes presented by the artwork (and subsequently adopting those attitudes). I argue that neither strategy is unproblematic. In the next two sections I

present background psychological and philosophical evidence for each option and give counterarguments against each.

**4.2.1 *Imagination's Influence on Belief.*** The first option for the moral harm theorist is to argue that imaginings can directly affect our beliefs. Specifically, the concern is whether or not these imaginings can influence a person's beliefs such that he will be less able to distinguish between good and bad when faced with such a dilemma. This concern can be restated in two ways. The moral harm theorist's concern hangs on the possibility that imaginative interaction with an immoral artwork will either (a) cause the viewer to adopt the immoral attitudes and beliefs, and as a result, the person will be unable to judge good from bad in the future; or (b) strengthen pre-existing immoral beliefs, thereby making the viewer a worse judge of good and bad.

Option (b) faces problems. While this claim is plausible, it is not one that the moral harm theorist can easily show to be true. The moral harm theorist seeks an objective criteria by which we can cite some artworks as the salient cause of a shift in moral beliefs. However because the viewer in question has diminished capacity to discern between good and bad, the viewer may see the artwork as endorsing her immoral beliefs, even if the artwork does not manifest such an attitude. We can state this concern in terms of capacities: in order to measure the capacity that the artwork has over its viewer to shape her beliefs, we need to control for interfering variables. I am not arguing that there is no way to know if a painting manifests some immoral attitude—this is an interpretative task. Instead, I argue that in order to establish that the manifested attitude has a causal capacity to change the viewer's attitudes, we need to isolate the artwork as the salient cause of the person's shift in beliefs.

This is an epistemic worry, but one that cannot be ignored. Consider an analogy. I know

it is the right thing to do to donate money to charity. Donating to charity also makes me happy. We might argue the following about determining the salient cause of my actions: The clearest case in which my moral principles (versus my pleasure in performing acts of charity) caused my action, are cases where I am *not* happy about donating money, but I donate anyway. The same can apply to the determination of moral harm in artworks. In order to establish that artworks can morally harm, we must isolate those situations in which the moral harm could not be caused by say, the viewer's poor character. It is possible that the immoral person, given their dysfunctional moral perception, will interpret *any artwork* as coherent with their immoral world-view. Therefore, the *clearest* cases of moral harm will involve people who do not already hold the immoral attitudes expressed by the artwork. It is possible that moral harm *could* take place in the manner given by (b). I simply argue that such a case is not epistemically accessible, and thus cannot establish the existence of moral harm.

Joshua Landy have given a similar argument regarding the difficulty of establishing corruption in cases where people already hold the beliefs the artwork endorses. Unlike myself, Landy sees this difficulty as sufficient to reject the moral harm theorist's thesis. Landy presents an argument from *underdetermination*, as I call it. He explains that we can extract many morals from literature. For instance, in Chaucer's *Nun's Priest's Tale*, we might find two morals: (1) dreams are prophetic and (2) prudence is a useful and valuable character trait. Landy argues that the only reason why we do not take the first moral to be what the story intends to teach us, is that we "already subscribe to a particular piety, [and so we] will happily take a story to be illustrating it" (2010, 71). Therefore, Landy says that literature is inert on its own. He does argue, though, that literature "can corrupt, ironically enough, when coupled with the belief that it is morally beneficial" (ibid., 73). This belief, he argues is traceable back to institutions. Without the

“sanction of an external authority” morals in literature will not be adopted by readers (Landy 2012, 35).<sup>14</sup>

Landy's point here supports my own. We both agree that our (moral) perception of artworks is dependent on a preexisting theory, i.e., we approach artworks with various interpretative and moral theories in mind. Viewers and readers often interpret the “moral” of the story as aligning with their own beliefs, and strengthen those beliefs due to this cognitive bias. Unlike Landy, however, I do not believe that Landy's argument from underdetermination—in which the proper moral of the story is underdetermined—refutes the moral harm thesis. Landy's argument focuses on the circularity of justification for beliefs about *individual* morals in a story. Based on this, it is wrong to conclude that artworks cannot morally harm us. Landy's argument presumes that readers *never* come to a story without opinions on its subject matter; moreover, that they will not have opinions which are in tension with those presented in the artwork. In other words, Landy does not consider our option (a)—that the artwork could cause us to adopt *new beliefs*—and thus, his argument falters. Sometimes stories lead to beliefs about people or places with which we have had no direct experience. A story could not give us these new beliefs *unless* it had some influence beyond mere corroboration. The moral harm theorist's claims are still pertinent.

