

*Ex Nihilo*

*The University of Texas at Austin*

*Spring 2013*

This year's submissions, though selected purely for their philosophical merit, share several themes. Most obvious is that of the *self*, or *self-reflection*, which guides our contributors' inquiries into morality, politics, aesthetics, and knowledge. We hope their work will inspire our readers to think about themselves as agents, as appreciators, and, of course, as philosophers.

We would like to thank Susan Somers for her patient guidance, Dr. Al Martinich for reading our final proofs, our readers for their tireless work, and our authors for their thoughtful contributions. We would also like to thank the previous student leaders of Ex Nihilo, especially Pavel Nitchovski and Sophia Portier, for their support.

Rebecca Besaw & Daniel Muñoz  
Co-Editors-in-Chief

EX NIHILO

## Ex Nihilo Staff

Daniel Muñoz & Rebecca Besaw  
*Co-Editors-in-Chief*

Robert Garcia  
*Data Manager*

Joshua Sklar  
*Reader*

Magdalene Burger  
*Reader*

Joshua Morrow  
*Reader*



Jacket design by Daniel Muñoz  
Adapted from *Before the Mirror* by T.W. Dewing  
Oil painting on canvas, 1910

## Table of Contents

<b>Paying for Past Sins: Climate Change and Compensatory Justice</b>	<b>1</b>
<i>Greg Yanke Arizona State University</i>	
<b>The Aesthetic Divergence of Mirror Images</b>	<b>16</b>
<i>Johnathan Ochs The University of Maryland</i>	
<b>Testimony, Evidence, and Epistemic Individualism</b>	<b>38</b>
<i>Kevin Dorst Washington University in St. Louis</i>	
<b>The Deep Self View and Wolf's Sanity Condition: An Unnecessary Addendum</b>	<b>52</b>
<i>Sabeen Ahmed The University of Virginia</i>	
<b>Autonomy and Integrity</b>	<b>67</b>
<i>Victoria Koc The University of Pennsylvania</i>	

# Paying for Past Sins: Climate Change and Compensatory Justice

*Greg Yanke*

***Abstract:***

*Ethical claims regarding the obligation of developed countries to pay developing countries for the impact of their greenhouse gas emissions are usually based on the principle of compensatory justice. However, framing the moral obligations of developed countries in terms of compensation for past emissions is problematic due to the complicated temporal and causal aspects of climate change. Identifying climate change offenders and victims is difficult due to changing emission trends, large differences between aggregate and per capita emissions in many countries, and obstacles to proving that climate change damage has occurred and to quantifying it.*

Much of the philosophical literature concerning the ethics of climate change relies on two key premises: that developed countries are primarily responsible for historical greenhouse gas emissions that are causing climate change and that developing countries will suffer most of the adverse effects that will result. These claims are generally true. Of the cumulative greenhouse gases that humans have emitted since 1850, developed countries are responsible for approximately 76% of them (Baumert, Herzog, & Pershing, 2005; p. 32). Yet, developed countries will likely fair better in response to global warming than developing countries, which tend to be situated in more vulnerable, low latitude regions. In fact, in terms of expected change in food production, many developed countries are expected to

benefit from climate change (Parry, Rosenzweig, Iglesias, Livermore, & Fischer, 2004; p. 63). In contrast, developing nations, particularly in Africa, Southeast Asia, Central America, and South America, will likely face the most harm from climate change impacts (Parry et al., 2004; p. 63).

Given these circumstances, compensatory justice arguments that assert that developed countries are obliged to pay developing countries for the impacts of their emissions should be very persuasive on their surface. As Pogge (2005) states in his article, *World Poverty and Human Rights*, it is unfair that “the global poor benefit least, if at all, from polluting activities, and also are least able to protect themselves from the impact.” (p. 6). However, greenhouse gas emission trends and other complexities challenge the generalization that all developed countries are climate change wrongdoers and that poor nations are the victims. The evidence suggests that framing the moral obligations of developed countries in terms of compensation for past emissions is problematic. The compensatory justice framework is ill-suited for incorporating the complicated temporal and causal aspects of climate change.

### **Compensatory Justice**

Issues of compensatory justice typically arise in cases where a person or group acts in a manner that results in harm to one or more victims. Each offender is expected to compensate each victim for the amount of harm he or she suffered due to the offender’s actions. Aristotle characterized compensatory justice as a means of restoring equality by removing an ill-gotten gain from an offender and returning it to the victim (1987; p. 156). It is consistent with the value that is reflected in the key principle of tort law that we should return the victim to the *status quo ante*,

the position he or she would have been in if the harm did not occur (Hill, 2002; p. 394). When this is not practical or possible, the offender's obligation shifts to making the victim whole, which requires the offender to make the victim at least as well-off as he or she was prior to the harm (Lomasky, 1991; p. 27). Climate change compensation tends to fit into this last category since polluters cannot undo their past emissions and the damages suffered are not reversible.

In the context of climate change, compensatory justice is encompassed by the "polluter pays" principle. This principle states that if an agent performs an action that causes pollution, the agent must pay for the damages resulting from that action (Caney, 2005; p. 753). While compensatory justice has intuitive appeal in the climate change realm, the adjudication of compensation claims requires that we are able to identify the offenders and the victims, establish causality between the conduct of the offenders and the damages that the victims suffer, and quantify the damages that each victim suffered. Due to the unique characteristics of climate change, this proves to be difficult, if not impossible.

### **Identifying Climate Change Offenders**

Greenhouse gas emissions are not necessarily harmful. Rather, it is the excessive discharge of these gases into Earth's atmosphere that is causing climate change. In order to determine who the climate change offenders are, it is necessary to determine how we should measure responsibility for excessive greenhouse gas emissions. This depends on whether our unit of reference is the country, the individual, or some other entity, and on the amount of permissible emissions per that agent. If the unit of reference is the country, it seems unreasonable to suggest that the emission cap for a small state, such as

Monaco, should be the same as the cap for China since their populations are so different. Although there may be a moral obligation for countries to control their population levels in order to impede environmental degradation, each person needs to emit some greenhouse gasses for survival. When determining a person's permissible emission levels, the population in one's country of residence does not seem like a morally relevant criterion. Thus, if we only consider each country's aggregate greenhouse gas emissions, nations with high populations that emit very little per individual, such as China and India, who rank 99<sup>th</sup> and 140<sup>th</sup> respectively on a per capita basis (Baumert, Herzog, & Pershing, 2005, p. 22), would be heavily penalized.

Considering per capita emissions as a basis for restitution is also fraught with difficulties. Although it would compel countries such as the United States, Canada, and Australia to compensate developing countries for climate change, it would also implicate much of the Middle East, parts of Eastern Europe, and many small-island states (Baumert et al., 2005, p. 21) that are not generally considered to be developed. The latter group is particularly vulnerable to the effects of climate change since their low-elevation land masses are disappearing due to rising sea level increases. If these small island nations are considered to be climate change offenders because of their high per capita emissions, which when aggregated have an infinitesimal impact on global warming, compensatory justice does not provide them with a moral claim for assistance.

Compensatory justice also fails to consider the effect of requiring a polluting country to pay for its emissions. Regardless of how responsibility for excessive greenhouse gas emissions is assessed, some developing countries will likely be identified as perpetrators primarily due to the actions of their wealthy residents. For example, Turkmenistan, a top 20 per capita emitter, and India, a top

ten cumulative, aggregate emitter, respectively rank 97<sup>th</sup> and 138<sup>th</sup> in per capita gross domestic product adjusted for purchasing power parity (Baumert et al., 2005, pp. 22, 32; International Monetary Fund, 2012). Funds that these countries could be morally obligated to pay as restitution under a compensatory justice structure would likely be better spent to alleviate domestic poverty. Arguably, developing countries could be excused from their obligations to compensate others for their climate-changing emissions, but such a decision would be based on another principle or theory, such as distributive justice or consequentialism, rather than compensatory justice.

Because greenhouse gas emissions have occurred over centuries, the identification of climate change offenders becomes more complex. Some authors have suggested that forcing current residents of developed countries to pay climate change compensation is unfair since they were not responsible for historical emissions (Caney, 2005; Posner & Weisbach, 2010). However, this argument is subject to two key criticisms: most emissions have occurred in the recent past and residents of countries that have polluted the most on a cumulative basis have tended to benefit from the resulting economic fruits of industrialization.

With respect to the first criticism, Boden, Marland, and Andres (2010) estimate that global fossil-fuel emissions account for 337 billion tons of carbon being released into the atmosphere since the start of the Industrial Revolution. Of this amount, approximately half of the emissions have occurred since the mid-1970s (Boden et al., 2010). Thus, much of the climate change impacts that we are experiencing today are a result of emissions from those who are still alive. As well, today's global residents share culpability by continuing to pollute at ever increasing rates. Global emissions of carbon dioxide increased by 5% in

2010 and by a further 3% in 2011 (Olivier, Janssens-Maenhout, & Peters, 2012).

Even if some current residents of developed nations were not alive when their predecessors emitted significant amounts of greenhouse gases, they may still be morally obligated to compensate low-emitters if they currently enjoy benefits derived from those emissions. Arguably, developed countries are relatively affluent today because their past, excessive emissions allowed them to generate wealth through industrialization processes. As Nozick (1974) states, “the holdings of a person are just if he is entitled to them by the principles of justice in acquisition and transfer, or by the principle of rectification of injustice.” (p. 153). If the current standard of living that today’s residents enjoy is due to ancestral pollution, that resulting wealth constitutes unjust enrichment and beneficiaries should compensate those who suffer from the impacts of climate change.

Pogge (2005) makes a similar argument in the context of world poverty. He contends that global wealth disparity is a result of historical wrongs that developed nations inflicted on developing countries during the colonial era. While Pogge acknowledges that “we indeed cannot inherit responsibility for our forefather’s sins”, he asks “how then can we plausibly claim the fruits of their sins?” (2005; p. 2). In *World Poverty and Human Rights*, Pogge (2002) states that developing countries’ uncompensated exclusion from the use of the planet’s natural resources violates a negative duty not to gain from the “unjust impoverishment of others” (p. 197). Assessing resource use from the perspective of the Lockean Proviso that we leave “enough, and as good” (Locke, 1689) for others when we unilaterally consume resources, Pogge concludes that developed countries have created radical inequality without justification (2002; p. 202). While Pogge does not specifically consider the ability of the Earth to

absorb carbon as a resource, his argument can be applied to climate change.

If the largest cumulative emitters of greenhouse gases have indeed enjoyed lifestyle advantages from their polluting activities, then there should be a significant, positive correlation between their emissions and their economic output. A comparison of the aggregate gross domestic product (“GDP”) figures for the countries that have emitted significant amounts of greenhouse gases confirms this (refer to Appendix 1 and 2 for the data and statistical results). There is an extremely high correlation of 0.916 (with 1.0 representing a perfect correlation) between an emitting country’s percentage of cumulative emissions and GDP adjusted to consider purchasing power parity. Though it is possible that population could influence this correlation since both GDP and emissions are expected to be greater in countries with more people, the correlation between population and emissions (0.271) and population and GDP (0.568) are far lower.

Because GDP is an imperfect proxy for standard of living and well-being, it is also useful to consider the correlation between a nation’s cumulative emissions and its Human Development Index (“HDI”) rating, a metric that Mahbub ul Haq and Amartya Sen created that includes threshold human health and education measures, along with income (Gertner, 2010). The correlation of emissions with HDI (0.309) is much lower than the one with GDP, which suggests that although countries may realize benefits from their polluting actions, residents share disproportionately in this benefit. Thus, although it is possible to identify those countries that have gained the most from greenhouse gas emissions, it would be unfair to deem all of their residents as offenders for the purpose of a compensatory regime. While this could be remedied with an appropriate tax code, as previously mentioned, such a method of identifying

climate change offenders does not overcome the serious difficulties associated with using aggregate emissions.

The preceding discussion is not meant to suggest that it is impossible to identify any climate change offenders based on the various criteria discussed. Certainly the United States, Canada, and Australia should be considered partially responsible for the effects of greenhouse gas emissions since they all rank high on the list of cumulative, as well as current aggregate and per capita emissions (Baumert et al., 2005). However, the general claim that all developed nations are morally responsible for paying climate change compensation is much more problematic.

### **Identifying Climate Change Victims**

Identifying climate change victims also involves major pitfalls. The generalization that developed countries are the greenhouse gas emitters and that developing countries will be the victims fails to acknowledge the dynamics of pollution trends. Though developed countries have historically been the largest polluters, developing countries now produce over half of greenhouse gases (Olivier et al., 2012; p. 17). As well, emission growth rates are highest among developing countries and this trend will likely continue (Baumert et al., 2005; p. 13).

Globalization complicates the issue of responsibility for emissions due to the prevalence of foreign corporations operating in developing countries. The penetration of foreign capital in developing countries has a significant impact on greenhouse gas emissions, which exceeds the impact of domestic investment (Grimes & Kentor, 2003). Foreign corporations tend to invest in energy intensive manufacturing in such countries, which also tend to be greater polluting industries due to weaker environmental controls in developing countries (Grimes & Kentor, 2003).

It becomes difficult to allocate responsibility for these emissions. The foreign corporations certainly share accountability due to the profits they earn from the emitting activities, but the governments of the developing countries consent to the foreign investment and enjoy economic benefits as well and it is therefore more difficult to identify them as climate change victims.

In addition, developing countries will emit substantial greenhouse gases as part of their industrialization. To the extent that they are able to develop more efficiently than previous nations, this will likely be the result of technological knowledge that they receive from developed countries. Thus, victimhood should not be based on a country's lack of historical emissions, but rather on damages that it has suffered from climate change impacts. The problem is that few groups can establish that they are victims because almost all of the harm from climate change is expected to occur in the future.

### **Establishing and Quantifying Damages**

Because compensatory justice requires that offenders compensate their victims for the amount of harm they suffer as a result of their misdeeds, the damages must occur and be reasonably quantifiable before payment is due. Though it may be likely that sub-Saharan countries will experience increased incidence of drought or that Caribbean countries will face more cataclysmic weather disasters as a result of climate change, the extent of the damage they will endure is unknown and not guaranteed. Thus, they are not currently entitled to redress under a compensatory justice framework. While it is true that compensation will be due to future victims after disaster strikes, those facing the adverse effects of climate change could use assistance now in order to undertake adaptation measures.

Adaptation involves modifying natural and human systems in order to reduce or eliminate the expected harm from potential climate change impacts (IPCC, 2001). In areas vulnerable to drought, it may involve using crop varieties that consume less water, employing better water management techniques, and developing an infrastructure to store excess food produced. In areas vulnerable to increased storms, it may involve creating better early warning systems, using different building construction techniques, and implementing disaster response regimes. Though different countries face diverse hurdles depending on their geographical location and available resources, they all have one thing in common: adaptation requires money. While developed countries that will face significant climate change impacts, such as Australia, have the capital necessary to pursue adaptation measures now, developing nations are not so fortunate. As a result, the impacts of climate change are likely to be exacerbated in developing countries since they will not have the infrastructure necessary to prepare for and cope with disaster events.

Even if disasters do occur and the resulting damages may be quantified, it is very difficult for the victims to assert a right to compensation due to the imprecise causal relationship between greenhouse gas emissions and specific weather events. For example, when Hurricane Katrina entered the Gulf of Mexico in 2005, surface water temperatures were two degrees warmer than normal, which likely intensified its effects (Center for Climate and Energy Solutions, n.d.). However, although scientists have been able to measure ocean temperature increases (Levitus, Antanov, & Boyer, 2005), they have not been able to “partition the observed warming to an anthropogenic component or a component associated with natural variability.” (Levitus et al., 2000).

Thus, when a weather-related disaster occurs, it is difficult for victims to demonstrate that they would not

have suffered the damages they did if not for the influence of greenhouse gas emissions. After all, devastating natural disasters occurred on Earth prior to the emergence of anthropogenic climate change. Yet, as scientific evidence accumulates over longer time periods, the hurdle of proving a causal link between greenhouse gas emissions and specific disaster events may diminish. Even with the recent disaster resulting from Hurricane Sandy, though there is no definitive association between the storm's magnitude and human-induced climate change, scientists can link the severity of the damage from the storm surge that impacted the East Coast of the United States and recent anthropogenic sea level increases (Gillis, 2012). However, scientific evidence can also impair a victim's compensatory justice claim as well. For example, though numerous hurricanes and tropical storms have ravaged Haiti's landscape over the past decade, dramatic domestic deforestation has greatly intensified the damage that the country has suffered (Malkin, 2012). In essence, Haiti has contributed significantly to the harm it experiences from natural disasters.

