

Stance

An International Undergraduate Philosophy Journal



Stance

An International Undergraduate Philosophy Journal

Stance: An International Undergraduate Philosophy Journal is produced and edited entirely by undergraduate students. We aim to enrich student learning by providing an opportunity for undergraduate students to have their original scholarly work reviewed by, or published in, a peer-reviewed academic journal.

Stance is published annually in April. The deadline for submissions is in mid-December. All papers are carefully considered in multiple blind reviews. Notification of initial decision is early February. All authors receive constructive feedback concerning submissions.

Stance requires first publication right. All other rights reside with the author. Having an international standard serial number [ISSN 1943-1880-print; ISSN 1943-1899-electronic], *Stance* is registered with the U.S. Library of Congress. *Stance* is listed in *The Philosopher's Index*.

Further information regarding *Stance* is available at:

<http://stance.iweb.bsu.edu>. Back issues can be found on the Ball State University Virtual Press at: **<http://www.bsu.edu/library/virtualpress/stance>**

Inquiries should be directed to David W. Concepción, Ph.D.,
dwconcepcion@bsu.edu.

Stance gratefully acknowledges Ball State University's support of the publication of this journal.

Stance

An International Undergraduate Philosophy Journal

EDITORS-IN-CHIEF

Lorrie Caskey
Rosamae Swoape

MANAGING EDITOR

David W. Concepción, Ph.D.

ASSOCIATE EDITORS

Stephanie Atkins
Vivek Hadley
Emily Hinman
Robert Wilbur

CONTRIBUTING EDITORS

Jeremy Christman
Chris Hoover
Whittley Lewis
Ben Rogers

EDITORIAL BOARD

Veronica Buckler
Jess DeGabriele
Trevor Krogman
Dominic Martyne
Caleb Mosier
Kiley Neal
Brian Schimpf
Jacob Taylor

COPY EDITORS

Jess DeGabriele
Whittley Lewis
Ben Rogers

PRODUCTION/DESIGN

Eric Wyman

WEB DESIGN SPECIALIST

Brian Schimpf

External Staff

ASSISTANT EDITORIAL BOARD

Nathan Couch
University of Illinois

Caleb Pittman
University of Louisville

Adam Croom
University of Pennsylvania

Jared Nathaniel Smith
Flagler College

Beth Mendenhall
Kansas State University

Peter Tan
University of Illinois

EXTERNAL REVIEWERS

Milo Crimi
Kutztown University of Pennsylvania

John C. Hill
Gettysburg College

Anthony Adrian
Texas State University

Róbert Mátyási
University of Pécs

John Stockton
University of Southern California

Kevin Dupree
University of Central Florida

Yona Rabinowitz
Florida Atlantic University

Grace Christman
Western Kentucky University

Lauren Townsend
University of South Florida

Kevin Scott
Canisius College

Caleb Hamman
Butler University

Contents

The Piety of Escape	3
<i>Phillip Shannon, Lycoming College</i>	
Certainly Uncertain: Nietzschean Pessimism for an Optimistic World	15
<i>Stephen Bailey, Samford University</i>	
How We Are Moral: Benevolence, Utility, and Self-Love in Hobbes and Hume	27
<i>Jenna Kreyche, Boston University</i>	
Ruminations on Intermittent Existence	39
<i>Anthony Adrian, Texas State University</i>	
A Multi-Causal Approach To Synchronicity	49
<i>Zachary Stinson, Flagler College</i>	
A Critical Assessment of George Klosko's Version of the Principle of Fair Play	61
<i>Leonardo Moauro, Brown University</i>	
Gettiering Goldman	69
<i>Kenneth Stalkfleet, University of Illinois</i>	
A Rawlsian Revitalization of Gewirth's Normative Structure for Action	79
<i>Bo Fox Pons, University of South Dakota</i>	
A New Approach to the Paradox of Fiction	91
<i>Pete Faulconbridge, University of Warwick</i>	
The Universe Began to Exist? Craig's Philosophical Arguments For A Finite Past	103
<i>Blake McAllister, Pepperdine University</i>	
Book Review: Why Some Things Should Not Be for Sale by Debra Satz	115
<i>Brittney Sovik, Georgia State University</i>	

The Piety of Escape

Phillip Shannon

ABSTRACT: This paper examines two seemingly contradictory views of piety found in Plato's *Euthyphro* and *Crito*. Using the Socratic dialogues for evidence of what Socrates actually believed and to piece together a Socratic account of piety, it seems that his argument in favor of remaining in prison is inconsistent with his own beliefs. The paper concludes that Socrates ought not to have thought it was impious to escape from prison.

While awaiting his death sentence in prison, Socrates hears many arguments attempting to convince him of the piety of escape. In Plato's *Crito*, Socrates gives perhaps his most explicit account of his reasons for remaining in prison. However, when his argument is examined beside a parallel argument found in *Euthyphro*, it seems that Socrates may have been mistaken. After reviewing the arguments, I will discuss what I believe is an inconsistency in Socrates' argument. Then, using the Socratic account of piety found in *Euthyphro* and evidence of his beliefs found in *Euthyphro*, *Apology*, and *Crito*, I will demonstrate how a faulty premise ought to be restated. Finally, I will conclude that Socrates should not have thought it would be impious to escape. It appears that when the assumptions in Socrates' arguments are restated to fit what Socrates would likely have believed in accordance with his account of piety, his life-saving escape would have been pious.