As a result, possibility (a)—a person can be morally harmed by an immoral artwork if it causes the viewer to adopt immoral attitudes and beliefs—is the one that we should consider when trying to construct an argument against the moral harm theorist. At this point, we must

---

<sup>14</sup>There is some ambiguity in Landy's argument here. It is worth noting that an charge of circularity against the adoption of specific morals from a story is different from a charge of circularity about the adoption of *any* beliefs from a story. In other words, it is one thing to claim that we will only interpret a story in line with what we already believe, and to claim that we will never adopt beliefs unless “external authority” has sanctioned the work as a whole. Landy's argument from underdetermination supports the prior claim, but not the latter.

introduce a desideratum. Possibility (a) turns on whether imaginings can produce in us *strong* moral beliefs. If a belief is strong, then it will not be easily dismissed in light of contrary evidence. If it were the case that these beliefs were easily discarded, then they would be unlikely to be said to be (partly) constitutive of one's faculty for moral reasoning. Moreover, if these beliefs were easily defeated, it would be epistemically difficult to justifiably trace the cause of a particular belief back to an artwork

Consider the following scenario. Ryouichi acquires the belief when looking at the woodblock print that Chinese people are morally inferior. Later he becomes friends with a Chinese person and changes his mind. Sometime later, that Chinese friend betrays him (e.g., cheats him in a business deal). Ryouichi then decides that Chinese people are by nature deceitful, and thus not as morally good as Japanese people. It seems in this scenario that Ryouichi has acquired a belief with the same content as his earlier belief (Chinese people are morally inferior to Japanese people). The woodblock print can no longer be considered the cause of this belief. Thus, if the woodblock causes Ryouichi to adopt a strong belief, it will be hard to determine if the woodblock is the salient cause of Ryouichi's faulty moral reasoning. In this case, it seems that the bad business deal is the salient cause of Ryouichi's standing racist belief. If it is not, and Ryouichi sees the incident as confirming evidence that he was “right all along,” this fact does not illustrate that the print was sufficient to warrant the belief. The print's expressed attitude will still need real-world confirmation; namely, Ryouichi's betrayal.

How might the adoption of such strong beliefs occur? How do these beliefs relate to the imaginative content we get from an artwork? Neil Van Leeuwen has argued that when we speak of our ability to imagine something, there are three senses of the word “imagine.”<sup>15</sup> Two are

<sup>15</sup>The other way is *imagistic imagining*, wherein the word “imagine” can indicate an imagistic representational format for what's happening in one's mind” (2013, 221).

pertinent to our discussion: *constructive imagining* and *attitude imagining*. *Constructive imagining* occurs when a person is “engaged in a temporally-extended constructive process of assembling mental representations” (2013, 221). *Attitude imagining* is wherein someone “has a cognitive attitude besides belief to the content [before them]” (ibid.). In other words, we take an attitude of imagining toward some content X, versus an attitude of belief toward X (i.e., I *imagine* X versus I *believe* X). Importantly, Van Leeuwen notes that constructive imagining can sometimes result in an attitude of *belief* rather than an attitude of imagining (ibid.). For instance a detective may (constructively) imagine how a particular murder was committed while walking through a crime scene. The detective may actually believe the consequent mental representation, e.g., the butler killed the duke.

Unfortunately, this account does not properly map onto the case of artworks. An artwork might suggest further imaginative content. For instance, the viewer might be called to imagine events only implied in a painting. Regardless, this process is relevantly dissimilar from cases of constructive imagining that lead to beliefs. The dissimilarity lies in the target of the mental representations—the viewer will not have a belief about anything *beyond* the artwork. Richard Wollheim explains that we play a kind of game with the artwork, asking about what can be imagined: when one does not wish to (or simply cannot) imagine something, then

for [the] spectator here and now, the limits of visibility in this surface have been reached.[...] [W]e can understand the [viewer's choice not to imagine further] as a refusal on his part to be forced beyond the appropriate experience, hence a refusal to force upon the picture something that it does not represent. (1998, 224)

The painting—which is composed of marks on a surface—“leads” our imaginings (ibid.).

Therefore, we can view the artwork as giving the viewer *guidelines*—content for initial imaginings—and the viewer fills in the details; however, the viewer is always constrained by the

artwork.