Even if climate change disaster victims are able to establish that they have suffered damages due to the actions of those who have polluted, the offenders arguably have a moral right of set-off. This is akin to the legal principle that allows someone who is found liable for harming a third-party to deduct an amount from damages for harm that the third-party has caused to the perpetrator. In the climate change context, a country such as Australia, which is likely to suffer from climate change impacts, could refuse to compensate developing countries for harms that it has caused because those countries have, or will, pollute at least as much in the future.

The Inuit constitute a cultural group that appears to have a strong moral claim as climate change victims, but could arguably be subject to such a right of set-off. Arctic

warming from greenhouse gas emissions has and will to continue impact the indigenous group's historical mode of transportation, its hunting and food-sharing culture, and its melting landscape (Watt-Cloutier; pp. 25-26). Yet, many of the Inuit are also embracing the economic benefits arising from the global demand from petroleum products and minerals, and are becoming significant greenhouse gas emitters.

The approximately 150,000 people who identify as Inuit reside in multiple countries, though about one-third of them located in Greenland, which accounts for almost 90% of the country's population (Hessel, 2006; pp. 9, 20). In addition to hosting billions of barrels of petroleum reserves (Stigset, 2011), ongoing exploration suggests that the region is also home to large mineral resources, which will become easier to access as ice and permafrost melt (Harvey, 2012). Even now, with most of its natural resources still in the ground, Greenland is ranked 33<sup>rd</sup> in per capita carbon emissions, which is significantly higher than almost all European countries (Carbon Dioxide Information Analysis Center, 2009). The Inuit realize that their greenhouse gas production is likely to rise as they extract their valuable resources and have expressed a desire to be exempt from international emissions reduction targets (Bell, 2009). Thus, even though the Inuit are losing aspects of their heritage as their landmasses slowly melt, their moral entitlement to compensation may be jeopardized by a moral right of set-off that considers their likelihood to pollute in the future.

### **Conclusion**

The moral claim that developed countries should compensate developing countries for the harm that they are likely to suffer as a result of climate change disasters initially seems compelling. However, an attempt to apply a

compensatory justice framework to the climate change situation is fraught with problems. It is difficult to identify offenders that should pay compensation since emissions are dynamic. Developing nations now pollute more than developed ones and will continue to do so. As well, cumulative, aggregate, and per capita emissions data identify significantly different lists of offenders, some of which are very poor and will suffer significantly from the expected impacts of climate change. Identifying the victims of climate change is also an imprecise assignment. Few countries can currently establish that they are victims and the damages that they will suffer are both contingent and unknown. Even when weather-related disasters do occur, it is difficult to know the causal role climate change played in its occurrence.

Despite the difficulties in applying principles of compensatory justice to climate change, we still possess the moral intuition that developed countries should aid the victims of climate change disasters and help them adapt to potential climatic events. It may be that distributive justice or luck egalitarian concepts that require developed countries to alleviate the suffering of those in developing countries who are less fortunate through no fault of their own may explain this. Further research should consider the application of these frameworks to the climate change issue. This would include the apparent implication that if society focuses on reducing inequality, rather than compensating for past wrongs, climate change would only be morally relevant to the extent that it increased the incidence of poverty.

## Appendix 1: Emission and Economic Data

	% of	GDP		
	Cumulative Emissions	PPP (Intl \$B)	Population	HDI
US	29.3	15.076	315	0.910
EU-25	26.5	15.853	504	0.867
Russia	8.1	2.383	143	0.755
China	7.6	11.3	1347	0.687
Germany	7.3	3.114	82	0.905
UK	6.3	2.288	62	0.863
Japan	4.1	4.444	128	0.901
France	2.9	2.214	65	0.884
India	2.2	4.421	1210	0.547
Ukraine	2.2	0.725	46	0.729
Canada	2.1	1.395	35	0.908
Poland	2.1	0.771	39	0.813
Italy	1.6	1.847	61	0.874
South Africa	1.2	0.554	52	0.770
Australia	1.1	0.915	23	0.929
Mexico	1	1.667	112	0.770
Spain	0.9	1.406	46	0.878
Brazil	0.8	2.294	194	0.718
South Korea	0.8	1.554	50	0.897
Iran	0.6	0.99	75	0.707
Indonesia	0.5	1.125	238	0.617
Saudi Arabia	0.5	0.688	28	0.770
Argentina	0.5	0.716	40	0.797
Turkey	0.4	1.075	75	0.699
Pakistan	0.2	0.488	181	0.504
Source:	Baumert et al. (2005)	IMF (2012)	various	UNDP (2011)

## Appendix 2:

# Correlation Matrix

	<i>Cumulative Emissions</i>	<i>GDP</i>	<i>Population</i>	<i>HDI</i>
Cumulative Emissions	1			
GDP	0.915576201	1		
Population	0.270997065	0.568731107	1	
HDI	0.309442212	0.171884156	-0.438946675	1

# The Aesthetic Divergence of Mirror Images

*Johnathan Ochs*

***Abstract:***

*This paper treats objects and their mirror images as two distinct objects (or stages of one and the same object) that are identical in every respect, except one is the mirror-reversed twin of the other. It asks, "Can an object differ aesthetically from its mirror-reversed self?" The greatest perceptible difference between such objects is the ostensive opposite orientation that they embody in relation to one another. This paper focuses its attention to orientation and explores whether a change in an object's orientation under mirror-reversal can cause an aesthetic divergence between the original object and its mirror-reversed self.*

**INTRODUCTION**

In this paper, I will explore the question: "Can an object differ aesthetically from its mirror-reversed image?" For my purpose, an object's *mirror-reversed image* is a real, space-occupying object that is identical in size and dimensions to the mirrored-object; as if one were to physically rearrange an object's particles in such a way that the resulting object would be the mirror-reversal of the original. An object's mirror reversal can come about from one of two ways. In the first, an object can be duplicated, in which case there would be two distinct objects that are exactly identical in all respects, except, one is the mirror-reversed representation of the other. In the second, an object may go through a process of mirror-reversal; in

which case, one and the same object becomes its own mirror-reversed self.<sup>1</sup> In both cases, I will refer to the mirror-reversed product as the original's "flop." Similarly, I will refer to the process of mirror-reversal as the "flopping" of an object.

I will begin by analyzing the opposing accounts of two philosophers who attempted to answer the question above. In "The Aesthetics of Mirror Reversal," Roy Sorensen proposes that there is no aesthetic difference between an object and its flop.<sup>2</sup> In "A Note on the Aesthetics of Mirror Reversal," Rafael De Clercq critiques and seeks to directly refute Sorensen's argument.<sup>3</sup> De Clercq argues that Sorensen assumes a highly controversial supervenience principle in establishing his argument and, therefore, the matter is still open for interpretation. I will argue that De Clercq has the stronger account, and that he successfully illustrates a situation in which an object can indeed differ aesthetically from its flop.

My paper will be divided into three main sections. In §1, I will discuss Sorensen's arguments against the aesthetic relevance of mirror reversal. In §2, I will explain De Clercq's counter-arguments to Sorensen's claims and evaluate their strengths and weaknesses. I will also address a possible response to De Clercq's argument for the aesthetic relevance of changes in orientation in order to show how his argument is particularly effective in refuting Sorensen's opposing claim. Finally, in §3, I will evaluate Sorensen's and De Clercq's accounts in relation to two important distinctions pertaining to the topic of aesthetics and mirror reversal in order to reveal a significant shortcoming about Sorensen's account. Thus, proving in my conclusion that an object can differ aesthetically from its flop.

## **1. ANALYSIS OF SORENSEN'S ARGUMENTS**

Sorensen argues that there is no aesthetic difference between an object and its flop. His reasoning is as follows: An object's aesthetic properties *supervene* on its *intrinsic* properties, never on its *extrinsic* properties. Further, an object's intrinsic properties are absolutely preserved under mirror-reversal; therefore, since nothing intrinsic changes, nothing aesthetic changes (Sorensen, 2000).

In order to make Sorensen's argument more intelligible for my purposes, I will explain some of the terms that he uses, and what I take them to mean. First, the difference between *intrinsic* and *extrinsic* properties: Intrinsic properties are those that are inherent to the object in question. For instance, in a painting, the colors used or the characters portrayed in a depicted scene are intrinsic to the painting. Extrinsic properties are, contrarily, not inherent to the object in question. For instance, whether the painting is located in New York or in London is extrinsic to it.

A second term that is relevant to Sorensen's argument is '*supervenience*'. The Stanford Encyclopedia of Philosophy defines supervenience thus: "A set of properties *A* supervenes upon another set *B* just in case no two things can differ with respect to *A*-properties without also differing with respect to their *B*-properties" (McLaughlin, 2005). Given this definition, Sorensen's claim—that an object's aesthetic properties *supervene* on its intrinsic properties—amounts to the claim that no aesthetic properties about the object can change without its intrinsic properties changing as well. Since, according to Sorensen, nothing intrinsic about an object changes when it is flopped, it follows that nothing aesthetic changes either. For this reason, Sorensen concludes that mirror-reversal is aesthetically irrelevant.

### **a. Left/Right Orientation**

Sorensen's second argument addresses the notion that when an object is flopped, there might be an aesthetic difference as a result of the change in orientation. If a *left-oriented* object is mirror-reversed, then its flop will be *right-oriented*; that is, the process of mirror-reversal produces an object that has an orientation exactly opposite of its original orientation. Suppose that an object has some aesthetic property which results from its being left-oriented. If the object were flopped, it would become right-oriented and the aesthetic property would ostensibly be lost during the process. If such is the case, then the difference in orientation between an object and its flop would be aesthetically relevant.

However, Sorensen rules out such a case because he deems an object's left/right orientation to be an extrinsic property. He explains that *left* and *right* are observer-dependent properties; that is, an object's left/right orientation is determined by an observer's orientation in relation to the object. For example, if I say that a painting is oriented from left-to-right, it is because I impose my left and right upon the painting in relation to my position. Since orientation is an observer-dependent extrinsic property, and Sorensen deems all extrinsic properties aesthetically irrelevant, so too is orientation aesthetically irrelevant on his account.<sup>4</sup>

### **b. A Contradiction Concerning Numerical Identity**

Sorensen's third argument appeals to the numerical identity of a flopped object; that since an object and its flop are numerically identical, a contradiction is derived under the claim that flopping causes aesthetic changes because

## EX NIHILO

the object would then both have and not have certain aesthetic properties.<sup>5</sup>

Sorensen makes this point in reference to "universal reversals", i.e. flopping an entire universe. He explains:

*"A universal reversal would not alter the aesthetic merit of any artwork. Each flopped object would be numerically the same as the original. Hence, we would derive a contradiction from the supposition that the asymmetric object is better than its flop. For an object that is flipped through the fourth dimension would be better than itself"* (Sorensen, p. 186).

Sorensen supposes a four-dimensional being who may mirror-reverse our three-dimensional universe through the fourth dimension; just as we can flip a two-dimensional being (a "flatlander") through the third dimension by twisting his two dimensional universe into a möbius strip. If our universe is flipped in this way, then all objects, including ourselves, would be mirror-reversed. In such a situation, we would not know that any change had occurred because, in relation to us, all objects appear to be oriented exactly the same as they were before the reversal took place.

Sorensen's point is that the mirror-reversed objects in a flopped universe are numerically identical to the objects in the original universe. Therefore, saying that mirror-reversal changes an object's aesthetic merit, and the resulting flop is "better", yields a contradiction because we would be saying that the object is better than itself.

## **2. ANALYSIS OF DE CLERCQ'S ARGUMENTS**

De Clercq's essay serves to directly refute Sorensen's arguments. He begins by establishing that, for the purposes of his essay, he will only be writing about *enantiomorphs*. Enantiomorphs are identical objects, except one is the mirror-image of the other. Moreover, De Clercq extends the definition of enantiomorphous objects to be "objects that change orientation [when mirror-reversed]"<sup>6</sup> (De Clercq; p. 554). In contrast, an object that does not change orientation under mirror-reversal would be something that is perfectly symmetrical along the axis which the object is reversed along; such that no distinction can be made whether it is oriented from left-to-right or from right-to-left (e.g. an untainted billiard cue ball).<sup>7</sup>

### **a. A Plausible Weakening of the Supervenience Principle**

De Clercq's first critique is of Sorensen's main argument regarding the supervenience of aesthetic properties on intrinsic properties. He argues that the supervenience principle, upon which Sorensen's argument relies, is highly controversial and that it is merely assumed, not argued for. Further, the principle entails that all extrinsic properties are aesthetically irrelevant. De Clercq asserts that "only extreme formalists will deny that the aesthetic value of an object can be determined by its representational, historical and/or functional properties" (De Clercq; p. 555). Thus, he finds Sorensen's argument to be too narrow to conclude that there is no aesthetic difference between an object and its flop.

De Clercq supposes that Sorensen's argument may be improved if the assumed principle is weakened to only include certain extrinsic properties. Instead of saying that all extrinsic properties are irrelevant, Sorensen could say

that observer-dependent extrinsic properties are irrelevant. Doing so would establish the aesthetic irrelevance of orientation if orientation is an observer-dependent property. Sorensen argues that this is the case; therefore, the plausible weakening of his argument should work in his favor. However, De Clercq supposes that even in accordance with this plausible account, orientation may still be an aesthetically relevant extrinsic property; which would serve to disprove Sorensen's argument. He explains that even if observer-dependent properties, such as being left-oriented or right-oriented, are aesthetically irrelevant, a change in orientation might still bring out an aesthetic change.

**b. 'Determinate' Orientation versus 'Indeterminate' Orientation**

In order to show De Clercq's reasoning for his claim above, I will first clarify an ambiguity surrounding the notion of 'orientation'. There is a distinction between an object having a *particular orientation* (e.g. being left-oriented or being right-oriented) and an object having *some orientation* or *another* (i.e. enantiomorphs). The former is known as "determinate" orientation, and the latter is "indeterminate" orientation. (De Clercq; p. 556) This is an important distinction to draw because only determinate orientation is regarded as observer-dependent; whereas, indeterminate orientation, specifically, enantiomorphism, can be defined in a way that does not make it observer-dependent. De Clercq uses a definition of this sort for his argument that a change in orientation may be aesthetically relevant. The new definition of enantiomorphism goes as follows:

*"An object O is an enantiomorph just in case there is a neighborhood N of O such that N is large enough to admit reflections of O and that the result of every reflection of O*

*in N differs from the result of every rigid motion of O in N"*  
(De Clercq; p. 555; from Earman 1989; p. 141).

This definition does not implicitly or explicitly refer to observers, so it succeeds in showing that there need not be an observer present in order to determine the orientation of enantiomorphous objects. De Clercq explains, "as a consequence, it will not suffice to say that observer-dependent extrinsic properties are aesthetically irrelevant" (De Clercq; p. 556); for left/right orientation was deemed aesthetically irrelevant (because it is observer-dependent), but a change in orientation (i.e. indeterminate orientation) might still result in an aesthetic change.

To see this, consider that under the supposition that the left/right distinction is aesthetically irrelevant, it is still possible to have aesthetic properties distributed in such a manner that no enantiomorph has exactly the same aesthetic properties as its flop. In such a scenario, turning an object into its own mirror image would be an aesthetically significant process, even though it would not matter whether the operation achieved a left-to-right reversal or a right-to-left reversal.

For the scenario to be consistent, it first must be the case that an object can be said to have some orientation or other without having any orientation in particular. That is to say, the consistency of the scenario depends on the consistency of the following three sentences:

1.  $O$  is right-oriented or  $O$  is left-oriented.
2. It is indeterminate whether  $O$  is right-oriented.
3. It is indeterminate whether  $O$  is left-oriented (De Clercq; p. 556).