After hearing Crito's arguments in favor of escaping his death sentence, Socrates states his reasons for remaining in prison. To escape, he ar-



The Piety of Escape

gues, would be flouting the verdicts of the city's courts, thereby decreasing the city's force. This would ultimately be harming the city. Taken to its logical conclusion, such an action would be an attempt "to destroy [the laws of Athens], and indeed the whole city."¹ Socrates believes that he has special moral obligations to Athens which are similar to the obligations owed to one's parents. He personifies the laws of Athens to create a metaphor comparing his relationship with his parents to his relationship with his city-state: "[C]ould you, in the first place, deny that you are our offspring and servant, both you and your forefathers? If that is so, do you think that we are on an equal footing as regards the right, and that whatever we do to you it is right for you to do to us?"² The conclusion is that just as a father has the right to do things to his son that his son cannot do to his father, the city has the right to do things to its subjects. "Do you think that you have this right to retaliation against your country and its laws? That if we undertake to destroy you and think it right to do so, you can undertake to destroy us, as far as you can, in return?"³ Socrates believes that, even though Athens is attempting to destroy him unjustly, he has no right to cause it harm.

Socrates then tries to demonstrate how escaping from his prison cell would be impious. "It is impious to bring violence to bear against your mother or father; it is much more so to use it against your country."⁴ The argument could be stated like this:

- I. It is impious to bring violence to bear against your country.
- II. To escape prison would be to bring violence to bear against your country.
- III. To escape prison would be impious.

Perhaps a better understanding of why Socrates believes each of these claims can be found elsewhere. Because of the analogy connecting

1. Plato, *Five Dialogues*, trans. G. M. A. Grube (Indianapolis: Hackett, 2002): 53.

2. Ibid.

3. Ibid., 54.

4. Ibid.

one's relationship with his parents to his relationship with his country, this argument runs parallel with the one being refuted by Euthyphro:⁵

- I. It is impious to bring violence to bear against your father.
- II. Euthyphro is prosecuting his father for murder, thereby bringing violence to bear against his father.
- III. Euthyphro is doing something impious.

Euthyphro denies the impiety of his actions. He believes that this argument is one of the misguided masses, and that "the pious is [...] to prosecute the wrongdoer, be it about murder or temple robbery or anything else, whether the wrongdoer is your father or your mother or anyone else; not to prosecute is impious."⁶ However, at the end of the dialogue, Euthyphro has still given no real reason why what he is doing is not impious. Socrates states that if Euthyphro had "no clear knowledge of piety and impiety," he would have never "ventured to prosecute [his] father for murder on behalf of a servant."⁷ We can thus hold Socrates to the claim that, when the piousness of an action is in question, the default position is to regard the action as impious unless one has a good reason to believe otherwise. With no such reasons, it is practical to believe that Euthyphro is doing something impious.

The first premise of the parallel arguments, restated as "It is impious to bring violence to bear against your country or father," is problematic. Socrates gives evidence in the *Apology* that there are cases in which it is permissible to bring violence to bear against your country. When Athens was ruled by the Thirty, Socrates was ordered to bring in Leon from Salamis to be executed. Believing this execution to be unjust, Socrates did not follow the order. He argues that, in making that decision, his "whole concern [was] not to do anything unjust or impious."⁸ Because he is disobey-

5. *Ibid.*, 5.

6. *Ibid.*, 6.

7. *Ibid.*, 20.

8. *Ibid.*, 37.



The Piety of Escape

ing the laws of his country, Socrates is “bringing violence to bear against his country.” We are presented with the following inconsistent triad in Socrates’ reasoning:

- I. It is impious to bring violence to bear against your country or father.
- II. Socrates disobeyed an order, thereby bringing violence to bear against his country.
- III. Socrates has not done anything impious.

At most, only two of these statements can be true. The most likely proposition to be correct is (II). Socrates did disobey an order, and, even if only to a small degree, this means that he brought violence to bear against his country.

Whether or not (III) is true depends largely on the account of piety one chooses. Because we are examining Socrates’ own argument, it is appropriate to use his account of piety. When Socrates refused to follow the orders of the Thirty, it is unclear whether he did it because he believed complying would be unjust or because complying would be impious, or both. He does not make it any clearer when he states, “[D]eath is something I couldn’t care less about, but that my whole concern is not to do anything unjust or impious.”⁹ Gregory Vlastos examines this in depth as he points to two doctrines that Socrates claims loyalty to in his actions. One is the reasoned argument; “I am the kind of man who listens to nothing within me but the argument that on reflection seems best to me,”¹⁰ while the other is the divine, which can be found in his explanation to the court of his pursuit of philosophy as being “enjoined upon [him] by the god, by means of oracles and dreams, and in every other way that a divine manifestation has ever ordered a man to do anything.”¹¹

9. *Ibid.*, 36.

10. *Ibid.*, 48.