Alternately, in typical constructive imagining, the target of the imaginings is some truth about this world. This is why the resulting mental attitude is a belief (versus an imagining). One interprets the imaginings as evidence for the truth of a proposition. In similar fashion, for constructive imagining to yield beliefs about the world from an *artwork*, the artwork's expressed attitudes would have to be interpreted as evidence for some belief about the world. I am not arguing that artworks *never* provide evidence for the views they express. Instead I am arguing that a belief that the artwork is evidence for its expressed views will rely on prior beliefs about, for instance, the power of artworks or their relation to morality.

One might insist that we construct beliefs out of the imaginings we have when we look at artwork. For instance, the belief that "Chinese are primitive." However this does not quite capture the nature of this belief. In fact, one's imaginings include an extra step. We do not without warrant move from the initial belief, "[the imagistic, represented] Chinese [here] are primitive," to the generalization that "[all actual or real] Chinese are primitive." Likewise, we will probably need the belief that the painting represents some actual event, and does so truthfully in all relevant aspects. These beliefs are ones that the viewer must bring to the artwork, ones without which the artwork cannot affect the viewer in any substantial way. The viewer will have to believe that the artwork is an authority on the evidence it presents; therefore, the charge of moral harm will not be a direct result of the artwork's power over the individual. In other words, if we deny that the features of the artwork, independent of whatever theory we bring to interpret the work, are ever a salient or important part of the explanation of how one forms a particular belief, we escape the worries of the moral harm theorist.

Finally, the moral harm theorist could appeal to certain contemporary work on

imagination which suggests that imaginative engagement is not necessarily confined to the “world” of imaginings. Tamar Szabó Gendler describes the phenomenon of contagion—when our imaginings can lead to changes in cognitive or affective attitudes. Gendler explains that this contagion “arises because certain features of our mental architecture are source indifferent, in the sense that they process internally generated and externally generated content in similar ways” (Gendler 2006, 183-84). As a result, many of the things we imagine can influence perception, emotions, and other attitudes *in the same way that perception can*. Likewise, Gendler argues for another instance of cognitive contagion at the unconscious level: imaginings or “mere contemplation of some emotionally neutral imaginary scenario” can prime us to be “(over-) sensitive to similar scenarios in the actual world” (2003, 131).

Neil Van Leeuwen presents psychological evidence to support Gendler's claim that certain imaginings are “source indifferent.” Based on this evidence he concludes that “[m]ental imagery exists” and “[m]ental imagery is processed largely in the same regions of the brain as perception” (2011, 73).<sup>16</sup> Because they are processed in the same regions of the brain, mental imagery, or “[p]erceptually formatted imaginings can be integrated with veridical perception, especially visual or auditory perception” (*ibid.*, 67). When we imagine an elephant in a (real) room, we *integrate* this mental imagery into our visual perception of the room.

If imaginings have a direct effect on actions and emotions in the same way that perception does, this does not entail that imaginings can result in the adoption of certain beliefs. In response to this objection, I argue first that we should distinguish between *pretense* and

---

<sup>16</sup>Among other studies he cites Kosslyn et al.'s study which indicates that in PET scans “visual area 17 is activated both in perception and in imaginative imagery. The same sorts of contents, whether perceived or imagined, yield similar activation patterns” (73). See reference page for full citation.

*imagining*. Cases of cognitive contagion that are relevant here apply to *pretense*.<sup>17</sup> Imagination, if we take it in the constructive sense, involves the manipulation of mental imagery; pretense, on the other hand, involves acting upon those imaginings (and therefore, some element of desire or motivation). In cases of pretense it is possible that contagion will occur, and imaginings will influence action. But this possibility should not lead to us to think that our beliefs have been affected. Indeed “beliefs *govern* imaginative inferences” (Van Leeuwen forthcoming, 9). In this way, beliefs are prior to imaginings.

How might contagion occur then? Van Leeuwen notes that “imaginings of the type [he posits] are closer to the generation of action insofar as *they are causally downstream from beliefs*” (2011, 67, emphasis added). Imaginings can influence our actions, and sometimes—in cases of contagion, cause us to act in a way contrary to our beliefs—but this does not imply that our beliefs are somehow different. Instead, this possibility implies that our beliefs have been bypassed in the performance of the action. Therefore, cognitive contagion does not allow for the conclusion that imaginings directly influence our beliefs.

In cases where the cognitive contagion takes the form of “priming” we can say something similar. Gendler gives the example of imagining that we are birdwatchers: “if I spend the morning imagining that I am birdwatching, I will be more likely [i.e., *primed*] to attend to actual bird-encounters in the afternoon,” and I will lower evidence for certain beliefs, e.g., “hearing a rustle in the tree outside my study window may be sufficient to convince me that the sound has been made by a bird” (2003, 135).