Here,  $O$  can be any enantiomorph. This combination is not inconceivable. In supervaluational semantics, a disjunction may be true without either disjunct being true. Therefore, there is at least one *prima facie* acceptable framework in which 1, 2, and 3 are consistent that cannot be rejected out of hand (De Clercq; pp. 556-557).

If 1, 2, and 3 above are consistent, then there's still one more consistency that needs to be investigated:

4. All enantiomorphs  $x$  and  $y$  are such that if one is the mirror image of the other, then there is an aesthetic property exemplified by  $x$  but not by  $y$ .

Here,  $x$  and  $y$  can be ordinary, persisting objects or stages of such objects (De Clercq; p. 559). Even if the antecedent of 4 is true, there need not be any object with a determinate orientation. The relation “being a mirror image of” implies that if one object is determinately left-oriented, then the other object is determinately right-oriented, and vice versa (De Clercq; p. 559). Therefore, 1, 2, 3, and 4 above form a consistent set. What this means is that orientation may matter aesthetically even if determinate-orientation does not.

### **c. De Clercq's “Possible Worlds” Analysis**

To make his argument more intelligible, De Clercq provides an example that illustrates a situation where the sentences 1, 2, and 3 above are all true. The example shows how cross-world comparisons of orientation are impossible, nonetheless, we are still able to conceptualize changes in enantiomorphs' relative, indeterminate orientations, which shows how a change in orientation can be aesthetically relevant.

Figure 1 below<sup>8</sup> shows two possible worlds  $w_1$  and  $w_2$ , each containing a pair of F-shaped enantiomorphs, where one is the mirror-reversed image of the other. For the purpose of the illustration, one object in each world is white and the other is shaded. Both worlds are two-dimensional Euclidean planes.

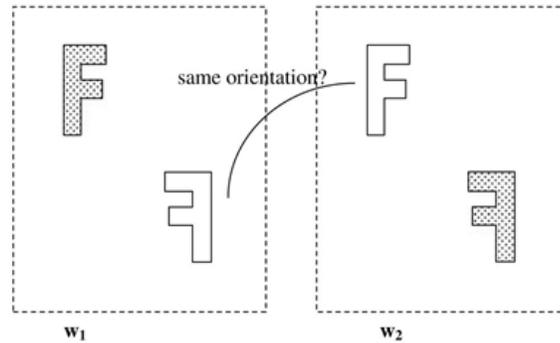


Figure 1. Comparing orientations across worlds.

Because the two objects are mirror images and enantiomorphs, they have opposite orientations. So, if one is left-oriented, then the other is right-oriented, and vice versa. Therefore, the forms can be said to have *some orientation or other*; namely, the orientation opposite to that of their mirror-reversed twin in the same plane. However, the two forms do not have a *determinate* orientation; for if they did, it would be possible, for instance, to compare the orientations of the forms in  $w_1$  with the orientations of the forms in  $w_2$ . If such were the case, then it would make sense to ask, "Does the white form in  $w_1$  have the same orientation as the white form in  $w_2$ ?" However, it is impossible to make such a comparison and, in principle, this question is unanswerable. The reason is that the forms do not have a determinate orientation (De Clercq; pp. 557-558).

De Clercq explains that if cross-world comparisons are impossible, then we cannot analyze a statement such as "I could have been oriented differently," because doing so requires positing a possible world where one is oriented differently. Thus, a cross-world comparison is made in the analysis, which was deemed impossible (De Clercq; p. 558). However, De Clercq's suggestion is to understand the statement, "I could have been oriented differently," in a relative manner; that is, "I could have been oriented differently *vis-a-vis* objects in my environment." This way, the statement can be analyzed without making cross-world comparisons of orientation (De Clercq, p. 558).<sup>9</sup>

As a result, we are able to regard the change in an enantiomorph's orientation (that occurs from mirror-reversal) as an aesthetically significant operation. Therefore, under the supposition that a change in an object's left/right orientation is aesthetically irrelevant (because it is observer-dependent), since we are able to express this change in a manner that is not dependent on an observer, a change in orientation (i.e. indeterminate orientation) still might matter aesthetically.

#### **d. A Possible Response to De Clercq's Argument**

Though De Clercq's argument for the aesthetic significance of changes in orientation seems to succeed in refuting Sorensen's claim by eliminating observer-dependence from orientation, it is open to the objection that it does not conclusively show that an object's orientation is intrinsic to it. If it had, then it would serve to directly refute Sorensen's entire account on the basis that orientation is *not* an extrinsic property. Rather, De Clercq's argument merely discounts Sorensen's claim—that extrinsic properties are irrelevant—since the supervenience principle Sorensen uses is controversial.

Though this rebuttal is legitimate, I find De Clercq's argument for non observer-dependent orientation to be equally, if not more, substantial than an argument that would make orientation out to be an intrinsic property. This is because there are certain observer-dependent extrinsic factors, such as physiology and culture, that are undeniably influential to an observer's preference to either *left-to-right* or *right-to-left* oriented objects; which make an attack on Sorensen's argument for the irrelevance of extrinsic properties somewhat trivial.

To see this, I will consider the arguments in Jarrold Levinson's article "Aesthetic Uniqueness." Levinson, like Sorensen, argues against the aesthetic relevance of the change in an object's left/right orientation that results from mirror-reversal.<sup>10</sup> Levinson notes that there are three factors which could be appealed to as sources of aesthetic divergence between an object and its flop. The first factor is human physiology, i.e., the possibility of physiological asymmetries in the human perceptual apparatus due to the left-brain/right-brain dichotomy. The second factor is widespread cultural associations. For instance, persons whose language writes from right-to-left (e.g., Hebrew users) may prefer right-to-left oriented objects from an aesthetic perspective. The third factor is symbolic functioning, i.e., right/left may be established as a symbolic element in certain regional traditions of art (Levinson 1980; p. 442). Though Levinson argues that mirror-reversal is an aesthetically irrelevant operation, he admits that these factors might make for aesthetic differences between paintings that were identical except for mirror-reversal or inversion (Personal conversation; December 3, 2012).

Sorensen alludes to such factors as well, but discounts them on the grounds that they are extrinsic factors and are therefore irrelevant. Rather than impose their irrelevance based on a controversial principle, he

makes a much stronger case to refute them individually. For this reason, it seems futile to refute the principle upon which Sorensen's argument is based when the principle itself is vacuously contentious. Therefore, the method by which De Clercq refutes Sorensen's claim (i.e., arguing that orientation need not be observer-dependent) is more substantial than counter-arguing the notion that orientation is an extrinsic property. Thus, De Clercq's account succeeds to prove that mirror reversal can be an aesthetically significant operation, on the grounds that a change in an object's orientation can bring out an aesthetic change even if a change in its left/right orientation does not.

**e. Sorensen's "Numerical Identity" Argument**

De Clercq refutes Sorensen's third argument explained above (see §1-b), concerning the numerical identity of objects and their flops. To reiterate, the argument goes as follows:

*"An object that is mirror-reversed is still, numerically, the same object; if mirror-reversal were to cause aesthetic changes, then one and the same object would both have and not have certain aesthetic properties; hence, a contradiction results from the assumption that mirror-reversal causes aesthetic changes"* (De Clercq; p. 559).

The argument is simple, but unconvincing. De Clercq provides a formally equivalent argument concerning changed objects. He explains that if Sorensen's argument were sound, then it can also be concluded that a changed object does not have different properties. De Clercq's counter-example goes as follows:

*"Each object that has undergone change is numerically the same as the object before the change; hence, we can derive a contradiction from the supposition that an object is different after a change, for an object that has changed would be different from itself"* (De Clercq, p. 560).

An argument of this form has gained popularity among some philosophers who feel that there is a tension between the notion of qualitative change and Leibniz's Law (Lowe, 2002).<sup>11</sup> In particular, it is used to support the perdurance view of persistence that objects persist by having different temporal parts at different times. However, according to De Clercq, if a perdurantist view is adopted, then Sorensen's argument quoted above (see §1-b) no longer makes sense. This is because one of the basic premises of perdurance theory entails that the object before the reversal and the object after the reversal are numerically identical; that is, they are not intended to be different temporal parts of some four-dimensional object (De Clercq; p. 560). De Clercq concludes that "if [Sorensen's argument] works in the case of orientation, then it should also work in the case of change; but if it works in the case of change, then it cannot work in the case of orientation" (De Clercq; p. 560). Thus, Sorensen's argument for the aesthetic irrelevance of orientation works against itself.

### **3. TWO DISTINCTIONS**

In this final section, I will address two distinctions that should be taken into consideration by any philosopher writing on the subject of mirror reversal. The first pertains to the subject-matter that one evaluates; specifically, whether one is investigating the nature of art objects or the nature of ordinary objects. The second distinction pertains to the process of mirror-reversal that one bases his account upon. I find the importance of these two distinctions to be particularly relevant to note, because both Sorensen and De

Clercq poorly establish their positions regarding each of these distinctions in their essays. Moreover, the validity of some of their arguments seems to depend on establishing such footing.

**a. Distinction 1: 'Art Objects' versus 'Ordinary Objects'**

An object may differ aesthetically from its flop depending on whether the object in question is an ordinary object or a work of art. Levinson explains that there is a generally accepted notion that every art object has a unique aesthetic merit in virtue of being a work of art:

*"Artworks are often accorded a special status or respect on the grounds that each is aesthetically unique. It is something of an unscrutinized dogma that every artwork will not fail to be one of a kind aesthetically, that every artwork is guaranteed to have something different and unduplicated to offer us as a work of art"* (Levinson 1980; p. 437).

If this is true, mirror-reversed art-objects may differ aesthetically from their originals on the grounds that each individual work of art *qua* art is aesthetically unique.

Sorensen uses several examples of mirror reversal in his essay; some of which pertain to art objects and others to ordinary objects. On the account that there is an aesthetic uniqueness about every individual work of art, Sorensen seems to defeat his own argument in his examples that address paintings, pictures, and music. De Clercq, however, does not use such examples in his essay. I am not proposing that De Clercq intentionally leaves examples of art-objects out of consideration for the reason that I provided; nonetheless, in doing so, his essay cannot be deemed illegitimate by this potential shortcoming.

### **b. Distinction 2: Modes of Mirror Reversal**

As I described in my introduction, there are two modes of mirror reversal.<sup>11</sup> The first entails two distinct duplicates, of which, one is the mirror-reversal of the other. The second mode entails an object going through a process of mirror-reversal, in which case it becomes its own flop. The former results in two objects existing at the same time, and the latter results in one and the same object, where the original and its flop exist at different times.

When writing on the aesthetics of mirror reversal, it is important that one establishes which mode of mirror reversal one is referring to; for certain changes that result from mirror-reversal may or may not be aesthetically significant depending on which mode one uses. For instance, a change in an object's left/right orientation resulting from mirror-reversal can be determined when the two distinct objects are present before an observer; whereas, in the case of an object that becomes its own flop, the original object's left/right orientation is indeterminate after the reversal process has occurred. Rather, it may only be inferred that the original was oriented exactly opposite the way it is oriented now (strictly speaking). If we consider this problem when evaluating Sorensen and De Clercq's opposing arguments regarding orientation, we can see how it is imperative that each author establish under which mode of mirror reversal he bases his argument.

The appeals that Sorensen and De Clercq make to possible worlds in support of their respective arguments are both based upon the mode of mirror reversal where an object turns into its own flop. In De Clercq's example of cross-world comparisons, if he were to refer to enantiomorphous objects using the other mode of mirror reversal--that they are two distinct objects where one is the mirror-reversal of the other--his argument still turns out to

be valid. Using either mode, whether we attempt to compare the orientations of the different temporal parts of one enantiomorph and those of its twin in a possible world or the orientations of two enantiomorphs and those of their respective twins, making cross-world comparisons of their orientations will be impossible. What this means is that De Clercq's argument works no matter which mode of mirror reversal is used.

In contrast, Sorensen's argument against the aesthetic relevance of changes in orientation, which result from a universal reversal, becomes invalid under the mode of mirror reversal in which an object is duplicated. Flipping an entire universe in this way would result in two distinct universes where one is the mirror-reversal of the other. However, his argument requires a scenario in which the original universe and its flop are numerically identical so that he is able to derive, from such cases, the contradiction that one and the same object may both have and not have certain aesthetic properties. If they are numerically distinct, then this contradiction is irrelevant. Therefore, whatever the shortcomings of Sorensen's argument,<sup>13</sup> it is only applicable under one mode of mirror reversal, but not both.

Thus, De Clercq provides a stronger account for the relevance of orientation to an object's aesthetic divergence from its flop than Sorensen provides for the irrelevance of orientation. Since De Clercq's argument maintains its validity under any mode of mirror reversal that may be applied, it provides an overall more comprehensive account than Sorensen's opposing claim provides.

**CONCLUSION**

This paper has treated mirror-reversal as a dimensional phenomenon with certain metaphysical implications. It asked, "Can an object differ aesthetically from its mirror-reversed twin?" In the introduction, I offered the accounts of Roy Sorensen and Rafael De Clercq, who maintain opposing answers to the posed question.

Whether a change in an object's orientation is an aesthetically significant operation presented itself to be the most contentious inquiry amongst the two authors. This was no surprise to me, as the greatest perceptible difference (if not the only perceptible difference) between an object and its flop is that their orientations are exactly opposite. For this reason, I focused my considerations in this paper on arguments for and against the aesthetic relevance of orientation.

In §I, I explained how Sorensen's argument relies on the principle that an object's aesthetic properties supervene on its intrinsic properties. I then showed how his view rules out changes in an object's left/right orientation as aesthetically relevant.

In §II, I explained De Clercq's argument concerning *indeterminate* orientation; showing how a change in an object's orientation under mirror-reversal can be aesthetically relevant, even if an object's left/right orientation is not. I then refuted a possible response to this argument, which helped to substantiate my claim that De Clercq succeeds to refute Sorensen's account of orientation.

In §III, I proposed two pertinent distinctions about the topic of mirror reversal, which, when imposed on Sorensen's and De Clercq's arguments regarding orientation

## EX NIHILO

across possible worlds, revealed a significant shortcoming of Sorensen's account while De Clercq's account turned out to be impervious to any such flaws.

Through my analysis in sections 2 and 3, I have shown that De Clercq successfully refutes Sorensen's argument--that there is no aesthetic difference between an object and its flop. Moreover, I have drawn on situations in which the change in an object's orientation resulting from mirror reversal is an aesthetically significant operation. Thus, the answer to the question I originally posed is "Yes, an object can differ aesthetically from its mirror-reversed image."

## Notes

<sup>1</sup> It may be helpful to note a third process of mirror-reversal by which an object is flopped through a higher dimension (this will be relevant beginning in §1-b of this paper). This sort of mirror-reversal is relevant in cases of "universal reversals" (See Sorensen, R., *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 100, No. 2 (Aug., 2000), p. 186), which is when an entire universe may be mirror-reversed by a higher-dimensional observer flipping a lower-dimensional universe through the higher dimension. A case of this would be if I were to flip a two-dimensional being, who lives on a strip of paper in my three-dimensional world, by twisting his world into a möbius strip. As the two-dimensional being travels along the strip of paper, he eventually reverses himself through the third-dimension without knowing that anything has changed. In theory, four-dimensional beings may do the same to us without us noticing any change.

<sup>2</sup> See Sorensen, R., *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 100, No. 2 (Aug., 2000), pp.175-191.

<sup>3</sup> See De Clercq, R., *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 132, No. 3 (Feb., 2007), pp. 553-563.

<sup>4</sup> Though Sorensen says an object's 'left/right orientation' is aesthetically irrelevant, he does not really say that an object's 'orientation' is aesthetically irrelevant. What I mean is that he rules out orientation as being 'observer-dependent', but such only pertains to noting whether an object is *left-oriented* or *right-oriented*, not whether it has *some orientation* or *another*. De Clercq refers to the former as 'determinate' orientation, and the latter as 'indeterminate' orientation. (See De Clercq, R., *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 132, No. 3 (Feb., 2007), p. 555) This distinction becomes more relevant in §2-b of my essay when I introduce De Clercq's cross-world analysis.

<sup>5</sup> This is under the supposition that an object and its flop are numerically identical; where the original is one temporal part and the flop is another.