11. Gregory Vlastos, *Socrates, Ironist and Moral Philosopher* (Ithaca: Cornell University Press, 1991): 157.

By openly committing himself to these two philosophies, reasoned argument and the divine, Socrates is accepting ideas that could be conflicting. Vlastos compares Socrates' rationalization of the gods with the way the Ionians rationalized the gods. The Ionians had a world in which the existence of "gods" was unnecessary, but if they existed, they would be "naturalized and thereby rationalized."¹² This merges the idea of god with ideas from science of nature, physiologia. Any Ionian gods would therefore have the same naturalistic limitations as man. Vlastos argues that Socrates rationalizes the gods in a similar manner, but instead of holding them to naturalist principles, he holds them to the same moral principles as man.

Vlastos' most convincing piece of evidence for this comes from book II of the *Republic*. While it is likely that the *Republic* was written in Plato's middle period, in which it is generally thought that Plato began expressing his own ideas rather than those of Socrates, and while what follows this passage is clearly Platonic, Vlastos points out that this section does not use "any premises foreign to the thought of the earlier dialogues."¹³

"What about what is good? Is it beneficial? [Yes.] So it is the cause of doing well? [Yes.] What is good is not the cause of all things, then. Instead, it is the cause of things that are good, while the bad ones it is not the cause. [Exactly.] Since the gods are good, they are not – as the masses claim – the cause of everything. Instead, they are a cause of only a few things that happen to human beings, while of most they are not the cause. For good things are fewer than bad ones in our lives. Of the good things, they alone are the cause, but we must find some other cause for the bad ones, not the gods."¹⁴

12. *Ibid.*, 159.

13. *Ibid.*, 163.

14. Plato, *Republic*, trans. C. D. C. Reeve (Indianapolis: Hackett, 2004): 60.



The Piety of Escape

“Socrates, no metaphysician, sticking to his own last, the moralist’s, taking the fact of such causation [of the good] for granted, is content to do no more than clamp on it moral constraints, reasoning that since the gods are good, they can only cause good, never evil.”¹⁵ The Greek gods were often credited with superior wisdom, and Vlastos argues that Socrates held this to be true, too. For Socrates, however, this superior wisdom was “not theoretical, but practical,” and because of the way he has rationalized the gods this “wisdom binds gods no less than men.”¹⁶

Because Euthyphro ended in *aporia*, a Socratic account of piety may seem out of reach. Despite this, to define Socratic piety, Vlastos believes we need only to look at the evidence found in *Euthyphro* and the *Apology*. Upon hearing Euthyphro characterize piety as “the godly and pious is the part of the just that is concerned with the care of the god,” Socrates asks, “To the achievement of what aim does service to the gods tend?”¹⁷ Vlastos points out that Socrates tells Euthyphro how important a question this is when he states, “if you had given that answer, I should now have acquired from you sufficient knowledge of the nature of piety.”¹⁸ While Euthyphro still does not have an answer for Socrates after this comment, by examining the *Apology* we can infer what Socrates’ own response to that question would be. Vlastos argues that Socrates saw his own work as pious, as it was “his own service to the god” by “summoning all and sundry to perfect their soul [...] at the god’s command.”¹⁹ This point of view is supported by Socrates’ claim in the *Apology* that he “was attached to [Athens] by the god [in order to] fulfill some such function,”²⁰ namely to talk to the people of Athens and “[persuade them] to care for virtue.”²¹ Taking these things into account, Vlastos believes that Socrates’ account of piety can be sum-

15. Vlastos, Socrates, *Ironist and Moral Philosopher*, 163.

16. *Ibid.*, 164.

17. Plato, *Five Dialogues*, 17.

18. Vlastos, Socrates, *Ironist and Moral Philosopher*, 174.

19. *Ibid.*, 175.

20. Plato, *Five Dialogues*, 35.

21. *Ibid.*

marized like this: “doing on the god’s behalf, in assistance to him, work the god wants done and would be doing himself if he only could.”²²

Applying this account of piety to (III) of the aforementioned inconsistent triad, we can assume that Socrates would have agreed that (III) is correct. Because he was working on the gods behalf, and because he had rationalized his god as being restricted by the same moral principles as his own, Socrates could not have justified doing something obviously impermissible like leading Leon of Salamis to his execution. With this in mind, there seem to be no good reasons to doubt (II) or (III). The only proposition left to be rejected is (I). If Socrates denies that it is impious to bring violence to bear against your country or father, then what does he actually believe? I have already shown that Socrates believes that Euthyphro’s prosecution of his father is impious, so it is impious in at least some cases to bring violence to bear against your country or father. Therefore, Socrates’ actual position must admit both cases in which it is impious and cases in which it is not impious to bring violence to bear against one’s country.

Because of this, there must be at least one significant difference between Euthyphro’s case and Socrates’ case with the Rule of Thirty. In the Rule of Thirty case, Socrates objects to what he is being ordered to do because it is obviously unjust. He states that he would rather be “on the side of law and justice” than take part in “an unjust course.”²³ “That government, powerful as it was, did not frighten me into any wrongdoing.”²⁴ While it is not completely certain that what Euthyphro is doing is impious, there is no reason given to believe otherwise. At the very least, we can say that a crucial difference between these two is that one involves engaging in an action which is certainly unjust, while the other involves engaging in an action which is possibly, but not necessarily, unjust. Since this appears to me to be the only significant difference between these two cases, Socrates’ position is likely to be that it is not impious to bring violence to bear

22. Vlastos, *Socrates, Ironist and Moral Philosopher*, 175.

23. Plato, *Five Dialogues*, 36.

24. *Ibid.*, 37.