Gendler is correct to point out that when we are actively incorporating our imaginings

---

<sup>17</sup>There are cases of cognitive contagion involving visual representation. For instance, *imagining* a black square results in the visual experience of an afterimage, in the same way that perceiving a black square does. This case is interesting, but orthogonal to our discussion—participants do not really believe that there is a white or black square beyond their mental image of it. See Gendler (2006) for a summary of this study.

into our actions (i.e., during pretense) then our evidentiary standards may be lowered in the way she describes: our beliefs will be influenced by imaginings *as if* they were percepts. However, when we are engaged in imagining what is presented to us by some artwork, we are not thereby engaged in the process of integrating those imaginings into the real world. As a result, any priming that occurs will depend on background beliefs about the artwork's application to the real world. Finally, the concern that imaginings may sometimes lead to *affective* contagion can be handled by the general question of whether or not emotions influence our beliefs. This possibility will be considered in the next section.

**4.2.2 *Emotion's Influence on Belief.*** The second available option for the moral harm theorist is to claim that we are driven via emotions to adopt the attitude the artwork endorses, or we consider our emotional response to the artwork to be evidence for the immoral attitudes presented by the artwork (and therefore we adopt those attitudes).<sup>18</sup> In this option we regard the emotions we feel when engaging with the artwork as support for the attitude we imaginatively adopt. For instance, if a painting makes us cry, we will be more inclined to believe that the imagined event is sorrowful or wrong.

There is psychological evidence that supports this claim. Nico H. Frijda and Batja Mesquita argue that emotions are intimately connected with belief formation. Emotions do this in a variety of ways. Our emotional state often leads us to notice new things in our environment that would elicit that emotion. Likewise, Frijda and Mesquita note that “[m]any beliefs are created to fill out the picture of the emotional situation. Others are formed to justify or explain one's

<sup>18</sup>I will assume that artworks can elicit genuine emotions. There is an extensive amount written about the nature of fictional emotions. Generally speaking though, people are not skeptical of whether or not art can elicit these emotions, but rather are concerned with how this is possible given a cognitive (belief-based) theory of emotions. Even then, there is good reason to believe that it is not strange or unlikely that fictions elicit emotions. For more on psychological support for the possibility of fictional emotions see: Van Leeuwen, “The Meanings of 'Imagine' Part II: Attitude and Action,” (forthcoming), and Tamar Szabó Gendler and Karson Kovakovich, “Genuine Rational Fictional Emotions” (2005). See bibliography for full citations.

emotional appraisal” (2000, 53). In effect emotions can lead to new beliefs through (1) readying us to notice new features about our environment and (2) in rationalizing the particular emotion. This influence is not without limitations. Frijda and Mesquita note that these beliefs are by nature short-lived (ibid., 53). Given our previous desideratum about the strength of beliefs (in 4.3.1), this account is still lacking.

Frijda and Mesquita do argue that these beliefs can become stable over time if the emotion becomes a *sentiment* (ibid., 55). Sentiments are “dispositional emotions,” that lead us to appraise new evidence in a way that aligns with the particular emotion (ibid.). Likewise, these attitudes are necessarily “attitudes towards what touches upon one's [personal] concerns” (ibid., 56). Frijda and Mesquita see beliefs as tied up with emotions. They explain that beliefs associated with an emotional object “pertain to intrinsic properties of the emotional object, [and therefore] they are about things that persist” (ibid., 55). In other words, if the emotion becomes stable (i.e., a disposition), then its corresponding beliefs will also become stable.

Frijda and Mesquita list four ways that sentiments can form: (1) If the emotional event lends itself to “dispositional attribution,” where, for instance, “personality dispositions are [seen as] permanent attributes”; (2) if there is “pertinence of the belief to an ongoing situation, and its usefulness for dealing with it”; (3) if there is “the presence of social encouragement or example for the belief”; and (4) if the belief is mediated by “emotional anticipation” (ibid., 56, 57).