<sup>6</sup> There is no need to challenge De Clercq's decision to do this because doing so is not problematic for the validity of his argument; he is

## EX NIHILO

allowed to narrow his focus to these sort of objects since, according to Sorensen, orientation is aesthetically irrelevant. Thus, if De Clercq yields a contradiction about enantiomorphous objects that follows from Sorensen's argument, then such will suffice for De Clercq to be able to refute Sorensen's argument in its entirety.

De Clercq does not make clear whether, for the purpose of his essay, enantiomorphous objects are different temporal parts of one and the same object or two distinct objects. (See De Clercq, R., *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 132, No. 3 (Feb., 2007), p. 554)

<sup>7</sup> I should clarify that such an object's orientation is not *literally* unchanged. Rather, it *appears* unchanged to an observer.

<sup>8</sup> Taken from De Clercq, R. (2007), p. 557

<sup>9</sup> This strategy, however, does not work for the statement "everything could have been oriented differently" (as Sorensen uses in his 'universal reversal' example). (De Clercq, p. 559) De Clercq comments: "Sorensen believes that the [aforementioned] statement is true. However, he provides only indirect support for this belief by countering arguments against the possibility of a universal reversal such as the argument that we can never come to know that a universal reversal has taken place". (De Clercq, p. 562)

<sup>10</sup> However, Levinson does not discount an object's extrinsic properties as aesthetically irrelevant.

<sup>11</sup> Leibniz's Law states that whatever is true of a thing is true of anything identical with that thing, since anything identical with that thing is that very thing itself. (Lowe, p. 23) Consider an object that undergoes a qualitative change in which the former and latter states of the object's noted quality are incompatible (e.g. a wire that goes from being straight to being bent). The tension that arises between such a case and Leibniz's Law, according to Perdurantists, is that it seems to be that the wire is both bent and straight (i.e. one and the same object seems to both have and not have a particular quality), which cannot be the case. (See Lowe, E. J. (2002). *A survey of metaphysics*. Oxford, NY: Oxford University Press., pp. 41-42)

<sup>12</sup> I actually mentioned three. However, two of which can be thought of as the same sort of operation; that is, both feature the concept of a *process* of mirror-reversal, whereby an object is turned into its own mirror image so that the original and its flop are one and the same object. The third mode results in two numerically distinct duplicates

where one is the mirror-reversal of the other. The distinction I am drawing here is whether mirror reversal results in one object or two.

<sup>13</sup> See §2-d above.

# Testimony, Evidence, and Epistemic Individualism<sup>1</sup>

*Kevin Dorst*

***Abstract:***

*Epistemology has traditionally focused on individual forms of evidence, such as memory and perception. However in recent years there has been a move toward analyzing the social aspects of cognition, including the important role that testimony plays in the justification of beliefs. In particular, John Hardwig argues that our pervasive reliance on testimony undermines the traditional individualistic approach to epistemic justification. In this paper I defend this traditional approach: not only can we make sense of testimony as a form of evidence, but in fact that our reliance on testimony is perfectly consistent with an individualistic approach to justification.*

Many of us feel the intuitive pull of an evidentialist, individualistic account of epistemic justification (I will refer to this type of account as EIJ). Roughly, this account is based upon the thought that the status of an epistemic agent's beliefs as rational and justified (or not) should be held accountable to all and only the facts within that agent's 'epistemic reach,' so to speak. When evaluating agents for these kinds of epistemic merits, we cannot

---

<sup>1</sup> I would like to thank Professor Roy Sorensen and Chris Dorst for helpful comments and discussion on earlier drafts of this paper.

simply expect them to arrive at the truth. The best we can hope for is the fulfillment of certain structural principles among their beliefs. The motto is ‘Think (rationally) for oneself.’ While intuitive, EIJ appears to be in tension with the fact that we must accommodate the possibility of *testimony* as a source of epistemic justification. Given the large number of our beliefs that are based upon the testimony of others, failure to account for testimonial justification will lead us to widespread and unacceptable skepticism. But accepting a proposition based upon testimony seems to be a paradigm case of *not* thinking for (or justifying) oneself, since it ultimately relies on trust in the authority of another agent. Thus it may appear that we must abandon EIJ. However, I will argue that this conclusion is unjustified: accounting for the justificatory force of testimony does not raise any *principled* barriers to upholding EIJ.

John Hardwig is one of the most vocal proponents of incorporating testimonial justification into our epistemologies and (hence) rejecting EIJ, so I will center my discussion around his arguments. First I will consider four conceptual worries Hardwig puts forward against understanding testimony as ‘evidence,’ and show how the familiar notion of memory-based evidence furnishes us with a response.<sup>2</sup> After arguing that there are no conceptual difficulties with the notion of ‘testimonial evidence,’ I then

---

<sup>2</sup> For a rejection of conceiving of testimony as ‘evidence’ based on different grounds see Goldberg (2006). Space precludes me from addressing the worries presented in that paper, but I think that they can be dealt with by clearing up that evidentialism is a thesis only about epistemic *justification*, not epistemic evaluation more broadly. See Feldman and Conee (1985): 28.

put forward a positive argument aimed to show that this notion is consistent with EIJ.<sup>3</sup>

### **I. Hardwig and Testimonial Evidence**

Hardwig (1985) argues that testimony that  $p$  provides a type of “reason [for belief] that does *not* constitute evidence for the truth of  $p$ ” (337). If correct, this conception of testimony would undercut EIJ by disconnecting an agent’s reasons for belief from her evidence, hampering her ability to ‘think for herself.’ Hardwig’s argument proceeds as follows. He first suggests a conception of evidence in which “‘evidence’ [is] defined roughly as anything that counts toward establishing the truth of  $p$ ” (336). Using this conception (which I accept), he aims to show that testimony does not count in favor of  $p$  in the way necessary to qualify as evidence.

Consider a case in which  $A$  has direct (non-testimonial) evidence that  $p$ , and in which  $B$  comes to justifiably believe that  $p$  due to  $A$ ’s testimony that  $p$  (where  $B$  has no non-testimonial evidence that  $p$ ).<sup>4</sup> Given our understanding of evidence as something that counts in

---

<sup>3</sup> Much of the debate over the epistemology of testimony centers on whether we should be *reductionists* or *anti-reductionists* about its justificatory force. The former camp claims that this force comes from testimony’s connection to other sources of knowledge, such as perception and memory, while the latter claims that there is always *prima facie* justification for accepting testimony (see Burge 1993 for a plausible example of the latter). Though it is an interesting topic, in this paper I will remain neutral between these two interpretations of testimony, as my argument works irrespective of how we decide this issue.

<sup>4</sup> There are certainly standards that must be met for an instance of testimony to succeed in providing justification to the receiver of said testimony (for instance, the receiver must not have evidence that the testifier is a pathological liar). In my discussion I will generally assume that these standards are met.

favor of (the truth of)  $p$ , Hardwig puts forward the following four considerations with the aim of showing that testimony does not function this way.

(1) He claims “although  $A$ ’s evidence counts toward establishing the truth of  $p$ , the case for  $p$  is not stronger after  $B$  discovers [via testimony] that  $A$  has this evidence...” (337). The thought is simply that since evidence counts in favor of  $p$  and  $B$ ’s testimony-based reasons for believing  $p$  do *not* count in favor of  $p$ ,  $A$ ’s testimony cannot be evidence. (2) Hardwig then points out that “the chain of appeals to authority must end somewhere” in order for there to be any support for  $p$  (337). So since an ungrounded chain of testimony cannot count in favor of the truth of  $p$ , and evidence *does* count in favor of  $p$ , testimony cannot be evidence. (3) Intuitively, “evidence that  $p$  counts against evidence that not  $p$ ” (337, fn 1). Now suppose we have our same situation with  $A$  and  $B$ , but, in addition,  $C$  has non-testimonial evidence that  $\neg p$ . When  $B$  comes to believe that  $p$  based on  $A$ ’s testimony, this fact does not make the case against  $C$  (and his belief that  $\neg p$ ) any stronger. And since  $B$ ’s testimony-based reasons for believing  $p$  do *not* count against  $C$ ’s evidence-based reasons for believing  $\neg p$ ,  $B$ ’s reasons cannot be evidence. (4) There are cases in which  $B$  comes to believe that  $p$  based upon  $A$ ’s testimony, even when there is, in fact, *no* evidence that  $p$  (337, fn 1). This may involve, for instance,  $A$  convincing  $B$  that he has evidence for  $p$  even though he has none. In this case nothing counts in favor of  $p$ , thus  $B$ ’s testimony-based reasons for belief cannot be evidence for  $p$ .

Though these objections have some surface plausibility, I submit that they cannot be correct. This is because exactly analogous objections succeed in showing that *memory* cannot be understood as providing evidence-based

justification. But since (as it is traditionally conceived) memory *does* provide evidence-based justification, these objections must be mistaken. Consider the following scenario that is analogous to the testimonial exchange of *A* and *B*.  $D_1$  (i.e. agent *D* at time 1) has direct<sup>5</sup> evidence that *p*, and  $D_2$  (i.e. *D* at time 2) remembers that  $D_1$  believed that *p* based on direct evidence, though  $D_2$  no longer remembers the specific evidence that justifies this belief. To make the example concrete, perhaps  $D_1$  has just completed a mathematical proof that *p*, and  $D_2$  remembers having completed such a proof but does not remember the specific steps required to do so. I take it as evident that in this example  $D_2$  has memory-based *evidence* that *p*.

Now return to Hardwig's objections. (1) He first claims that the testimony-based reasons *B* has for believing that *p* cannot make the case for *p* stronger (since they are parasitic on *A*'s evidence). Analogously, the justification provided by  $D_2$ 's memory of  $D_1$ 's evidence-based belief is also parasitic on  $D_1$ 's original evidence (i.e. his proof). (2) Second, Hardwig points out that the chain of appeals to testimonial authority must bottom out in direct evidence if it is to count in favor of *p*. But clearly the chain of appeals to memory must also bottom out in direct evidence. Remembering having a memory that one had a belief that *p* only counts in favor of *p* if the original belief was based on direct evidence counting in favor of *p*. (3) Third, Hardwig claims that testimony that *p* fails to *count against* evidence that  $\neg p$ , and therefore testimony cannot itself be evidence. Analogously, suppose that in addition to the proof (that *p*) possessed by  $D_1$ , *C* has a (apparent) proof that  $\neg p$ . Clearly  $D_2$ 's *memory* of  $D_1$ 's proof does not count further against *C*'s proof than  $D_1$ 's original proof does. Thus  $D_2$ 's memory-based reasons that *p* do not count against *C*'s

---

<sup>5</sup> By *direct* evidence, I mean non-testimonial, non-memory-based evidence, such as perceptual evidence.

evidence-based reasons that  $p$ . (4) Finally, Hardwig points out that  $B$  may believe that  $p$  based on  $A$ 's testimony even when nothing at all counts in favor of  $p$ . But of course we can construct an analogous case for memory. For example suppose  $D_2$ 's memory is mistaken and  $D_1$  never actually had a proof of  $p$ , or there was an invalid step in said proof.

We can now see that certain cases of memory-based belief are precisely analogous to the cases of testimony-based belief presented by Hardwig. So if his objections (1)-(4) show that testimony does not provide evidence, they likewise show that *memory* does not provide evidence. But since memory *does* provide evidence, something must have gone awry with Hardwig's objections.<sup>6</sup>

To assess what has gone wrong, I want to tentatively suggest two refinements that these objections seem to have overlooked. These are that (i) sometimes evidence only counts in favor of a proposition from a certain epistemic perspective, and (ii) testimony (like memory) does not generate evidence, but merely transfers it from one epistemic perspective to another. The thought behind (i) is that not only can we ask, "What evidence is there that  $p$ ?" but also, "What evidence does  $A$  have that  $p$ ?" or "What evidence that  $p$  is there *from  $A$ 's perspective*?" Certainly there is a sense in which evidence can be objective, such as when some facts make other facts

---

<sup>6</sup> Hardwig may now object that a memory of an evidence-based belief (as in the cases I have been using) does *not* provide evidence. But this suggestion would be implausible. Certainly we must admit that *some* forms of memory provide evidence for our beliefs if we are not to be complete skeptics (and not to be quite uncharitable to evidentialists). And must we truly remember the direct evidence we had in order for a memory to provide us with evidence? Must I remember every step in a proof in order for memory of that proof to provide me with evidence of its conclusion? If we are to avoid significant skepticism, we must answer in the negative to these questions.

objectively more probable. But there also appears to be a sense in which evidence can be subjective, such as when an instance of perception, memory, or testimony leads an agent to (justifiably) increase her subjective probability that  $p$ . And this may happen even when the objective probability is not increased, as in the example of testimony or memory.

The thought behind (ii) is straightforward: testimony and memory provide non-grounding forms of evidence. Such evidence is ultimately parasitic on the direct evidence (e.g. perception) on which it is based, and simply serves to transfer the justificatory force of the direct evidence to a new perspective. The general sentiment behind this view (sometimes called the *transmission view*) is widespread, with good reason.<sup>7</sup> What else is testimony (or memory) presumed to do, other than to bring justification from one perspective to another? Lackey (2011; 83-6) offers interesting arguments against specific formulations of the transmission view, but these concerns are not relevant here. Her arguments are directed against proposals for necessary and sufficient conditions for transferring knowledge via testimony, but the claim from (ii) does not (and need not) go into these specifics to explain away Hardwig's objections.<sup>8</sup>

It should be fairly clear how (i) and (ii) allow us to maintain the intuition that evidence is something that counts in favor of  $p$ , while also admitting that memory and testimony are forms of evidence. The basic thought is that these mechanisms serve to count in favor of  $p$  (or count against  $\neg p$ ) from a given agent's perspective, and that this

---

<sup>7</sup> See Lackey 2011, fn. 34 and 35 for a substantial list of references.

<sup>8</sup> For a specific and defensible formulation of the transmission view see Burge (1993: 485-7). I believe his formulation avoids the brunt of Lackey's objections.

happens via a transfer of information across epistemic perspectives. That being said, I do not want to push the specifics of (i) and (ii) too strongly, as more precise and careful formulations could surely be developed. My basic point is that due to the strong analogies between memory and testimony, we should be able to model our understanding of testimonial evidence off of our understanding of memory-based evidence, whatever the specifics of the latter turn out to be.<sup>9</sup>

## **II. Utilizing Information and Drawing Arbitrary Lines**

Having laid out these arguments and refinements, I conclude that Hardwig's considerations do not pose any conceptual difficulties for conceiving of testimony as "evidence." But Hardwig has a very clear response to this fact. He can allow me to call testimony "evidence," but then claim that all I've done is pull a verbal maneuver.<sup>10</sup> Perhaps we can incorporate the justificatory force of testimony while still being "evidentialists" in name, but I have done nothing to capture the *individualistic* intuitions that underlie EIJ. Whether we call testimony "evidence" or not, accepting that *p* based upon testimony still ultimately relies on trust in authority. So the true epistemic individualist should now simply assert, "claims to know or to rationally believe must rest on *nontestimonial* evidence" (Hardwig 1988: 311, my emphasis). This is because only

---

<sup>9</sup> Admittedly, there are *some* disanalogies between memory and testimony, such as the fact that memory appears to be a more basic and integral faculty for functioning as an epistemic agent. Though I cannot put forward a fully adequate case for it here, I believe the structural functions of these faculties are obviously similar enough to warrant my claim in the text.

<sup>10</sup> He makes this claim explicitly against a similar line of thought in Hardwig (1988; 311). Goldman (2011; 15, fn 4) seems to agree that the debate over whether we should conceive of testimony as "evidence" is not overly important, as it rests largely on verbal stipulation.

“nontestimonial evidence... establishes a sufficient link between *B*'s beliefs and the real world” (313) to fulfill the aim (set out by EIJ) of thinking for oneself. Thus the fact that we must incorporate testimony into our epistemology still provides us with a principled objection to EIJ.