The Piety of Escape

against your country just in case you would be doing something obviously unjust by doing otherwise.

Using this as a premise to replace Socrates' original inconsistent belief, we can now reexamine the original argument about why he should not attempt to escape prison:

- I. It is impious to bring violence to bear against your country unless doing otherwise would be engaging in something obviously unjust.
- II. To escape prison would be to bring violence to bear against your country.
- III. To escape prison would be impious.

The conclusion in this adaptation of the original argument does not necessarily follow from the first two premises. If it can be shown that Socrates would be doing something obviously unjust by "doing otherwise," namely staying in prison, there would be two significant implications. The first is that Socrates would be engaging in an unjust action by remaining in prison. The second is the rejection of the conclusion, thus demonstrating that escaping from prison would not be an impious form of bringing violence to bear against one's country.

The impiety of Socrates escaping prison depends on whether or not he would be doing something obviously unjust by remaining there. Crito gives several arguments in favor of this. The first is that it is unjust for Socrates to "give up [his] life when [he] can save it."²⁵ This is not always true. If I am on a sinking ship and can save myself but only by tipping over and killing a lifeboat full of people, it would be unjust to save my life. Another argument from *Crito* is that Socrates is doing something unjust by doing something which would "hasten [his] fate as [his] enemies would hasten

25. *Ibid.*, 48.

it.”²⁶ This does not provide a good enough reason to believe that Socrates is doing something unjust by remaining in prison because there are examples which show that a just action is one which hastens someone’s fate as his enemies would. For example, a prisoner who has been justly and rightly convicted but has escaped authorities would be doing a just thing by turning himself in, while at the same time also hastening his fate as his enemies would.

At this point, it could be reasonably objected that if Socrates believed it is worse to do wrong than to be wronged, thus my argument has yet to go far enough. However, one could respond to this objection by examining Crito’s argument that Socrates would be aiding the city in doing something unjust, and therefore doing something obviously unjust. If one accepts that supporting an unjust action is wrong, and he also accepts that the sentence is unjust, then he must also believe that submitting to the sentence is unjust just in case there are practical alternative options available.²⁷ Socrates clearly believes that Athens has done something unjust, which he expresses by his warning the city of its oncoming guilt for killing him,²⁸ and by pointing out that his accusers are “condemned by truth to wickedness and injustice.”²⁹ Crito argues that sitting in prison hastens the unjust sentence and helps in carrying it out, thereby making Socrates a guilty party in his own death sentence. If Socrates would have attempted escape, it seems clear that there would have been very little resistance by the city. The amount of harm done, if any, is almost certainly less than the amount of harm done by complying in the murder of a man who was unjustly convicted. Thus, Crito’s argument makes it clear that Socrates would be doing something obviously unjust by remaining in prison.

Upon reaching this conclusion, we can reexamine Socrates’ argument for escape one last time with the new premise that Socrates is doing something obviously unjust by remaining in prison.

26. Ibid.

27. Ibid.

28. Ibid., 42.

29. Ibid.



The Piety of Escape

- I. It is impious to bring violence to bear against your country unless doing otherwise would be engaging in something obviously unjust.
- II. To escape prison would be to bring violence to bear against your country.
- III. Socrates would be doing something obviously unjust by remaining in prison.

The conclusion that could logically be deduced is this:

- IV. Socrates' escape from prison is not a case of impious violence against one's country.

That does not confirm Socrates' escape from prison is not impious, but only that it is not impious in the sense that it is impious to bring violence to bear against your country. It could still be a possibility that his escape is unjust for other reasons. However, it seems that there are no other reasons given in Plato's writing, and we can therefore conclude that Socrates' escape from prison would not be impious.

I would like to thank the anonymous referees for raising the objection that it is possible that Socrates respects procedural justice. If Socrates believed that one has an obligation to a just government to obey its decisions, then this would account for the significant difference between Socrates' refusal to cooperate with the Rule of Thirty and his unwavering decision to comply with his death sentence. While this may seem like a reasonable interpretation of Socrates' behavior at first glance, evidence from the *Apology* suggests otherwise. Had the Athenian court said to Socrates, "We acquit you, but only on condition that you spend no more time on this investigation and do not practice philosophy, and if you are caught doing so you will die,"³⁰ Socrates would have had to respect and comply with

30. *Ibid.*, 34.

this decision. Surely, this is no less just than his death sentence. Despite this, Socrates confesses that he would never follow such a decision: "As long as I draw breath and am able, I shall not cease to practice philosophy."³¹ This provides a contradiction for any claim that Socrates' respect for the procedural justice of Athens is what held him in prison.

Why would Socrates comply with a death sentence but not with an order to cease practicing philosophy? If his account of piety is "doing on the god's behalf, in assistance to him, work the god wants done and would be doing himself if he only could"³² requires that he practice philosophy, and to escape prison would mean he would never be allowed to practice philosophy, then it is possible that Socrates could never justify this type of life as not being impious. Because "the unexamined life is not worth living,"³³ perhaps death appeared to be the best alternative. In spite of this, I believe that Socrates is doing something much worse by accepting his death sentence: giving up on his gods-given mission. While it is likely that he would be put to death for acting in a way that he believes is pious, surely it would be more pious to escape prison and continue practicing philosophy than to accept death and never practice philosophy in Athens again.