*Emotional anticipation* here refers to the feelings we commonly have when considering alternative scenarios. In any given context we might form “solid and persisting beliefs” if the

alternative beliefs have perceived emotional, social, or physiological negative consequences.<sup>19</sup>

I argue that in light of the moral harm theorists' concerns about *artworks*, these four conditions are not sufficient to cause the formation of strong beliefs. I address each condition in order. Condition (1) says that we might see the particular artwork as having certain properties that elicit the relevant emotion. We have, for example, a novel belief like, "this artwork is disgusting," rather than (or in addition to) the belief, "this artwork elicits disgust." This may well be true, but it does not lead us to the conclusion that artworks can morally harm us. The viewer in attributing such a disposition to the artwork is unjustified—these properties are merely propositional attitudes of the viewer.

Condition (2) suggests that our emotions might become sentiments, if their corresponding beliefs have particular relevance to an ongoing situation. This condition encounters problems as well. When viewing an artwork we are not in situation that is prolonged or sufficiently emotionally charged to be considered an "ongoing situation." Consider the examples that Frijda and Mesquita use for (2). They cite such situations as a "deteriorating marital relationship, political oppression, and an ongoing fight" (*ibid.*, 56). Beliefs that may be perceived as useful in these highly emotional situations are adopted in order to "survive" the situation. Our interactions with artwork so brief that seeing them as the kind of "situations" in question is misguided.

Likewise, the artwork has no causal power to place us in those situations. Even if one walked away from the woodblock believing that Chinese are inferior to Japanese, the painting does not

---

<sup>19</sup>Specifically Frijda and Mesquita say that "Emotion anticipations are the anticipations, foresights, and imaginings of actual emotions that might emerge under certain envisaged circumstances" (58). They give examples that align with my presented interpretation, including our desire to hold onto a belief because we perceive that abandoning it would conflict with our sense of self, or social norms. When considering whether or not to adopt a belief or keep one's belief we might experience "[f]lashes of feeling, images of future emotional situations of derision and of the wailing and gnashing of teeth cross one's mind at moments of decision. We just made mention of thoughts like 'I could no longer have faced myself' that make decisions flip the way of accepting a grave risk [e.g., hiding Jews during WWII]" (58).

cause the person to enter situations in which this belief would be expressed. The person would need to seek these situations out on his own.

One might argue that during Japan's imperialist period, the relevance of such emotions and beliefs prompted by the woodblock print could turn a fleeting emotion into a sentiment: such emotions are relevant to the Japanese citizen's life; therefore, they become persistent and the corresponding beliefs become stable. Granting this situational stipulation makes the concept of moral harm so narrow as to remove all its content: given the right contexts, the right social pressures, and a receptive audience, art can make us worse people. However, given the right contexts any particular object can be a weapon. This fact does not justify the belief that these objects are then always weapons. Likewise, an appeal to context will not explain people's concerns that certain films like *Triumph of the Will* can morally harm us. The moral harm theorist believes that the power to harm is a timeless and an *intrinsic* property of the work.

Condition (3) states that we might develop sentiment if there is social encouragement to adopt the belief. It is obvious why we should reject this in our case. Any moral harm that might result from adopting the particular attitudes in the artwork can be traced to immoral background beliefs, an appeal to authority, or the belief that moral norms are determined by one's culture. If the person adopts the moral attitudes of the artwork because society endorses those attitudes, then it is not the artwork that morally harms us; rather, it is the background beliefs or social contexts that cause us adopt those moral attitudes that are immoral.

Finally, condition (4) states that the emotions elicited by the artwork might point to counterfactual consequences of adopting (or rejecting) their corresponding beliefs. To make this condition more clear, consider an example. One might adopt the anti-feminist attitude presented by Allen Jones' work because one *feels fearful* that doing *otherwise* would cause the social

structure to unravel. One might think (or unconsciously imagine) that if women were *not* fundamentally meant to serve others that women might stop doing this altogether—chaos would ensue. Here it seems clear that the adoption of the attitude is not triggered by the artwork itself, but by some counterfactual state of affairs. The adoption of the immoral belief (and thus the potential for moral harm) hinges upon the viewer's ability to counterfactually imagine a situation that is contrary to the one is presented in the artwork. The artwork's content then cannot be cited as the cause of moral harm. If anything, the critic would be forced to conclude that any moral harm was the result of the person's background beliefs, suspicions, or prejudices: if the viewer did not hold such emotional and cognitive biases, then they would not adopt the attitude manifested by the artwork.