In response to this line of argument, I want to object to the claim that testimony poses a unique or principled problem to the epistemic individualism of EIJ. I will attempt to show that we can rely on testimony while still thinking for ourselves. I begin by noting that perception, reasoning, and memory are obviously all individualistic, evidential sources of justification based upon utilizing information presented to an epistemic agent. With this in mind, I contend that there is *no non-arbitrary line* that makes a *principled* distinction between the information utilized by these sources, and the information utilized in considering the testimony of others. There are two ways I will attempt to illustrate this. First, suppose a scary movie is playing, and consider the following series of evidence statements:

- 1) I see that the monster jumped out (and so have individualistic evidence through perception).
- 2) I see through glass that the monster jumped out.
- 3) I see through binoculars that the monster jumped out.
- 4) I see through a digital camera display that the monster jumped out.
- ...
- m) I discern through observing your reaction that the monster jumped out.
- ...
- n) I am told via testimony that the monster jumped out.

There is a very clear continuity between (1) and (n), which can also be brought out in another way. To say that *B* is (to varying degrees) *transparent* to *A* means that *A* has (varying degrees of) knowledge of *B*'s internal mental processes and states.<sup>11</sup> Given this terminology, if *B* is *completely* transparent to *A*, and *B* says that *p* (supposing this testimony is based on good evidence), then *A* has direct, individualistic evidence that *p*, since he simply has the same evidence as *B*. Now consider:

- 1) *B* is 100% transparent to *A*, and *B* testifies that *p* (so *A* has individualistic evidence).
- 2) *B* is 99% transparent to *A*, and *B* testifies that *p*.
- 3) *B* is 98% transparent to *A*, and *B* testifies that *p*.
- ...
- n) *B* is as transparent to *A* as most people are to most people (that is, *A* has some understanding of the mental processes of *B* due to common sense and standard psychological and social knowledge of humans), and *B* testifies that *p*.

The point I am driving at is that testimony can be understood simply as a complex way of obtaining information through a medium. The only difference between testimony and standard forms of individualistic evidence is that instead of garnering information through (e.g.) binoculars, one is garnering information through a person. While there may be differences in degree between the media this information comes through, there is no principled difference in kind.<sup>12</sup>

---

<sup>11</sup> I draw the notion of agential transparency from Parfit (1984): 18-19.

<sup>12</sup> Of course, this is not to say there are no differences in kind *in any respect* between binoculars and persons. I merely mean that they both can and do function as information transmitters in a similar manner.

Now some may contend that there *is* a difference in kind due to the fact that information received through testimony is intruded upon by some alien agency. However, the above progression based on *B*'s increasing transparency hints at my response to this objection (since *B*'s agency is not called into question, even when he is 100% transparent).

Regardless of one's opinion on the metaphysics of 'free will,' clearly the behavior of individuals *does* follow predictable causal and psychological patterns (whether or not it *must* follow such patterns): the entire fields of psychology and neuroscience are based on this fact. And the existence of these patterns leads people to function as information transmitters in predictable ways that are analogous to the predictable effects of binoculars.

This understanding of testimony already has some acceptance in the literature, though (to my knowledge) its implications for EIJ have not yet been drawn out. Thus Kelly (2011) claims, "In general, the way the judgments of some other mind come to play the role of evidence does not differ from the way other states of the world do" (197).<sup>13</sup> Likewise, in a different context Noë (2004) cites a thought experiment from Sanford (1997) about "Chris-the-amazing-human-hearing-aid." Chris is a person who "has superhuman powers of mimicry" that allow her to play the role of a hearing aid to a listener by replicating all the surrounding sounds (Noë 173). Though Sanford does not think Chris provides a form of perception, Noë argues that this is because Chris does not fulfill the correct subtle counterfactuals (such as changing pitch when the listener moves his head). He contends that after suitable alterations are put in place to make Chris fulfill these counterfactuals, her testimony should be understood as providing a form of

---

<sup>13</sup> Similarly, in their famous defense of evidentialism Feldman and Conee take it as obvious that testimony functions just like other forms of evidence (see their 1985: 29).

prosthetic hearing (174). My intuitions agree with Noë's, and I believe the case helps illustrate the continuity between testimony and other forms of evidence.

Now there is a plausible objection to my attempt to reconcile EIJ with the justificatory force of testimony, which goes as follows. There may be no sharp, principled *line* separating the information of testimony from that of perception. But the amount of direct, individually accessible evidence that a piece of information provides to an agent decreases as the mechanism that provides it becomes (a) less understood and (b) less consistent and reliable. Further, the processes governing a testifier's actions are not well understood by any given agent (thus the role of *trust* in testimony), and the reliability of testimony is incredibly variable and situation-dependent. Thus the amount of individual evidence provided by testimony is quite minimal and cannot support the degree of justificatory force that is in fact provided by testimony.

In response to claim (a), I simply deny that the degree of evidence provided by a piece of information depends in this way on how well the mechanism is understood. After all, very few people have any knowledge of how their perceptual or memory systems work, and yet we all agree that these mechanisms provide some of the strongest forms of evidence. Further, it is worth noting that our justification in accepting perceptual evidence ultimately rests on *trust* in these mechanisms, just like testimony does.<sup>14</sup>

---

<sup>14</sup> As I mentioned above, I want to be neutral where exactly this trust comes from (i.e. whether it is based on a reductionist or non-reductionist account). But clearly for both testimony and perception the trust can be overridden by situational factors. If you are blindfolded you should not trust your visual system; if you are talking to a known liar you should not trust his testimony.

In response to (b), I agree that situation-dependence and lack of reliability can undermine the justificatory force of a type of evidence. But I claim that this difference in degree actually matches our intuitions about testimony. Certainly it is the case that perception is not *absolutely* reliable, and its reliability *is* situation-dependent.<sup>15</sup> On the other hand, simply note that if epistemologists truly thought that testimony was utterly riddled with problems of *unreliability*, then they would not spend their time trying to incorporate the justificatory force of testimony into their theories. Granted, testimony is in general (though not always) less reliable than perception. But this matches our intuitions that *seeing* that *p* is (usually) a much stronger form of evidence than being told that *p*. Thus it turns out that the imperfect reliability of testimony actually fits well with our intuitions about its justificatory force.

Having considered this line of objection, I maintain my contention that there is no principled distinction between the information provided by testimony and that provided by standard forms of individualistic evidence. Hence testimony does not pose a unique or principled problem for maintaining EIJ.

### **III. Individualism without Skepticism**

Hardwig has succeeded in demonstrating that any acceptable epistemology must incorporate the justificatory force of testimony. However, I have argued that his bold claims that doing so “undermines the methodological individualism that is implicit in most epistemology” and calls for “some very basic changes in our epistemologies”

---

<sup>15</sup> For example, if I have recently taken a hallucinogen, then my visual experiences do not provide much justificatory force.

are unjustified (1985: 349). His argument goes through only if testimony provides us with a unique form of justification, one that cannot be assessed individualistically and evidentially. But I have attempted to show (1) that there are no conceptual problems with understanding testimony as a form of evidence, and (2) that there is no principled distinction between the information provided by standard forms of individualistic evidence and that provided by testimony. Granted, the epistemology of testimony is a very complex and puzzle-ridden issue. But it does not pose a uniquely dangerous threat to our evidentialist, individualistic intuitions about epistemology.

# The Deep Self View and Wolf's Sanity Condition: An Unnecessary Addendum

*Sabeen Ahmed*

In her article “Sanity and the Metaphysics of Responsibility,” Susan Wolf expounds upon previous philosophers’ attempts to offer an adequate compatibilist theory of responsible agency. The theory, which Wolf dubs the “Deep Self View,” states that an agent is responsible for her action if it is controlled by her will, insofar as the will is governed by the agent’s deeper self. Nevertheless, Wolf does not think that the Deep Self View (hereafter referred to as the “DSV”) on its own is sufficient to ensure our status as *morally* responsible agents. According to Wolf, though an agent’s action may be governed by her deeper self, it may be the case that her deeper self is incapable of distinguishing between right and wrong. In such a circumstance, the agent should not be held blameworthy for her actions, as she lacks the appropriate capacity. Wolf calls this capacity the “sanity condition,” and posits it as a condition on top of the DSV that ensures an agent’s full moral responsibility.

In this paper, I address Wolf’s sanity condition and argue that it is insufficient to account for the relief from moral responsibility that Wolf aims to establish. I will begin by offering a more comprehensive evaluation of the DSV, and then move on to examine each side of the sanity condition: the cognitive side and the normative side. In

doing so, I will show how the sanity condition fails on both accounts. On the one hand, the cognitive sanity condition is implicitly encompassed in the DSV, and although it does ensure our moral responsibility, it does not do so by providing the so-called missing link that Wolf thinks the DSV lacks. On the other hand, the normative sanity condition is an impossible condition to lack if an agent meets the DSV; meaning, the capacity to distinguish between normatively right and wrong actions cannot be lacking in an agent whose actions are ultimately governed by her deeper self.

Before moving into my argument, I want to note that I will not be arguing for or against the compatibilist view of free agency, nor will I address Wolf's defenses of the DSV against metaphysical objections. Rather, I will assume the compatibilist view for the sake of taking for granted the possibility for moral responsibility as well as the validity of the DSV, and will focus my argument on the plausibility of Wolf's condition specifically. Furthermore, I will be focusing on the conditions necessary for *moral* responsibility rather than free agency in general, as Wolf's inclusion of the sanity condition is meant to account for cases in which an agent may be acting freely, but still relieved of blame.<sup>1</sup> Wolf also assumes a realist picture of morality, in that a moral agent is able to distinguish between normatively right and wrong actions. I will take this for granted, and thus not address objections stemming from moral subjectivism or moral nihilism. Nevertheless, I do find it necessary to expound on the DSV more explicitly before moving into my critique of Wolf's sanity condition.

---

<sup>1</sup> The normative implication of 'blame' as used by Wolf is indicative of moral responsibility, and will be interpreted as such in my paper.

## I. The Deep Self View

For Susan Wolf, the Deep Self View is the notion that ensures our status as responsible agents: “responsible agents are those for whom it is not just the case that their actions are within the control of their wills, but also the case that their wills are within the control of their *selves* in some deeper sense.”<sup>2</sup> She examines the theories of free agency offered by Harry Frankfurt, Gary Watson, and Charles Taylor, and coins the DSV as the basic notion that underlies each philosopher’s theory: “[their] accounts may be understood as alternate developments of the intuition that in order to be responsible for one’s actions, one must be responsible for the self that performs these actions.”<sup>3</sup> This view is sufficient to explain why certain agents, such as victims of hypnosis or brainwashing, are not blameworthy for their actions, as their wills are not truly governed by their deeper selves. Furthermore, the DSV accounts for such agents as kleptomaniacs and unwilling addicts, whose wills are not controlled by forces physically external to the agent, but forces external to the agent’s *true* self. Nevertheless, it is not enough to say that a responsible agent is one whose actions are merely governed by her “deeper self,” as Wolf states. As illustrated by Watson and Frankfurt, there are significant characteristics of the deeper self that are worth noting. In discussing their views, I aim to shed light on the qualities of the Deeper Self that must be highlighted by the DSV as a whole, beyond the description of the view that Wolf offers.

Watson’s view of free agency is centered on an agent’s valuational system as the central quality that

---

<sup>2</sup> Wolf, Susan. “Sanity and the Metaphysics of Responsibility.” *Reason & Responsibility: Readings in Some Basic Problems of Philosophy*. Ed. Joel Feinberg and Russ Shafer-Landau. Belmont: Thomson Learning. 503. Print.

<sup>3</sup> Wolf, 502.

governs the agent's actions: "the valuational system of an agent is that set of considerations which, when combined with his factual beliefs (and probability estimates), yields judgments of the form: the thing for me to do in these circumstances, all things considered, is *a*."<sup>4</sup> The 'valuational system' can be taken as analogous to the 'deeper self,' as Wolf would suggest, though it is important to highlight Watson's mention of 'considerations' which drive the agent to act. These considerations, or the agent's "factual beliefs and probability estimates," when combined with her valuational system, are what enable the agent to make evaluative judgments for action. This is not to say that Watson's view of the deep self (the agent's valuational system) *includes* these "factual beliefs and probability estimates." However, it is not enough that the agent is acting out of her deeper self alone; the deeper self itself must be, to an extent, shaped by her (factual) beliefs about the world and her mental faculties of reasoning, sensation, inference, deduction, reflection, etc., in order to make appropriate probability estimates about the world.<sup>5</sup> These beliefs are what allow the agent to act in a manner that is compatible (and, in fact, governed) by her deeper self. Thus I hold that the capacities needed to acquire factual beliefs are a necessary facet of the agent's being able to make such judgments.

The same would apply in Frankfurt's theory, which states that a responsible agent is one who holds second-order desires, or desires about her desires. For Frankfurt, the agent must be able to control her will in some way:

---

<sup>4</sup> Watson, Gary. "Free Agency." *The Journal of Philosophy*. Vol. 27, No. 8. Journal of Philosophy, Inc. 1975. 215.

<sup>5</sup> These capacities are not explicitly mentioned by Watson. Rather, I assert that he would agree that the functional use of these capacities is implicit in an agent's ability to form factual beliefs and probability estimates about the world.

“[she] is free to will what [she] wants to will, or to have the will [she] wants.”<sup>6</sup> The free will, insofar as the agent’s formulation of second-order desires, is what differentiates responsible agents from non-human animals, infants, and wantons. However, unless the agent also had the capacities to reason and reflect, she would be unable to deliberate between her desires: she requires some sense of rationality that enables her to form second-order desires.<sup>7</sup> Otherwise, her actions would be moved merely by her brute desires or passions, and she in no way would be different from the wanton or animal: “the capacity for reflective self-evaluation [...] is manifested in the formation of second-order desires.”<sup>8</sup> Thus, in the same way that Watson acknowledges the necessity of (the capacity to have) factual beliefs and probability estimates about the world, Frankfurt acknowledges the necessity for an agent’s capacity for self-reflection.

Taking into consideration the above, I posit that a more comprehensive interpretation of the DSV is required, before introducing Wolf’s sanity condition. The DSV will hereafter be considered a condition of responsible agency, insofar as the agent’s actions are governed by her deeper self, which is itself the comprehensive set of beliefs, desires, and values the agent holds in virtue of her capacities of understanding, interpreting, reasoning, and

---

<sup>6</sup> Frankfurt, Harry G. “Freedom of the Will and the Concept of a Person.” *The Journal of Philosophy*. Vol. 63, No. 1. Journal of Philosophy, Inc. 1971. 15.

<sup>7</sup> By this I mean that the agent would need to have an accurate picture of the world in order to gauge which desires she wants to be ‘effective.’ For example, if she decides to stay in and study in favor of going to the cinema, she would need to deduce the consequences of each of these actions, as well as the implications of doing poorly on her test if she fails to study.

<sup>8</sup> Frankfurt, 7.

reflecting about the world. The DSV as stated does have an element of externality about it, in that the agent is in some way being controlled by the world. This is due to her necessity for holding factual beliefs about the way the world is, as well as her values and desires being shaped by the world in some way.<sup>9</sup>

Nevertheless, Wolf herself accepts this, and does not find it problematic for the view: “my deepest self will still be governed by something that must, logically, be external to myself altogether, [but] as long as we can revise ourselves, [and] reflect on what sorts of beings we are, [...] it is hard to find reason to complain.”<sup>10</sup> Furthermore, it accounts for the agent’s ability to understand and interpret the world in a way that allows her to make evaluative judgments about both the world and her values and desires. This is itself a unique characteristic that distinguishes rational agents from other animals, which both Frankfurt and Wolf assert: “the possession of a deep self that is effective in governing one’s actions is a sufficient distinction [between] [...] us [and] other members of the natural world.”<sup>11</sup> Thus, I hold that my comprehensive conception of the DSV is both consistent with Wolf’s interpretation of the view, as well as an elucidation that her predecessors would accept.

---

<sup>9</sup> By the latter I am referring to an agent’s values being affected by her religious, economic, and social upbringing, as well as her parents and peers.

<sup>10</sup> Wolf, 504-505.

<sup>11</sup> Wolf, 505.