I have thus shown why I believe Socrates' argument to remain in prison to be unsound. One premise of it appears to be inconsistent with Socrates' own beliefs and actions. Using Socrates' account of piety, if all actions which brought harm to bear against your country were impious, then Socrates would have been doing something impious by not obeying the Rule of Thirty order to capture Leon of Salamis. However, it seems clear that Socrates also believes that Euthyphro was doing something impious by prosecuting his father. Modifying the premise to accommodate these two seemingly contradictory cases, it appears his argument that escaping prison is impious was insufficient. Because Socrates believed his sentence

31. Ibid.

32. Vlastos, *Socrates, Ironist and Moral Philosopher*, 175.

33. Plato, *Five Dialogues*, 41.



The Piety of Escape

was unjust, he could not reasonably deny Crito's argument that he is engaging in something unjust by remaining in prison. After examining an objection about the nature of Socrates' reasons for remaining in prison, I have not been given cause to doubt my position. I have ultimately concluded that Socrates would have not been doing something impious by escaping prison. ❖

Certainly Uncertain: Nietzschean Pessimism for an Optimistic World

Stephen Bailey

ABSTRACT: In this paper, I contrast pre- and post-Socratic Greek thought, particularly with respect to Apollonian optimism and Dionysian pessimism. I show how Socrates' judgment of a "right" way of living undermined Greek pessimism and was the first step towards modern scientific optimism, the belief that the world can be understood. I then argue that new developments in quantum physics make this optimism untenable, and I finally assert that Nietzschean pessimism is a coherent and beneficial metaphysical perspective.

Introduction

...it is always a metaphysical belief on which our belief in science rests – and... even we knowing ones of today, the godless and anti-metaphysical, still take our fire from the conflagration kindled by a belief a millennium old, the Christian belief, which was also the belief of Plato, that God is truth, that the truth is divine... But what if this itself always becomes more untrustworthy, what if nothing any longer proves itself divine, except it be error, blindness, and falsehood; what if God himself turns out to be our most persistent lie?

– Friedrich Nietzsche, *The Gay Science*



Certainly Uncertain

It is hard to be pessimistic in a society where cell phones have made physical separation meaningless, where designer medicines are gradually eradicating disease, where new technology has allowed the sequencing of entire human genomes. The contemporary reader of Friedrich Nietzsche may be quick to admit that God is dead – but lo, Science certainly lives! The seemingly unstoppable progress of knowledge appears to give a strong reason for believing that the human situation is improving; that we are coming out of the darkness of not-knowing; that maybe one day, mankind will finally be able to put down its telescopes and spectrometers, take a deep breath, and relax, having attained full knowledge of the world. This intellectual optimism can be a stable and comforting pillow on which to rest a man's existential burden. Yet, on the highest peaks of physical science, the scientist can get vertigo. She may begin to think that maybe Nietzsche was right all along, that maybe it was not only the God of the Christians that is dead but also the god of Science. In this discussion, I seek to elucidate the historical foundations of scientific optimism; show that this belief, since the advent of quantum mechanics in physics, contradicts itself; and employ Nietzschean pessimism to give new hope to the scientific man.

The Olympic Synthesis

Nietzsche was formally a philologist, an expert in classical civilizations, and his philosophy is grounded in the historical development of human culture. He had a special admiration for the Greeks, and his first book, *The Birth of Tragedy*, extensively covers the changes in Greek thought between the 6th and 4th centuries B.C.E. He pays particular attention to their devotion to Apollo and Dionysus, two of the twelve Olympians. In this dichotomy, Apollo, "the god of all plastic energies," and the "soothsaying god," represents sculpture, individuality, and science; Dionysus, the

god of wine and revelry, represents music, oneness, and mysticism.¹ Nietzsche contends that the tension between the two modes is fundamental to human existence and that the early Greeks were remarkable for not only recognizing but embracing both of these forces in their culture. This synthesis was artistically embodied in the great Greek tragedies of Aeschylus and Sophocles, plays that managed to affirm life despite meaninglessness. Some pre-Socratic philosophers like Heraclitus and Democritus metaphysically reinforced this ordered-yet-chaotic approach to the universe.

To understand the peculiarity of this balance, I must give a brief overview of Greek religion. These early humans, Nietzsche says, “knew and felt the terror and horror of existence,” so they created the Olympians to exist alongside them.² To the modern reader, the gods can seem almost silly – they are more adulterous and pugnacious than nearly any mortal – and one may wonder what their religious significance could really have been. In Nietzsche’s analysis, it was precisely the fallibility of the Olympians that made them holy to those Greeks. The Olympians gave a divine justification for human emotions, desires, and follies. Their pantheon honored the whole breadth of human experience: there was a god for revenge, for motherhood and family, for work, for wine and revelry.³ Though the Greeks created an Apollonian form (i.e. a mythology) for their instincts, there is a Dionysian acknowledgement of the significance of all emotions.