Having rejected all these possibilities, we can subsequently reject the possibility of moral harm. Indeed, people's convictions about moral harm are strong. Parents still are concerned for the moral development of their children and seek this out through literature and the arts. But insofar as we believe artworks to have total causal sway over our moral faculties, I can only offer an error theory. Moral harm theorists wish to point out a property of immoral artworks, namely, that they have intrinsic immoral features that have causal force over our beliefs. Artworks have no such properties. Exposure to immoral artworks alone is not sufficient to affect our attitudes or beliefs permanently. We must willing viewers—ready to adopt such beliefs because of other beliefs we hold. Thus when we say that some artwork can morally harm us, this does not point to some power the artwork holds. Instead, it points to the human tendency to attribute our own woes, fears, and beliefs to others—whether they be artworks or people.

## 5 CONCLUSION

In this thesis I have argued that moral harm claims are in error. They attempt to argue for a strong disposition of moral artworks: they have causal sway over the viewer's moral beliefs such that the viewer will, with few background conditions adopt strong beliefs that echo the ones expressed in the artwork. I argue that if the viewer ever does adopt beliefs due to engagement with an artwork, these cases will be rare, and the conditions that need to hold will be numerous. Specifically, the viewer must be in a very particular context and have specific backgrounds beliefs if she is to adopt the beliefs the artwork expresses. Thus, we are in error to think that immoral artworks, generally, have the strong causal power of their viewers that is necessary for moral harm claims.

I have argued for this position by first presenting a definition for moral harm where moral harm is any case in which some artwork A impairs person B's faculty for moral understanding or perception such that B has diminished capacity to discern the morally better from the morally worse. I then argued for an expanded ethicist conception of immoral artworks. I suggest that immorality in artworks relies not only on attitudes expressed in the artwork, per Gaut and Eaton's claims, but also on the artwork's disposition to cause the viewer to adopt the attitudes presented in the artwork. Next, I narrowed this definition's extension to exclude features that reference disgusting or offensive subject matter. Finally, I posited that there are two ways in which artworks can morally harm their viewer: either through imaginative engagement leading to beliefs, or through emotional response leading to beliefs. I propose that since neither imagination nor emotion is sufficient for us to adopt beliefs that parallel ones presented in artworks, artworks cannot morally harm us.

It is an interesting question as to how persons have come to believe that art has this power

to corrupt one's character. While there is not space to give a definitive answer, I suggest that the answer may lie in our belief that art is an intimate form of human expression. Not only do we view art as an expression of emotions or beliefs, but also, a way of making these attitudes tangible. This intuition also applies to the creation of art. Persons who create art may believe that doing so is a way of realizing the mental (e.g., beliefs or emotions) as physical. If people feel this way about art, it is not surprising that they might believe that art has a kind of power that, for instance, verbal communication between agents does not. In viewing art, we have direct contact with expressive qualities, like joy, versus just the appearance of these expressions (as we might have when a person smiles).

If we are wrong in attributing to art this special ability to strongly influence our attitudes, then what does this say about the power of art? At the beginning of this paper, I presented a concern from both Danto and Julius that theory might remove (or sideline) powerful iconoclastic elements present in art. Indeed, one might worry that given the arguments presented here, artworks have little to no power to influence our beliefs. Moreover, one might worry that my argument also strips artworks of their power to make us better people. The critic may argue that just as artworks are unlikely to cause us to adopt immoral attitudes, they are unlikely to cause us to adopt good ones. So, Danto's worries seem corroborated: the theory of art given here leads one to the conclusion that artworks rarely affect their viewers.

Nevertheless, arguing that artworks are unlikely to affect our moral beliefs in the way the moral harm theorist believes, does not make art powerless. I argue that such an inference, would ignore the many other values that art can endorse. If we endorse a plurality of values, then the number of ways in which artwork can influence us increases. An artwork may occasionally cause us to adopt moral beliefs, but in the instances in which it does not, we need not believe its causal

disposition lies dormant. Although conditions may not often be right for some artwork to influence moral beliefs, they may be right for the artwork to influence other beliefs that are a source of value. For instance, art can cause us to reexamine a political issue, appreciate nature, or have a new emotional experience. It can alert us to some new way of perceiving the world, or inspire curiosity. None of the belief-adjustments that go along with the examples are necessarily moral in nature. If appreciation for Degas's sketches of ballerinas causes us to pursue ballet, this seems a valuable project; though not necessarily a moral one.