## II. The Sanity Condition

Wolf's limited account of the DSV leads her to believe that the view is insufficient to ensure moral responsibility. She illustrates this by offering an example of an individual, JoJo, the leader of a small and undeveloped country, who acts out of his deeper self, but whose deeper self is incapable of determining which actions are right and which are wrong. Due to his very sheltered upbringing under the care of Jo the First, a cruel and sadistic dictator, JoJo's actions, including torturing citizens and sending people to prison on the basis of whim, are driven by desires that, Wolf argues, "are the desires he wholly *wants* to have."<sup>12</sup> Thus, although "from the inside, he feels as integrated, free, and responsible as we do," JoJo is, "at the deepest level, unable cognitively and normatively to recognize and appreciate the world for what it is."<sup>13</sup> Because of this, Wolf believes that JoJo is an individual who meets the DSV, but is still not morally accountable for his actions. To handle such cases as JoJo, Wolf adds the sanity condition to her conception of the DSV.

Wolf defines the sanity condition as "the minimally sufficient ability cognitively and normatively to recognize and appreciate the world for what it is."<sup>14</sup> The DSV coupled with the sanity condition is comprehensively sufficient for an agent to be morally responsible for her actions. Nevertheless, the sanity condition as stated is ambiguous. In order to fully examine whether or not Wolf's condition is necessary for moral responsibility, it is imperative to break down the condition into its constituent features: the cognitive aspect and the normative aspect.

---

<sup>12</sup> Ibid.

<sup>13</sup> Wolf, 506-507.

<sup>14</sup> Wolf, 506.

### III. The Cognitive Sanity Condition

The cognitive aspect of the sanity condition is the functionality of an agent's mental processes, and primarily in the context of Wolf's writing, the ability to be controlled by and interpret the world in the appropriate manner. Wolf's cognitive sanity condition, "the ability cognitively to recognize and appreciate the world for what it is," has dual implications.<sup>15</sup> It has to do not only with the internal workings of the agent's cognitive faculties, but also the external world, insofar as the agent's interpretation of the world needs to align with the actuality of the world to ensure that her cognitive faculties are 'sane.'

Accepting this description as such, it is clear to see how an agent who fails to meet the cognitive sanity condition should not be held morally accountable for her actions. Here I highlight agents who suffer from mental illnesses, which cause them to experience a distorted perception of reality and/or inability to correctly analyze or reflect on their interactions with themselves and with others.<sup>1617</sup> Essentially, such an agent is incapable of acting out of her 'deeper self,' since her deeper self itself is not exercising full or correct use of her cognitive faculties, or whose beliefs about the working of the world do not align with the actuality of the world. I hold that we do not assign moral responsibility to agents who fall under this category, as this inability to cognitively recognize right from wrong renders them incapable of understanding the moral

---

<sup>15</sup> Vis-à-vis Wolf's definition of the sanity condition.

<sup>16</sup> The former addresses the external aspect of the cognitive sanity condition and may include such individuals as schizophrenics, while the latter addresses the internal aspect of the cognitive sanity condition and may include such individuals as psychopaths.

<sup>17</sup> Perring, Christian. "Mental Illness." *Stanford Encyclopedia of Philosophy*. Spring 2010. <<http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=mental-illness/>>.

implications (or even, perhaps, the non-normative consequences) of their actions. Further, this assertion is validated by the justice system's employment of the insanity defense, which Wolf herself references: "criteria for sanity have historically been dominant in legal questions about responsibility."<sup>18</sup>

This is certainly not to say that an agent who suffers from a psychological disorder is *never* morally responsible for her actions. For example, an agent who is bipolar may, when taking the appropriate medication, be fully functioning. Nevertheless, during periods of manic highs, the same agent may suffer from extremely vivid aural or visual hallucinations which make her believe or feel obligated to perform an action that goes against her valuational system or true desires.<sup>19</sup> Further, she may, at this time, be unable to foresee what consequences may follow her actions, moral or otherwise. Thus, any actions she would have performed during this time would not be driven by her deeper self. It is not merely that the agent *does* not cognitively distinguish right from wrong in this case; it is that she *cannot*.

Further, for agents who suffer permanent cognitive disabilities, such as severely mentally disabled agents, I argue that we never consider them responsible agents, either morally or otherwise. Though their actions are their own, they do not enjoy the capacity for self-reflection or deliberation that allows them to determine the consequences of their actions, or the understanding and significance of choosing between actions. Consequently, it would be the case that their actions are more instinctual or 'brute' than truly intentional, and these actions would not

---

<sup>18</sup> Wolf, 506.

<sup>19</sup> Further, these hallucinations themselves are an indication that she is not perceiving the world accurately.

be governed by a ‘deeper self.’<sup>20</sup> Like some cases of agents with psychological disorders, agents suffering from severe cases of mental retardation would not be responsible agents, as they fail to meet the cognitive sanity condition; the only difference between the two is that actions performed by severely mentally disabled agents are *never* governed by a deeper self.

Nevertheless, although the cognitive sanity condition is a necessary condition for an agent to be responsible for her actions, I argue that it fails for Wolf on the grounds that it is already implicitly contained in the DSV. As discussed, the DSV requires that an agent’s deeper self govern her actions. The deeper self itself however, requires that the agent’s evaluative judgments for her actions stem from her factual beliefs and probability estimates about the external world, which can only be accomplished if she is both internally cognitively sane (has full functionality of her mental processes) and externally cognitively sane (is able to recognize the world for what it is). Thus, the DSV requires that the agent’s actions are governed by her deeper self, but this is only possible if her deeper self is already cognitively sane. Accepting this, the cognitive aspect of Wolf’s sanity condition is redundant.

#### **IV. The Normative Sanity Condition**

There is still another side to the sanity condition worth addressing: the normative sanity condition, or “the ability normatively to recognize and appreciate the world for what

---

<sup>20</sup> I argue that a philosopher such as Frankfurt would hold that such agents lack free will as a whole, since they would be incapable of forming second-order desires. Their actions would be equivalent to that of the wanton or infant, and thus Frankfurt would not consider them either free or responsible agents.

it is.”<sup>21</sup> This means that the agent is able to recognize which actions are normatively right and which are normatively wrong. Wolf does not believe that the DSV accounts for this capacity, and thus finds it a necessary condition that an agent must meet (assuming they have already met the ‘deep self’ criteria) to ensure her moral responsibility.<sup>22</sup>

As stated previously, I will not address objections from moral subjectivism and moral nihilism, and instead take for granted a moral realist picture. I hold that moral realism is what allows Wolf to distinguish between the cognitive and normative aspects of the sanity condition; unless there did actually exist moral facts, agents would have no need to be able to distinguish between moral claims that are true and those that are false: “moral claims do purport to report facts and are true if they get the facts right.”<sup>23</sup> Accepting, then, that there are moral facts and it is possible for agents to determine which moral claims are true, there must be some capacity an agent has that allows her to evaluate them. Otherwise, an agent’s acceptance of particular moral claims would be arbitrary. This capacity must then be a sort of mental process or faculty, and in turn, is a part of her cognitive make-up. Thus, in order to accept that the normative aspect of Wolf’s sanity condition exists at all, we must accept that agents have a certain cognitive capacity for moral reasoning.

I hold, then, that it is impossible to imagine a case in which an agent who acts out of her deeper self would lack the *capacity* for moral reasoning, unless a feature of

---

<sup>21</sup> Vis-à-vis Wolf’s definition of the sanity condition.

<sup>22</sup> She uses the case of JoJo as a means of justifying this claim.

<sup>23</sup> Sayre-McCord, Geoff. “Moral Realism.” *The Stanford Encyclopedia of Philosophy*. Summer 2011.

<<http://plato.stanford.edu/archives/sum2011/entries/moral-realism/>>.

their cognitive make-up were also impaired.<sup>24</sup> If this were the case, then that agent would fail to meet the cognitive sanity condition and, consequently, would fail to be acting out of her deeper self at all. Of course, this is not to say that an agent who meets the DSV but appears to be normatively insane does not have a severely *limited* moral understanding, but there certainly seems to be a significant difference between lacking moral understanding and lacking the *capacity* for moral understanding that must be addressed.

#### V. Ignorance vs. Insanity: The Case of JoJo, Revisited

For Wolf, the case of JoJo is meant to be an example of an agent who meets the DSV, but is unable to recognize normative rightness from wrongness, and thus is not to be considered a morally responsible agent. His actions are still fully his own, as they are a result of his deepest desires, but his inability to understand that his actions are wrong are an indication of his normative insanity: “in the cases of JoJo and the others, there are certain features of their characters that they *cannot* avoid even though these features are seriously mistaken, misguided, or bad” (emphasis mine).<sup>25</sup>

However, in a study conducted by David Faraci and David Shoemaker, an important alternative intuition has been suggested, that rather than being normatively *insane*, JoJo is normatively *ignorant*: “what JoJo lacks is knowledge of what *is* right and what *is* wrong [...] just because he does not know that expressions of ill will are

---

<sup>24</sup> In fact, I hold that an agent who lacks the capacity for moral reasoning is cognitively impaired for this very reason (since the capacity for moral reasoning is itself a cognitive capacity), if not due to other additional cognitive deficiencies.

<sup>25</sup> Wolf, 508.

immoral, that does not yet mean that he lacks the ability to know it.”<sup>2627</sup> This intuition also justifies why we view certain groups of agents as less morally blameworthy as others, given the social and moral constraints of their time: “this [...] explains why we give *less than full* responsibility to persons who [...] act in ways that are strongly encouraged by their societies – the slaveowners of the 1850s, the Nazis of the 1930s, and many male chauvinists of our fathers’ generation” (emphasis mine).<sup>28</sup> Nevertheless, in no sense do we withhold full moral responsibility from these agents, which Wolf herself seems to be admitting. Rather than total relief from moral responsibility, we are less inclined to hold these agents as *fully* blameworthy due to their severe moral *ignorance*. Thus, just because JoJo does not see the difference between right and wrong, this is not enough to say that he *cannot* do so at all, particularly since he does meet the cognitive sanity condition (since his actions are governed by his deeper self): “it is not necessarily that JoJo *cannot* see the basic demand; perhaps, instead, he simply *does not*.”<sup>29</sup>

This discussion is not meant to justify why agents cannot be normatively insane; we have already accepted that if this is the case, it is due to some cognitive impairment. Nevertheless, Faraci and Shoemaker’s critique of Wolf’s JoJo example is significant in that it brings up an

---

<sup>26</sup> Faraci, David and David Shoemaker. “Insanity, Deep Selves, and Moral Responsibility: The Case of JoJo.” *Review of Philosophy and Psychology*. 2010. 14.

<sup>27</sup> Faraci and Shoemaker reached this conclusion by assessing students’ intuitions regarding the JoJo case. The students were given three different scenarios of JoJo (one of which was Wolf’s original) and asked to rate the blameworthiness of each agent on a numerical scale. In every case, JoJo was found blameworthy, and under no circumstance did any student intuitively find JoJo free of moral responsibility or normatively insane.

<sup>28</sup> Wolf, 507.

<sup>29</sup> Faraci and Shoemaker, 19.

important distinction – that of moral insanity and moral ignorance – which may explain our intuitions about blameworthiness. It does *not* aim to determine whether or not JoJo is truly cognitively insane or merely morally ignorant. In fact, I hold that either of the two are distinct possibilities: it may be the case that JoJo is not acting out of his deeper self at all (in which case, he is not blameworthy), or that he is merely ignorant of what is morally right and wrong (in which case, he is blameworthy).<sup>30</sup> Nevertheless, it may allow us to see why we uphold a scalar assessment of moral responsibility: “differences in degrees of assessment may correspond inversely to the degree of difficulty in the target’s coming to moral recognition [or] lack of exposure to moral alternatives.”<sup>31</sup>

## **VI. Lingering Questions and Concluding Remarks**

Having examined the DSV and both aspects of Wolf’s sanity condition, I have concluded that the sanity condition is either unnecessary as an addendum to the DSV or irrelevant for moral responsibility. This is because the cognitive aspect of the sanity condition is implicitly inherent in the DSV itself, and thus unnecessary, or because it is impossible for an agent acting out of her

---

<sup>30</sup> This is to say, if JoJo truly lacks the capacity to distinguish between normative right and wrong actions, he is not only normatively insane but also cognitively insane. If this is the case, his deeper self is not truly governing his actions, and he fails to meet the DSV and is not a responsible agent. On the other hand, if JoJo *does* meet the DSV (and thus, is cognitively sane), then he is merely morally ignorant. If this is the case, we may hold him less blameworthy than an agent who has a fuller exposure to moral alternatives (due to his stringent upbringing), but he is not fully *relieved* of moral responsibility, as Wolf holds.

<sup>31</sup> Faraci and Shoemaker, 11-18.

deeper self to fail to meet the normative sanity condition, and thus it is irrelevant.

Nevertheless, there are important residue questions that Wolf's sanity condition brings up that, although lie outside of the scope of this paper, are certainly worth noting. Primarily, it is important to explore whether the DSV can account for differences in degrees of assessment of blame, given that we accept the existence of a scalar evaluation of blameworthiness. Further, the link between cognitive science and philosophical evaluation of moral responsibility merits further study, as a means of assessing the extent to which agents who seem morally insane truly *are* cognitively impaired (as opposed to morally ignorant). Finally, further discussion of the role of moral luck (particularly constituent and circumstantial) in assessing an agent's severity of blameworthiness is significant. Although I will not discuss possible resolutions to these issues, as these issues go well beyond the scope of this paper, I acknowledge that the need for further exploration of the responsible agent's moral blameworthiness would not be as significant without Wolf's writing on the subject.

# Autonomy and Integrity

*Victoria Koc*

***Abstract:***

*In this paper, I cast Ronald Dworkin's idea of political integrity as a plausible response to Robert Paul Wolff's philosophical defense of anarchism. Wolff claims that any account of de jure legitimacy, which he defines as the genuine possession of a "right to be obeyed," conflicts with our fundamental moral obligation to be autonomous. In contrast, Dworkin asserts that a legitimate political institution characterized by integrity in fact aids the ideal of self-legislation. Here I seek to expand upon this underdeveloped claim.*

## **I. Introduction**

In his provocative article "In Defense of Anarchism," Robert Paul Wolff denies the possibility of a legitimate government—that is, one that possesses the "right to be obeyed"—on the grounds that the correlative duty to obey is in conflict with the moral ideal of autonomy. Central to Wolff's thesis, as I shall explain in greater detail in the following section, is his insistence that our politics be founded on such principles of morality. Since this article's publication, many scholars in legal and political philosophy have engaged with Wolff's radical claim, whether explicitly or indirectly, offering accounts of legitimate political authority that do appeal to fundamental moral values. In *Law's Empire*, Ronald Dworkin hints at one such

case, suggesting that his conception of political integrity has moral and expressive benefits in its role of supporting the ideals of self-legislation. Dworkin moves on rather quickly, turning instead to the grounds of associative obligations produced by true communities (which must possess the attitudinal characteristic of political integrity) to find support for legitimate political authority. However, in this paper, I would like to explore the possibility of a defense of legitimacy based on this notion of integrity. I will attempt to develop Dworkin's position, demonstrating how political integrity might indeed be a safeguard to autonomy in the sense that it protects the ideal of self-legislation within a true political community. By doing so, I hope to offer a response to Wolff's claim that political institutions inherently strip their constituents of their autonomy and therefore never can be regarded as truly legitimate in their claims to authority.

## **II. The Connection Between Morality and Politics**

Before turning to Dworkin's philosophy, however, we should provide a fuller account of the connection between politics and morality addressed in Wolff's article. He begins the paper by distinguishing authority in the descriptive sense from authority in the normative sense. Under the former definition, having authority is merely having the claim of the right to be obeyed "acknowledged and accepted by those at whom it is directed" (Wolff, 5). The normative meaning, on the other hand, consists of actually having that right. To illustrate, consider the case of a dictator. If members of the state make no attempt to overthrow his rule, he possesses authority in a descriptive sense, but our intuitions prevent us (from an external perspective) from wanting to say that he actually has that right or that his rule is entirely legitimate on these grounds.

In other words, we do not consider him to have authority in the normative sense. It is this latter meaning of the term “authority” that is of interest to the overwhelming majority of political philosophers and central to my discussion here.