Apollo, the “shining one”, is the god of light, and in the Nietzschean opposition, he stands for the art of reasoning and speculative philosophy.⁴ The Apollonian ideal is immensely attractive to mankind for its beauty and perfection. For example, the perfect Greek sculpture seems to exist independently of all else: it is proportional, symmetrical, untainted but it is also untouchable. For the man entrenched in harsh reality, Apollonian

1. Friedrich Nietzsche, *The Birth of Tragedy*, in *Basic Writings of Nietzsche*, ed. Walter Kaufmann (New York: Random House, 2000): 36.

2. *Ibid.*, 42.

3. Friedrich Nietzsche, *The Genealogy of Morals*, in *Basic Writings of Nietzsche*, ed. Walter Kaufmann (New York: Random House, 2000): 465.

4. Nietzsche, *The Birth of Tragedy*, 35.



Certainly Uncertain

constructions like Justice, Wisdom, Liberty, and the Good are tantalizing. However, it is one thing to recognize perfection and another to believe that it is humanly attainable. Dionysian rituals, such as orgies, demonstrated that the pre-Socratic Greeks tempered this rational understanding with an emotional mysticism and a belief that the world could not be completely understood by forms and language.

Any balance between the two forces is fragile, as each strives for dominance over the other. Leaning too far towards Dionysus tends to weaken a society's moral and social structure; moving the opposite way makes society rigid and more prone to destruction from emotional rebellion. In the 5th century B.C.E., Greek civilization hit its cultural and political Golden Age, and the two forces were synthesized into a culture so brilliant that modern civilization still looks back in awe. Like a star that burns brightly but quickly, the brilliant culture could not radiate forever, and a devastating civil war left Greek culture reeling at the end of the 5th century B.C.E. Some of the brightest Greek minds began to rebel against the sophistry of powerful aristocrats, and a stable but rigid Apollonian society began to solidify.

The Socratic Revaluation of Values

The “demon” that dissolved the delicate synthesis was Socrates.⁵ His three maxims – “Virtue is knowledge; man sins only from ignorance; he who is virtuous is happy” – would reshape Greek culture and set the stage for the next 2,400 years of Western intellectual development.⁶ Socrates does not acknowledge a mystical, Dionysian element to life, only illusory or confusing ideas, and he stakes his life on the existence of one True way of living. With this judgment, Nietzsche believes that Socrates has spilled the “magic potion” of the Greek genius into the dust.⁷ A dis-

5. *Ibid.*, 82.

6. *Ibid.*, 91.

7. *Ibid.*, 88.

connect has been introduced between instinct and the right way of living. No longer must a person be of noble birth or do heroic deeds to be good; he need only be knowledgeable – and incorrect knowledge meant that a person was not good.

Socrates' central conflict was with the sophists, and it is easy for the modern reader to side with Socrates and say, "Those sophists! Where is their integrity? Where is their pride?" There lies the problem: the Sophists' pride was actually in their sophistry. In Plato's dialogues, Socrates questions all those around him and concludes that they know nothing, that they operate "only by instinct."⁸ What is Greek piety? Instinct. Greek justice? Convention. Many of Socrates' conversations conclude in confusion, and in fact, Socrates admits that he truly knows nothing. Nevertheless, he advocates the search for sure knowledge as the path to virtue. His disciple, Plato, would make this valuation even more extreme by stating that true existence could only be attained through contemplation of the forms. In doing this, he affirms a world in which only absolute knowledge can make life understandable and provide comfort to the existing human. Consequently, he rejects the Dionysian wisdom that acknowledges ignorance yet embraces physical, human existence.⁹

Nietzsche asserts that the natural human instinct is creative-affirmative, while the consciousness is critical. In Socrates, the instinct became critical and the conscious creative. Here, "the logical nature is developed[...] excessively."¹⁰ Nietzsche should not be misunderstood as saying that Socrates' excessive logical instinct was an abomination; rather, he was a re-valuator of values, an übermensch. It took a character of his enormous intellectual strength to topple Dionysus and pursue rationality to his death. After his martyrdom, men flocked to try the Socratic method themselves and found great pleasure in it. Nietzsche acknowledges that no craving is

8. *Ibid.*, 87.

9. Friedrich Nietzsche, *The Gay Science*, trans. Thomas Common (New York: Barnes & Noble Books, 2008): 177.

10. *Ibid.*, 88.



Certainly Uncertain

more violent than the one a person gets for the thrill of a Socratic insight, because it embraces the “whole world of appearances” and makes life more tolerable.¹¹ However, it is still an embracing of illusions, and the pleasure of the devotee depends on the belief that the nature of things is fathomable. This Socratic faith would eventually generate the scientific method, and despite the new instruments and approach, the initial optimistic insight remains intact: if we question enough, if we search enough, we can find Truth in this world.

The Scientific Fixation

In theory, science is value-free. An experiment is supposed to be repeatable by anyone, anywhere, and at any time, thus preventing inaccurate or biased conclusions from being made by overeager scientists. Like Socrates, a scientist begins in a state of doubt, researching and questioning her predecessors. She must remove herself from the process as far as possible, and “all [of her] convictions must condescend to the modesty of a hypothesis.”¹² The honest scientist never hopes to come up with facts, only supported hypotheses. However, human beings are not good at building a world made solely of conjectures, no matter how well supported. Every worker approaches a problem with a vested interest in solving it – for power, for knowledge, for money, for grades – and she also brings the belief that it can be solved. When a theory is supported with enough evidence, humans intuitively accept it as fact into their world, regardless of how much “science” disapproves of the leap. After enough intellectual victories, scientists become particularly vulnerable to this error, and, adopting the hubris of the optimist, reduce themselves into theoretical men.