Likewise, art is powerful in its ability to spark dialogue. It need not be the efficient cause of our adoption of some moral belief (or any belief for that matter) for it to play a valuable role as evidence. Because art encourages imagination and consideration of counterfactual states of affairs, it can allow us to consider ideas we might not normally consider. We can view art then as providing a space, a starting point, for moral inquiry. The importance of contradictory and varying opinion to moral and intellectual development has been championed by many critics such as Hume, Mill, and Wayne C. Booth. Booth suggests that we do not leave artworks with complete understanding: "We may leave a fine performance of [some] play feeling in some sense purged...but we are not purged of all moral confusion or questioning. Rather we are steeped in the task of interrelating, comparing, contrasting rival moral judgments: we are caught up in moral inquiry" (2006, 259). If one accepts that art can influence a large range of beliefs, not just moral ones, Danto's concerns are lessened: we should not judge art to be powerful or weak based on particular cases of influence, such as moral harm; rather, we should base our judgment on art's breadth of influence. Artwork may have limited ability to influence any particular belief, but the range of art's influence is wide. In this respect, art is powerful.

## REFERENCES

- Booth, Wayne C. 2006. *The Essential Wayne Booth*. Edited by Walter Jost. Chicago: University of Chicago Press.
- Carroll, Noël. 1998. "Moderate Moralism versus Moderate Autonomism." *British Journal of Aesthetics* 38 (4): 419.
- Chapman, Hanah A., and Adam K. Anderson. 2013. "Things Rank and Gross in Nature: A Review and Synthesis of Moral Disgust." *Psychological Bulletin* 139 (2): 300-27.
- Curtis, Val, Robert Aunger, and Tamer Rabie. 2004. "Evidence that Disgust Evolved to Protect from Risk of Disease." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271 (4): S131–S133.
- Danto, Arthur Coleman. 1992. "Dangerous Art." In *Beyond the Brillo Box: The Visual Arts in Post-Historical Perspective*, 179–98. University of California Press.
- . 2001. "Beauty and the Beastly." *Nation* 272 (16): 25-9.
- Darwin, Charles. 2009. *The Expression of the Emotions in Man and Animals*. Cambridge: Cambridge University Press.
- Devereaux, Mary. 2004. "Moral Judgments and Works of Art: The Case of Narrative Literature." *Journal of Aesthetics & Art Criticism* 62 (1): 3–11.
- Dower, John. 1986. *War without Mercy: Race and Power in the Pacific War*. New York: Pantheon Books.
- Dubin, Steven C. 1992. *Arresting Images: Impolitic Art and Uncivil Actions*. Routledge.
- Eaton, A. W. 2003. "Where Ethics and Aesthetics Meet: Titian's Rape of Europa." *Hypatia* 18 (4): 159–88.
- Feinberg, Joel. 1985. *Offense to Others*. Vol. 4. New York: Oxford University Press.
- Frijda, Nico H., Antony S. R. Manstead, and Sacha Bem. 2000. *Emotions and Beliefs: How Feelings Influence Thoughts*. Studies in Emotion and Social Interaction. New York: Cambridge University Press.
- Gaut, Berys. 1998. "The Ethical Criticism of Art." In *Aesthetics and Ethics: Essays at the Intersection*, 1st ed., 182–203. Cambridge: Cambridge University Press.

- Gayford, Martin. 2007. "Allen Jones: The Day I Turned down Stanley Kubrick." *Telegraph.co.uk*, sec. art. <http://www.telegraph.co.uk/culture/art/3668398/Allen-Jones-The-day-I-turned-down-Stanley-Kubrick.html>.
- Gendler, Tamar Szabó. 2003. "On the Relation between Pretense and Belief." In *Imagination Philosophy and the Arts*, edited by Matthew Kieran and Dominic McIver Lopes. Routledge.
- . 2006. "Imaginative Contagion." *Metaphilosophy* 37 (2): 183–203.
- Gendler, Tamar Szabó, and Karson Kovakovich. 2005. "Genuine rational fictional emotions." In *Contemporary Debates in Aesthetics* Matthew Kieran (Ed.), 241–253. Oxford: Blackwell.
- Haidt, Jonathan, Clark McCauley, and Paul Rozin. 1994. "Individual Differences in Sensitivity to Disgust: A Scale Sampling Seven Domains of Disgust Elicitors." *Personality and Individual Differences* 16: 701-13.
- Hickey, Dave. 2000. "Vanessa Beecroft's Painted Ladies." In *VB 08–36: Vanessa Beecroft Performances*. 5–8.
- Hofer, Carl. 2008. "Introducing Nancy Cartwright's Philosophy of Science." In *Nancy Cartwright's Philosophy of Science*, edited by Luc Bovens, Stephen Hartmann, and Carl Hofer. Routledge Studies in the Philosophy of Science. Taylor & Francis.
- Jardine, Lisa. 1996. "Robert Mapplethorpe: Squandering Beauty." *Modern Painters* 9: 68-71.
- Jones, Peter. 2011. "Religious Belief and Freedom of Expression: Is Offensiveness Really the Issue?" *Res Publica* 17 (1): 75-90.
- Julius, Anthony. 2002. *Transgressions: The Offences of Art*. University of Chicago Press.
- Kass, Leon. 1997. "The Wisdom of Repugnance: Why We Should Ban the Cloning of Humans." *The New Republic* 216 (22): 17-26.
- Kelly, Daniel. 2011. *Yuck!: The Nature and Moral Significance of Disgust*. Cambridge, MA: MIT Press.
- Koppelman, Andrew. 2005. "Does Obscenity Cause Moral Harm?" *Columbia Law Review* 105 (5): 1635–79.
- Kosslyn, S.M., A. Pascual-Leone, O. Felician, S. Camposano, J.P. Keenan, W.L. Thompson, G. Ganis, K.E. Sukel, and N.M. Alpert. 1999. "The Role of Area 17 in Visual Imagery: Convergent Evidence from PET and rTMS." *Science* 284, no. 5411: 167-70.