Having made the necessary clarifications, Wolff turns to a consideration of the criteria for the possession of supreme authority. Here he focuses on the difference between acting on a persuasive argument and following an authoritative command. When a person or entity truly possesses authority, his demands will be met because of who he is, not what he says. For example, I am not exactly following the commands of any authority when I choose not to commit murder, for I would behave in the same way even in the absence of its existence—or, if the authority changed its nature and suddenly demanded homicide, I would deny its power over my actions and still refrain from killing. However, I may submit to the authority when I follow traffic laws; while I have practical reason to obey these in general (e.g., public safety), I have no moral obligation based on their substantive content (it is not the case that I must drive on the right side of the road; in fact, if I am in another country I am likely to be driving on the left). To see a figure or institution as possessing supreme authority, though, does not mean that I recognize it as authoritative in enforcing arbitrary legislation like traffic laws alone. In the case of supreme authority, my primary (perhaps even sole) reason for acting in accordance with every law must be that it is that authority’s command. Wolff notes, “That men accede to such claims of supreme authority is plain. That men *ought* to accede to claims of supreme authority is not so obvious. Our first question must therefore be, ‘Under what conditions and for what reasons does one man have supreme authority over another?’” (Wolff, 8). With this introduction of the question of “ought,” particularly in the context of interpersonal

relations, Wolff brings us to the connection between morality and politics.

To further develop the task of political philosophy, Wolff uses as a model Kant's process of deduction, in which the goal is to demonstrate the legitimacy of a concept (as opposed to the sort of deductive reasoning we use in formal logic). In this case, the concept under investigation is that of *de jure* authority (or, to use previous terminology, the right to be obeyed possessed by states of the normative brand). Wolff observes that we are at least in possession of such a concept, demonstrated by the existence of *de facto* states; whether or not we ultimately find it legitimate thus becomes political philosophy's main subject matter.

### **III. Wolff's Anarchistic Argument**

Wolff dedicates the remainder of his article to making a case for the negative answer to this basic question based on its apparent conflict with the idea of autonomy. He begins by noting that "The fundamental assumption of moral philosophy is that men are responsible for their actions," (Wolff, 12) an uncontroversial characterization of the discipline. From this assumption along with the elements of free will and the capacity for human reason, we quickly arrive at the Kantian claim that men are metaphysically free and therefore mostly capable of choosing their actions. The truth of this statement is enough to make us responsible, but to actually take responsibility incurs the burden of attempting to determine the right course of action, constantly working to gain knowledge, reflect on motives, predict outcomes, criticize principles—all dependent on the capacity to reason implicit in the original assumption that we are responsible for our actions. Being in possession of

this ability in conjunction with free will leads to the moral obligation to utilize them and actively take responsibility. Doing so will not always lead to the right action, necessarily, but as long as we are engaged in this reflective process, we may be called responsible individuals.

Wolff provides further description of what it means to be responsible, stating that a responsible individual is still bound by moral constraints, but what those constraints are must be determined by the individual alone (this is not to say that he cannot consider the arguments of others, as long as the ultimate determination of their validity is achieved by him and no one else). From these requirements, Wolff concludes that we may characterize the responsible individual as self-legislating, as he gives laws to himself in the sense that he expresses his moral decisions to himself in the form of imperatives (Wolff, 13-14). A self-legislating individual, then, can be said to be morally autonomous according to the Kantian definition (the combination of freedom and responsibility).

Here Wolff arrives at the heart of the conflict between individual autonomy and our conception of a state legitimate in the *de jure* sense; he states, “The autonomous man, insofar as he is autonomous, is not subject to the will of another. He may do what another tells him, but not *because* he has been told to do it” (Wolff, 14). In other words, an autonomous individual cannot recognize the supreme authority of another over him; and it follows that in a community of autonomous men, legitimate *de jure* authority cannot exist, for no member can claim this right over another—nor can any autonomous member accept this kind of obligation.

This conclusion begs a related question—could a community of non-autonomous men even exist? Or, put differently, can an individual forfeit his autonomy? Wolff

suggests that yes, he can, recalling that to be autonomous requires the active use of reason and free will. A man can choose not to take full responsibility (I can, for example, refuse to think before I act, or sign a form allowing a guardian to act on my behalf in certain circumstances) and therefore forfeit autonomy; however, this does not relieve him of the moral responsibility. Thus, Wolff claims that in submitting to a political authority in the *de jure* sense (which Wolff problematically assumes is external and entirely distinct from the community—this implies that we have no significant influence over the laws that govern us, a feature not true of Dworkin’s personified political community)—by forfeiting autonomy—we would be acting against our highest moral obligations. Such a scenario is logically possible, then, but it ought not to be realized. Wolff therefore concludes that “anarchism is the only political doctrine consistent with the virtue of autonomy” (Wolff, 18).

#### IV. Political Integrity

In *Law’s Empire*, Ronald Dworkin appears to take a very different (and, in my view, more appealing) stance, stating that the ideal of self-legislation on which Rousseau and Kant based their conceptions of freedom (clearly the same virtue reference by Wolff in the context of his Kantian explanations) “needs integrity [...] for a citizen cannot treat himself as the author of a collection of laws that are inconsistent in principle” (Dworkin, 189). He unfortunately does not dwell on this claim; however, given the claim that we ought to build a politics on such a moral ideal, it is very much worth considering whether this so-called political integrity really does support self-legislation and, if so, how we might find support for the legitimacy of *de jure* states after all, despite Wolff’s skepticism.

Dworkin selects the term “political integrity” to make explicit its connection with the “parallel ideal of personal morality” (Dworkin, 166). As discussed in the section on taking personal responsibility, adhering to the ideal of political integrity does not demand that the absolute right action be taken in every case, but rather that the options pertaining to each situation be given serious consideration indicative of a sincere effort to reach the best decision. In essence, then, Dworkin is treating the state as a moral agent much in the same sense that we view an individual (for example, we may describe a nation like the United States as “greedy” much in the way we would an individual person). This comparison becomes increasingly evident in his subsequent discussion on the deep personification at work with respect to the political community.

By characterizing the body politic as a collective moral agent held to the same standards as the individual moral agent, Dworkin remarks that we are “endorsing a complex way of reasoning about the responsibilities of both officials and ordinary citizens that finds a natural expression only in the personification of community and cannot be reproduced by a reductive translation into claims about officials and citizens one by one” (Dworkin, 168-169). By this he means to say that the community’s standards and principles are more than just an amalgamation of those of its constituents; they belong to the group as an indivisible whole.

Recognizing this personification in conjunction with our presumed understanding of individual integrity provides a decent basic idea of Dworkin’s political version of the virtue, but in order to properly address Wolff’s argument, we must develop a more thorough understanding of what political integrity in its own right involves. At its most simple level, Dworkin describes the notion with the catch phrase that “we must treat cases alike” (Dworkin, 165).

More explicitly, the virtue of political integrity requires the community as a whole (Dworkin views the political community as a special entity via a deep personification) to “act in a principled, coherent manner, treating all citizens under the same standards of justice or fairness” (Dworkin, 165). He explains that acting with political integrity is not merely seeking consistency by constantly conforming to past decisions; on the contrary, it may require us to depart from such standards in search of fidelity to the principles that are more fundamental to the scheme as a whole. To illustrate this claim, Dworkin refers us to the 1954 Supreme Court decision in *Brown v. Board of Education* (1954). In this hearing, the plaintiff was successful “in showing that an important part of what has been thought to be law is inconsistent with more fundamental principles necessary to justify law as a whole” (Dworkin, 221). Specifically, in the *Brown* case, the justices recognized the need to overturn the principles underlying the earlier decision in *Plessy v. Ferguson* (1896) that upheld the doctrine of “separate but equal” in accepting the constitutionality of segregation in public facilities. In *Brown*, the justices concluded that while the earlier court had perceived this as maintaining legal or formal equality, by allowing the resulting social inequality, the principle of equality behind the law was being completely undermined.

The integrity at work in the *Brown* decision is adjudicative in nature; Dworkin notes that integrity must be present in the legislative process as well. The virtue can thus be successfully maintained as legislators strive to “keep the law coherent in principle,” (Dworkin, 167) and judges attempt to see and enforce it in that manner—in essence, it permeates the entire political process. Here, he explains, a fidelity to integrity may come into conflict with the other standard political ideals of fairness (a more empirical matter of distributing political power so that

citizens have approximately equal influence in the decisions that govern them) and justice (where the ultimate concern is the what decision the standing institutions ought to make, regardless of the presence of fairness in procedure). Our real political world, unlike a utopian one, is far from perfectly fair or just, so we have no guarantee of coherence in legislative principles and thus need to recognize integrity as an independent ideal to keep the law-making procedure in check. When these virtues conflict, fairness and justice must sometimes be sacrificed in order to maintain integrity—but this is not to say that integrity in either of its forms should reign sovereign over the other virtues, necessarily.

To illustrate how we already balance these ideals, Dworkin bring up the issue of internal compromises, specifically “checkerboard” solutions, which he defines as “statutes that display incoherence in principle and that can be justified, if at all, only on grounds of a fair allocation of political power between different moral parties” (Dworkin, 435 n.6). Dworkin proceeds to set up a puzzle that questions our intuitive distrust of such solutions. He demonstrates that it does not stem from the belief in fairness, as from that virtue combined with the recognition of diametrically opposed moral values with respect to many hot button issues, such as that of abortion, “it would seem to follow [...] that legislation on these moral issues should be a matter not just of enforcing the will of the numerical majority, as if it were unanimous, but of trades and compromises” (Dworkin, 178). He proposes that we could turn to a Solomonic model of compromise that “treats a community’s public order as a kind of commodity to be distributed in accordance with distributive justice, a cake to be divided fairly by assigning each group a proper slice” (Dworkin, 178-179). However, as the biblical allusion implies, we quickly come to recognize that the solution

resulting from this process seems arbitrary and therefore unappealing when considering strong matters of principle. Instead, as Dworkin observes, we follow the model that requires “that each point of view must be allowed a voice in the process of deliberation”—so far, so good; we have achieved procedural fairness—“but that the collective decision must nevertheless aim to settle on some coherent principle whose influence then extends to the natural limits of its authority”—in other words, we do not in this model produce a checkerboard solution, which would contain bits and pieces of various moral principles, obscuring any unity at this foundational level. The reasoning behind our avoidance cannot be the same principle of fairness that applies to the procedural aspect of this model; checkerboard solutions are in a theoretical sense fairer than any others (everyone’s influence would be apparent in the compromise law).

We might think the failure of checkerboard solutions lies in the realm of justice, then (remembering that political ideals do conflict), but Dworkin maintains that this is not really the case either. The winner-take-all result, whichever substantive side we settle on, is always the least just outcome for the losing party; checkerboard solutions, while they do produce injustices for both parties, result in a less extreme lack of balance. Not knowing the outcome in advance, checkerboard solutions are the less risky option, as they are never absolutely unjust to either side the way winner-take-all results inevitably are. Thus, a commitment to justice in the ultimate decision gives us no logical reason for rejecting checkerboard legislation, at least a priori.

Dworkin also considers the possibility that checkerboard solutions seem unjust to us by their very definition, because they “treat different people differently for no good reason, and justice requires treating like cases alike” (Dworkin, 180). He rejects this hypothesis, however,

suggesting that justice as we have defined it does not explain why this is always objectionable, “for in the circumstances of ordinary politics the checkerboard strategy will prevent instances of injustice that would otherwise occur, and we cannot say that justice requires not eliminating any injustice unless we can eliminate all” (Dworkin, 181). He also denies a third possible reason derived from the ideal of justice that suggests our hesitation lies in the belief that “no one should actively engage in producing what he believes to be injustice” (Dworkin, 181). Dworkin correctly observes that accepting this claim explains only legislating behavior (why members of the community would abstain from voting), not the rejection of the legislative outcome.

Having thoroughly refuted the rejection of checkerboard solutions based on terms of fairness and justice, Dworkin proposes that we can explain our instincts through his ideal of political integrity: “we say that a state that adopts these internal compromises is acting in an unprincipled way,” (Dworkin, 183) as the solutions could not be attributed to any coherent set of morals, but rather fragments of different value systems as selected by the individual members; moreover, the state would have to “endorse principles to justify part of what it has done that it must reject to justify the rest” (Dworkin, 184).

#### **V. Political Integrity and Self-Legislation**

Now equipped with an understanding of political integrity as defined by Dworkin, we can begin to seek evidence in support of his claim that integrity as a political virtue supports the moral ideal of self-legislation (and consequently, autonomy). It should be rather apparent that this statement applies when considering the state as a deeply personified, collective moral agent in possession of the same capacities and held to the same standards of the

individual. But the question remains whether the same can be said to apply to each member of the state—this is important if we are to consider Dworkin’s theory as a response to Wolff’s, as the latter was expressly concerned with the moral status of individuals under the state’s governance.

Dworkin does not provide a thorough answer to this question, but his brief development of the thought is a good place to start. Here he introduces the following set of related sub-claims: “Integrity expands and deepens the role individual citizens can play in developing the public standard of their community because it requires them to treat relations among themselves as characteristically, not just spasmodically, governed by these standards;” “Integrity [...] insists that each citizen must accept demands on him, and may make demands on others, that share and extend the moral dimension of any explicit political decisions;” and, therefore, “Integrity [...] fuses citizens’ moral and political lives” (Dworkin, 189). I will now consider these propositions, evaluating their accuracy and relevance to the ideal of self-legislation.

Dworkin’s first claim can be further broken down into two parts, the first limited to the suggestion that integrity changes citizens’ view of relations with each other—this element is crucial to the subsequent statements. Why should coherence in law connect citizens at this deeper level? Dworkin suggests that it does so by placing all encounters among citizens into the category of falling within the scope of some past political decision. In contrast, if we view all legislative decisions as negotiations applying to totally discrete problems, politics only enters our interactions in the context of those negotiations. There is no implication that these political decisions are united by a common set of principles, or that the community as a whole even possesses one. However, in a community

characterized by political integrity—where all decisions are based on a set of moral principles—it follows that not only does the community possess a common value system, but that the members of this community share in it, or at least have an interest in shaping it. Thus a commitment to political integrity implies that citizens are bound to a common set of morals (by the assumption that citizens will follow the law—though we have not yet reached the conclusion that they are morally obliged to do so, as Wolff pointed out, we have gathered plenty of evidence that they do). This phenomenon is in a way similar to what occurs in a school or workplace, where everyone is connected by the commitment to follow a particular set of rule for the good of the group, or perhaps even more closely tied to a religious community, where members share a true belief in this common set of moral principles (such is really the aim of a state of political integrity). In the political state, as in all of these examples pulled from personal life, all active participants have reason to treat each other as connected on a relatively deep level.

This connection influences not only our attitudes, but our actions as well. When we are all bound to a common value system, we have motivation to actively pursue a public standard that complements our private standards, whether by overtly political actions such as grassroots campaigning or more covert, social efforts to develop a shared understanding of abuses (physical violence, obscenity, even rejection of certain traditions). In order to put into effect a successful plan for either kind of work, we must first develop our private standards, which requires something akin to the process of reflection crucial to actively taking responsibility, as discussed earlier. As a product of this consideration, the private standards we settle on are essentially moral laws we determine for ourselves. Whether or not these ultimately are translated

into the public set of laws is another issue, but by giving us reason to reflect first at the personal level, it would seem that striving for political integrity does indeed promote self-legislation.

Dworkin also appears to be correct in stating that political integrity “fuses citizens’ moral and political lives” (Dworkin, 189). We can already see how the political influences the personal in encouraging the development of private standards of principle. In order to create a correspondence between these standards and those of the community to which we belong, we start to regard all social action—even simple conversation with neighbors—as potentially political, since even these most personal interactions can influence the development of community standards. Similarly, political action becomes morally charged, since decisions in this realm are not discrete negotiations but rather reflect the morals of the community and its members. Political thought then ceases to be entirely practical in nature and takes on a philosophical element.

How does all of this relate to the concept of self-legislation, though? Well, in real (that is, not hypothetical or utopian) political situations, our actions are necessarily constrained by the standards of the community in a very practical sense (note that these constraints differ from those grounded in morality, for the latter would result in the conclusion that all communities restrict self-legislation). We are far less likely to act on our own laws if they are in conflict with those of the surrounding community, for the consequences of acting against the latter (e.g., being sentenced to jail time), would put a great burden on our daily lives and potentially hinder our ability to act on other principles of our own (for example, if incarcerated, it becomes difficult to pursue positive moral action, such as actively helping others, as opposed to merely refraining from hindering their freedom). When the fusion of private

and public takes place, conflict between principles in the two realms ceases to exist, and we are no longer required to make a decision between them, either choosing to stand by our morals and suffer political repercussions, or following political laws at the expense of our personal standards. The principles become one and the same, so we can comply with both sets equally and fully.