We say “reduced” because Nietzsche affirms both the Apollonian and Dionysian aspects of existence, and the theoretical man embraces only one of these. This man “in whom the scientific instinct... for once blossoms

11. Nietzsche, *The Birth of Tragedy*, 97.

12. Nietzsche, *The Gay Science*, 176.

and blooms to the end is certainly one of the most precious instruments there; but he belongs in the hand of one more powerful.”¹³ Nietzsche does not deny that there is a certain innate appeal to the search for knowledge; indeed, he sees that very desire in himself.¹⁴ However, the pleasure that drives men to believe that everything is knowable, that they might in the end close the system, is destructive. He who finds his identity in such a belief is a “man without substance,” and his soul is a mere mirror of his metaphysical conviction that the universe is knowable.¹⁵

The tendency for scientists to fall into optimism is troubling to Nietzsche, and he sees in the scientific method an inherent decadence. He calls it, “most recent and noblest refinement” of the ideal which humans strive for in order to negate themselves and feel fulfillment.¹⁶ He exclaims, “How often the real meaning of all this [scientific industry and craftsmanship] lies in the desire to keep something hidden from oneself!”¹⁷ He understands that science has technologically pushed civilization in incredible ways, but he also knows that we are still the same animals, torn by the same impulses as the Greeks: to destroy and to create. By negating the Dionysian aspect of life, the optimist denies herself life in the real, existing world, and consequently sets herself up for destruction. In 1871, Nietzsche wrote: “Science, spurred by its powerful illusion, speeds irresistibly toward its limits where its optimism, concealed in the essence of logic, suffers shipwreck.”¹⁸ In 1925, the scientific enterprise crashed on that shoreline.

13. Friedrich Nietzsche, *Beyond Good and Evil*, in *Basic Writings of Nietzsche*, ed. Walter Kaufmann (New York: Random House, 2000): 317.

14. Nietzsche, *The Gay Science*, 177.

15. Ibid.

16. Nietzsche, *The Genealogy of Morals*, 582.

17. Ibid., 583.

18. Nietzsche, *The Birth of Tragedy*, 97.



The Advent of Quantum Mechanics

Physical scientists at the turn of the twentieth century were thoroughly optimistic. Some even thought that the chief goals of physics, at least, were nearly attainable. As new avenues, chiefly concerning light and atomic structure, were explored, several anomalies arose that Newtonian mechanics could not explain.¹⁹ Several enormous leaps were then made by the great thinkers of the time: Albert Einstein discovered that energy and mass were inherently related and also postulated that light could be understood as both wave and photon; Max Planck derived a quantum theory of energy; and Niels Bohr developed a new atomic theory of the hydrogen atom. It was not until 1925 and 1926 that German physicist Werner Heisenberg and Austrian physicist Erwin Schrödinger electrified the scientific community with a series of models comprising what is now known as quantum mechanics. It was this breakthrough that revitalized physics and has led to a wealth of new discoveries about the nature of the universe.²⁰ For the scientific optimist, however, this advance produces a very unsettling situation.

Schrödinger and Heisenberg found that tiny masses like electrons, being much smaller than their larger proton and neutron counterparts, operated in strange ways. Like light, electrons function as both a particle and wave. Schrödinger developed a mathematical equation that was able to describe some properties of an electron in an atomic system, and Heisenberg shortly thereafter developed his “Uncertainty Principle,” which “states that there are ultimate limits to how exact certain measurements can be.”²¹ Specifically, when one wants to observe an electron at an instant, she can choose to either measure the position or the momentum, but she cannot measure both. This principle is especially important in the instance of an electron because it is so minuscule, but on some level, this principle holds

19. David Ball, *Physical Chemistry* (Pacific Grove: Thomson Learning, 2003): 253.

20. *Ibid.*, 270.

21. *Ibid.*, 280.

for every particle in the universe, meaning uncertainty is a fundamental aspect of physical existence. If basic properties like position and momentum have “ultimate limits,” and the scientific optimist believes that we can know with absolute certainty everything in existence, then it would seem that he is stuck in a contradiction.

There are at least two objections the optimist could make here. The first is that there may be an underlying reason for our uncertainty which we simply have not discovered yet. The optimist could reject the Uncertainty Principle as not being ultimately authoritative, believing that we will eventually discover a way to know both the position and the momentum of an electron at one moment. This is quite possible, but staking one’s worldview on a prediction that contradicts the current best evidence seems suspiciously like the proverbial ostrich burying its head in the sand. Even the brightest scientists are not able to explain many of the phenomena found on a sub-atomic level, but they cannot deny that the phenomena exist and that they defy our best understanding of reality. A more realistic approach than dogmatic insistence would be to approach these discoveries with curiosity and humility. Instead of asserting “We will know,” we should ask, “Can we know?”