- Landy, Joshua. 2012. *How to Do Things with Fictions*. New York : Oxford University Press.
- . 2010. "Corruption By Fiction." *Republic of Letters*. 1(2): 67-76.
- Nussbaum, Martha C. 2004. *Hiding from Humanity: Disgust, Shame, and the Law*. Princeton, NJ: Princeton University Press.
- Olivera La Rosa, Antonio, and Jaume Rosselló Mir. 2013. "On the relationships between disgust and morality: A critical review." *Psicothema* 25 (2): 222-26.
- Plato. *Republic*. 1997. In *The Complete Works of Plato*, Edited by John M. Cooper, 971-1223. Indianapolis: Hackett Publishing Company.
- Rozin, Paul, Maureen Markwith, and Carol Nemeroff. 1992. "Magical Contagion Beliefs and Fear of AIDS." *Journal of Applied Social Psychology* 22 (14): 1081-92.
- Rozin, Paul, Carol Nemeroff, Matthew Horowitz, Bonnie Gordon, and Wendy Voet. 1995. "The Borders of the Self: Contamination Sensitivity and Potency of the Body Apertures and Other Body Parts." *Journal of Research in Personality* 29: 318-40.
- Rozin, Paul, Laura Lowery, Sumio Imada, and Jonathan Haidt. 1999. "The CAD Triad Hypothesis: A Mapping between Three Moral Emotions (Contempt, Anger, Disgust) and Three Moral Codes (Community, Autonomy, Divinity)." *Journal of Personality and Social Psychology* 76 (4): 574-86.
- Rozin, Paul 2008. "Hedonic 'Adaptation': Specific Habituation to Disgust/Death Elicitors as a Result of Dissecting a Cadaver." *Judgment and Decision Making* 3 (2): 191-94.
- Russell, Pascale Sophie, and Roger Giner-Sorolla. 2013. "Bodily Moral Disgust: What It Is, How It Is Different from Anger, and Why It Is an Unreasoned Emotion." *Psychological Bulletin* 139 (2). 328–51.
- Steinmetz, Julia, Heather Cassils, and Clover Leary. 2006. "Behind Enemy Lines: Toxic Titties Infiltrate Vanessa Beecroft." *Signs: Journal of Women in Culture & Society* 31 (3): 753–83.
- Tajima, Atsushi, and Michael Thornton. 2012. "Strategic Solidarity: Japanese Imaginings of Blacks and Race in Popular Media." *Inter-Asia Cultural Studies* 13 (3): 345–64.
- Van Leeuwen, Neil. 2011. "Imagination Is Where the Action Is." *Journal of Philosophy* 108 (2): 55–77.
- . 2013. "The Meanings of 'Imagine' Part I: Constructive Imagination." *Philosophy Compass* 8 (3): 220–30.

———. Forthcoming. “The Meanings of ‘Imagine’ Part II: Attitude and Action.” *Philosophy Compass*.

Vermazen, Bruce. 1986. “Expression as Expression,” *Pacific Philosophical Quarterly* 67: 196-224.

Wollheim, Richard. 1998. On pictorial representation. *Journal of Aesthetics and Art Criticism* 56 (3):217-226.