In addition, when we successfully combine our personal lives with the life of the body politic, the commands issued by the “authority” are wholly derived from our own maxims. Though the details of application and practice may come from another party, we are the creators of the underlying theoretical principle, and in striving for self-legislation, this is the important aspect—what matters is which morals we choose to live our lives by, not necessarily how we implement them. Moreover, as active citizens, we may view ourselves as inseparable from the authority (which encompasses the entire community as a unit—refer back to Dworkin’s deep personification)—all social interaction takes on the goal of constructing a morally desirable public set of standards. In essence, then, we become the basic unit of the legislative body of the state (the authority), as the legislature’s task is to create laws that will achieve compliance with said standards. Therefore, the commands we follow are in fact laws of our own creation, and we may say without hesitation that we act in accordance with them because of who they come from—to grant ourselves supreme authority is not problematic; on the contrary, it is the ultimate taking of responsibility. If we are not active members of a community charged with political integrity—that is, if we choose to totally and completely withdraw from the community, merely existing in its territory—we cannot claim our status as a part of the state/authority; in rejecting all social opportunities, we take on no legislative role. The commands that are ultimately

enforced, and the body that enforces them (both officials and fellow citizens invested in the community) remain external, and the conflict between following internal (self-legislated) and external orders returns.

Dworkin also states that integrity is necessary to fuel self-legislation “for a citizen cannot treat himself as the author of a collection of laws that are inconsistent in principle” (Dworkin, 189). This requirement is fairly self-evident, but I will dwell upon it for a moment due to its importance. Turning again to Kant—from whose writings we obtain the ideal of self-legislation—we may recall that his categorical imperative (the fundamental principle on which we base all moral decisions) demands universality, which Kant maintains requires consistency; i.e., any given maxim can only pass the universalizability test if it can consistently be willed as a universal principle (see *Groundwork of the Metaphysics of Morals*). Consistency among maxims is quite obviously necessary as well, then, for if one maxim comes into conflict with another, it cannot be said to be consistently (that is, without contradiction) willed. For example, a maxim to create economic equality through whatever means possible could come into conflict with a maxim not to steal, so perhaps the former maxim cannot be consistently willed by all. Therefore, it cannot pass the universalizability test and we cannot accept it as a moral principle under the constraints of the categorical imperative. To sum, Kant’s moral philosophy does depend on the sort of consistency pursued via Dworkin’s notion of political integrity (which may, to return to the example, put certain constraints on the legally available means for pursuing socio-economic equality), and the central ideal of self-legislation is in fact supported, as Dworkin claims in *Law’s Empire*.

## **VI. Self-Legislation and Legitimacy**

According to Robert Paul Wolff, we cannot grant legitimacy to a state that requires its constituents to forfeit their autonomy, as doing so would force us to deny our most fundamental moral obligations. Ronald Dworkin's philosophy demonstrates that contrary to Wolff's conclusion that we cannot attribute legitimacy to *any* state, one that holds political integrity as one of its primary virtues actually helps us, as members of this political community, realize the principle of self-legislation and thereby autonomy. Thus legitimacy should not be a priori denied to such a state. However, the bigger question remains: is the fact that a state aids individuals in complying with the moral ideal of self-legislation sufficient grounds for granting it legitimacy?

I should think the answer to this question would be yes. Achieving self-legislation in the fullest sense of the term requires not only constant reflection on our principles, but actually acting on our most thought-out set of laws. Action for human beings is necessarily social in nature; thus, society always puts constraints on our capacity to act (whether those be formal laws or informal but perhaps equally imposing customs). When society's standards match those of our own, restrictions cease to be problematic, for they then only add practical reasons on top of our moral reasons to comply with rather than deny our personal maxims.

Furthermore, when we truly achieve self-legislation, it necessarily follows that we are taking responsibility in all aspects of life, for the self-legislating process of reflection based on reason and morality is the very definition of taking responsibility as described by Wolff. The ability to take such responsibility, we should recall, implies that we have personal autonomy and we are thus metaphysically

free. The state (provided it is characterized by political legitimacy), rather than limiting our freedom, actually helps sustain it by aiding self-legislation.

Assuming (as we generally do) that securing this freedom and fulfilling our duty to be autonomous is the ultimate moral obligation, derived from our given capacities to reason and the assumed element of free will, it should follow that we must choose a lifestyle in which such a goal is not only realizable, but realizable to the fullest extent and by the most efficient means possible. It appears from our discussion that building a political community imbued with integrity in the legislative and adjudicative processes constitutes such a lifestyle. Therefore, submitting to this community becomes an obligation rooted in the basics of morality—in short, we are morally obliged to take on the duty to obey the political body of which we are a part. The correlative of this obligation, as stated in the discussion of Wolff's argument, is the state's possession of legitimate authority in the normative in addition to the descriptive sense of the term.

## **VII. Conclusion**

By studying Ronald Dworkin's theory of political integrity in detail and evaluating his claims of the moral and expressive benefits such an ideal provides—specifically in its relation to self-legislation—I have rejected Robert Paul Wolff's radical conclusion that states can never be granted *de jure* legitimacy—that is, true possession of what he calls the right to be obeyed—due to their inherent restriction of self-legislation and by natural association, autonomy. Dworkin's picture of a state characterized by political integrity (coherence in legal principle, in both the legislative and adjudicative processes)

demonstrates that some political communities can actually aid pursuit of this ideal. When this is the case, I suggest that not only do we have practical reasons to submit to the authority of the state, thereby granting it *de facto* legitimacy, but that we incur genuine obligations in addition due to the fundamentality of pursuing this moral ideal, and we can thus grant the state *de jure* legitimacy. I have suggested that these grounds are sufficient to constitute a successful theory of political obligation and authority, despite both Wolff's claim that such a model is altogether impossible, and Dworkin's unwillingness to base his philosophy of legitimacy upon it (as I mentioned, he goes on in the chapter to find grounds for legitimacy rooted in what I consider to be the more vague obligations stemming from "true" communities). Justification of legitimacy, in my view, needs no further support beyond assistance in complying with our highest moral obligation.

## **BIBLIOGRAPHY**

- Aristotle.** (1987). *The Nicomachean Ethics*, trans. J. E. C. Welldon. Amherst, N.Y.: Prometheus Books. Print.
- Baumert, Kevin A., Herzog, T., & Pershing, J.** (2005). *Navigating the numbers: Greenhouse gas data and international climate policy*. Retrieved from [http://pdf.wri.org/navigating\\_numbers.pdf](http://pdf.wri.org/navigating_numbers.pdf)
- Bell, P.** (2009). Exempt northerns from emission cuts: Inuit leader. *CBC News North*. Retrieved from <http://www.cbc.ca/news/canada/north/story/2009/12/10/copenhagen-n-north-exemption.html>
- Boden, T.A., Marland, G., & Andres, R.J.** (2010). *Global, regional, and national fossil-fuel CO<sub>2</sub> emissions*. Retrieved from [http://cdiac.ornl.gov/trends/emis/tre\\_glob.html](http://cdiac.ornl.gov/trends/emis/tre_glob.html)
- Burge, Tyler.** "Content Preservation." *The Philosophical Review* 102, no. 4 (1993): 457-488.
- Caney, S.** (2005). Cosmopolitan justice, responsibility, and climate change. *Leiden Journal of International Law*, 18(4), pp. 747-75.
- Carbon Dioxide Information Analysis Center.** (2009). Fossil-fuel CO<sub>2</sub> emissions by nation. Retrieved from [http://cdiac.ornl.gov/trends/emis/tre\\_coun.html](http://cdiac.ornl.gov/trends/emis/tre_coun.html)
- Center for Climate and Energy Solutions.** (n.d.). Katrina and global warming. Retrieved from <http://www.c2es.org/science-impacts/extreme-weather/hurricane-katrina>
- De Clercq, R.** *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 132, No. 3 (Feb., 2007), pp. 553-563.
- Earman, J.** (1989): 'Style and Supervenience', *British Journal of Aesthetics* 39, 255-262.
- Dworkin, Ronald.** *Law's Empire*. Cambridge, MA: Belknap, 1986. Print.
- Faraci, David and David Shoemaker.** "Insanity, Deep Selves, and Moral Responsibility: The Case of JoJo." *Review of Philosophy and Psychology*.
- Feldman, Richard, and Conee, Earl.** "Evidentialism." *Philosophical Studies* 48, no. 1 (1985): 15-34.
- Frankfurt, Harry G.** "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy*. Vol. 63, No. 1. Journal of Philosophy, Inc. 1971.
- Gertner, J.** (2010, May 16). The rise and fall of the G.D.P. *The New York Times*, p. MM60.
- Gillis, J.** (2012, November 1). Are humans to blame? Science is out. *The New York Times*, p. A29.

- Goldberg, Sanford.** “Testimony as Evidence.” *Philosophica* 78, no. 2 (2006): 29-51.
- Goldman, Alvin.** “A Guide to Social Epistemology” in *Social Epistemology: Essential Readings*, Alvin I. Goldman (ed.) and Dennis Whitcomb (ed.). New York: Oxford University Press, (2011): 11-37.
- Grimes, P. & Kentor, J.** (2003). Exporting the greenhouse: Foreign capital penetration and CO2 emissions 1980-1996. *Journal of World-Systems Research*, 9(2), pp. 261-75.
- Hardwig, John.** “Epistemic Dependence.” *The Journal of Philosophy* 82, no. 7. (1985): 335-349.
- . “Evidence, testimony, and the problem of individualism – a response to Schmitt.” *Social Epistemology* 2, no. 4. (1988): 309-321.
- Harvey, F.** (2012, July 31). Europe looks to open up Greenland for natural resources extraction. *The Guardian*. Retrieved from <http://www.guardian.co.uk/environment/2012/jul/31/europe-greenland-natural-resources>
- Hessell, I.** (2006). *Arctic Spirit: The Albrecht Collection of Inuit Art at the Heard Museum*. Vancouver: Douglas & McIntyre.
- Hill, R.A.** (2002). Compensatory justice: Over time and between groups. *The Journal of Political Philosophy*, 10(4), pp. 392-415.
- International Monetary Fund.** (2012). *World economic outlook database, October 2012*. Retrieved from <http://www.imf.org/external/pubs/ft/weo/2012/02/index.htm>
- IPCC third assessment report: Glossary of terms.** (2001). Retrieved from <http://www.ipcc.ch/pdf/glossary/tar-ipcc-terms-en.pdf>
- Kant, Immanuel.** *Groundwork of the Metaphysics of Morals*. Ed. Mary J. Gregor. Cambridge, U.K.: Cambridge UP, 1998. Print.
- Kelly, Thomas.** “Evidence.” *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/evidence/>
- . “Peer Disagreement and Higher Order Evidence” in *Social Epistemology: Essential Readings*, Alvin I. Goldman (ed.) and Dennis Whitcomb (ed.). New York: Oxford University Press, (2011): 183-217.
- Lackey, Jennifer.** “Testimony: Acquiring Knowledge from Others” in *Social Epistemology: Essential Readings*, Alvin I. Goldman (ed.) and Dennis Whitcomb (ed.). New York: Oxford University Press, (2011): 71-91.
- Levinson, J.** (1980). Aesthetic uniqueness. *The Journal of Aesthetics and Art Criticism*, 38(4), 435-449. Retrieved from <http://www.jstor.org/stable/430325?&Search=yes&searchText=aes>

[thetic&searchText=uniqueness&list=hide&searchUri=/action/doBasicSearch?Query=aesthetic+uniqueness&fromHomePage=true&acc=on&wc=on&fc=off&prevSearch=&item=4&ttl=11146&returnArticleService=showFullText](#)

- Levitus, S., Antonov, J. & Boyer, T.** (2005). Warming of the world ocean, 1955–2003. *Geophysical Research Letters*, 32, doi:10.1029/2004GL021592.
- Levitus, S., Antonov, J., Boyer, T. and Stephens, C.** (2000). Warming of the world ocean. *Science*, 287(5461), pp. 2225-9.
- Locke, J.** (1689). An essay concerning the true original, extent, and end of civil government. In J. Laslett (ed.), *John Locke: Two Treatises of Government*. Cambridge: Cambridge University Press.
- Lomasky, L.** (1991). Compensation and the bounds of rights. In J.W. Chapman (ed.), *Nomos XXXIII: Compensatory justice*, New York: New York University Press pp. 13–44.
- Lowe, E. J.** (2002). *A Survey of Metaphysics*. Oxford, NY: Oxford University Press.
- Malkin, E.** (2012, October 30). Yet another blow to Haiti from a natural disaster. *The New York Times*, p. A9.
- McLaughlin, B.** (2005, July 25). “Supervenience.” *Stanford encyclopedia of philosophy*. Retrieved from <http://plato.stanford.edu/entries/supervenience>
- Noë, Alva.** *Action in Perception*. Cambridge: The MIT Press, (2004).
- Nozick, R.** (1974). *Anarchy, state, and utopia*. New York: Basic Books.
- Parfit, Derek.** *Reasons and Persons*. New York: Oxford University Press, (1984).
- Parry, M.L., Rosenzweig, C., Iglesias, A., Livermore, M., & Fischer, G.** (2004). Effects of climate change on global food production under SRES emissions and socio-economic scenarios. *Global Environmental Change*, 14, 53-67. doi: 10.1016/j.gloenvcha.2003.10.008
- Perring, Christian.** “Mental Illness.” *Stanford Encyclopedia of Philosophy*. Spring 2010. <<http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=mental-illness/>>
- Pogge, T.** (2002). *World poverty and human rights*. Malden, MA: Blackwell Publishers.
- . World poverty and human rights. *Ethics & International Affairs*, 19(1), 1-7.
- Posner, E.A. & Weisbach, D.** (2010). *Climate change justice*. Princeton, NJ: Princeton University Press.

- Olivier, J.G.J., Janssens-Maenhout, G., & Peters, J.A.H.W.** (2012) *Trends in global CO<sub>2</sub> emissions*. Retrieved from <http://edgar.jrc.ec.europa.eu/CO2REPORT2012.pdf>
- Sanford, David.** "Some puzzles about prosthetic perception." Paper presented at the 1997 meetings of the Society for Philosophy and Psychology, New York (1997), cited in Noë, Alva. *Action in Perception*. Cambridge: The MIT Press, (2004): pp. 173-175, 266.
- Sayre-McCord, Geoff.** "Moral Realism." *The Stanford Encyclopedia of Philosophy*. Summer 2011. <<http://plato.stanford.edu/archives/sum2011/entries/moral-realism/>>
- Sorensen, R.** *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 100, No. 2 (Aug., 2000), pp.175-191.
- Stigset, M.** (2011, January 10). Greenland steps up its independence calls as oil ambitions grow. *Bloomberg*. Retrieved from <http://www.bloomberg.com/news/2011-01-11/greenland-steps-up-demands-for-independence-as-oil-revenue-spurs-ambitions.html>
- United Nations Development Programme.** (2011). Human development report 2011. Retrieved from [http://hdr.undp.org/en/media/HDR\\_2011\\_EN\\_Complete.pdf](http://hdr.undp.org/en/media/HDR_2011_EN_Complete.pdf)
- Watson, Gary.** "Free Agency." *The Journal of Philosophy*. Vol. 27, No. 8. Journal of Philosophy, Inc. 1975.
- Watt-Cloutier, S.** (2010). The Inuit right to culture based on ice and snow. In K.D. Moore & M.P. Nelson (Eds.), *Moral ground: Ethical action for a planet in peril* (25-7). San Antonio: Trinity University.
- Wolf, Susan.** "Sanity and the Metaphysics of Responsibility." *Reason & Responsibility: Readings in Some Basic Problems of Philosophy*. Ed. Joel Feinberg and Russ Shafer-Landau. Belmont: Thomson Learning (2001). 503. Print.
- Wolff, Robert P.** "The Conflict Between Authority and Autonomy." *In Defense of Anarchism*. New York: Harper & Row, 1970. 3-19. Print.