The second objection is that the uncertainty inherent in quantum mechanics becomes statistically insignificant as the system being observed becomes much larger. For example, there is no reason to question whether we can actually know if a car is on the road and moving at 100 miles per hour. (Even the best physicist could not weasel his way out of that ticket.) Classical physics still describes the world of sensory perception adequately. Then the optimist could say that science, even if it cannot be absolutely certain, can be probabilistically certain about reality. This is true: we have no practical reason to doubt, on account of quantum mechanics, whether the car will start in the morning or if the floor will support our next step. However, just because these principles do not affect our daily lives now,



Certainly Uncertain

does not mean they will not in the future. The lives of many physicists throughout the world are centered on the problems of quantum mechanics, and it is not impossible that new technology affected by this inherent uncertainty will infiltrate homes in the future. If the quantum mechanical picture of reality is accurate, then our metaphysical understanding of the world should reflect it as much as possible.

The scientific, Apollonian understanding of the universe, as it has delved deeper and deeper into the structure of matter, has undermined its own belief in itself. Without the cushion of scientific certainty, many optimists might resign themselves to despair and nihilism. Nietzsche, however, has proposed another way out, calling for a return to a time before Socrates and the scientific fixation.

Pessimism as the Cure to Optimism

The primary difference between the pre-Socratic Greek and the post-Socratic scientist is the lack of homage paid to Dionysus by the modern man. Those Greeks had a Dionysian pessimism which enabled them to see the world in its cruel reality, yet affirm life to a nearly unsurpassed extent. To adhere to this pessimism, one must be strong enough to learn the “art of this-worldly comfort” and of laughing in the face of reality. In 1925, Heisenberg revived the possibility of Dionysian revelry by giving it some scientific credibility, and today the cultural landscape is fertile for a new Apollonian/Dionysian synthesis.²² Before I expand on what this might look like, we must come to a better understanding of Nietzsche’s Dionysian pessimism.

Dionysian pessimism is neither a moral nor metaphysical code. It is a prescriptive attitude to help cope with the chaos and lack of universal meaning in the world. Scientific optimism had given man a pleasurable, tolerable world to enjoy and control. The pessimist must learn to live in a

22. Ball, *Physical Chemistry*, 280.

much less comfortable world in order to live a much more genuine life.²³ However, he does not need to reject scientific progress outright. He should acknowledge that “without a constant falsification of the world[...] man could not live.”²⁴ It is the theoretical man’s will to absolute truth that makes the implosion of scientific optimism so painful; for the pessimist who rejects that initial value judgment, this blow is hardly felt! She understands that the human-dependent illusions of mathematics, science, and logic are not fundamental to human existence and that by not setting them up as gods, she is free to interact with them as tools. The Apollonian instinct is not harmful if it is controlled by Dionysian wisdom. Conversely, Dionysian urges are not nihilistic if tempered by personally meaningful Apollonian constructions.

Balancing the Dionysian element with logic is critical, as it is easy for the unanchored man to slip into a destructive state of mind. The siren’s call of nihilism can particularly affect the former optimist, but for the constructive pessimist, the knowledge that the world is chaotic should merely lead one to constantly qualify her goals. We no longer have the authority to condemn self-centeredness, but pessimism does not necessarily insist that one should only tend one’s own garden; it insists rather that one should “[know] the limits to one’s actions, however ambitious.”²⁵ Just as it frees us from a universal purpose, so does it free us from enslavement to the past. The pessimist believes himself to be in a constant state of transformation and does not worry about the past so much as he looks towards the future.²⁶ Nietzsche’s pessimism ultimately motivates men to localize their lives and to create their own meaning in the context of this very real world.

23. Joshua Dienstag, “Nietzsche’s Dionysian Pessimism,” *The American Political Science Review* 95, No. 4 (2001): 933, <<http://www.jstor.org/stable/3117722>> (April 28, 2010).

24. Nietzsche, *Beyond Good and Evil*, 202.

25. Dienstag, *Nietzsche’s Dionysian Pessimism*, 936.

26. *Ibid.*, 935.



Conclusion

Pre-Socratic Greek civilization developed in a world that could not insulate itself from the harsh realities of the world, so it carved a place in its culture for them, even while creating architecture, political institutions, and artwork that rivals that of any modern society. This society produced Socrates, whose extreme push for rationality provided a way of looking at the world which could not be resisted by men of knowledge. As the scientific disciplines developed in the Western world, it became increasingly easy to become overconfident in man's ability to explain the world, and consequently, those who adopted this belief forfeited their appreciation of the Dionysian aspects of existence. New discoveries in quantum science, however, have made scientific optimism untenable, and an admission of uncertainty seems inescapable.

History has led us to the point where a new synthesis may be possible: the Dionysian scientist. Such an individual would be a new and brilliant star – an artistic Socrates. Humankind needs an outlet for the tragic despair that arises out of the knowledge of his relationship to the world – so he must be an artist. Humankind must also “become the best students and discoverers of all the laws and necessities in the world” – so he must be Socratic. The path to this man lies directly through scientific endeavor. Only through such work can he attain the instruments and tools he needs to create himself, this new man... Only for the accomplishment of such a task would Nietzsche exclaim, as he does in *The Gay Science*, “Three cheers for Physics!” ❖

How We Are Moral: Benevolence, Utility, and Self-Love in Hobbes and Hume

Jenna Kreyche

ABSTRACT: In this paper, I reconstruct Hobbes' theory of self-love. I then examine Hume's arguments that (i) self-love does not properly account for moral behavior and (ii) self-love is unnecessary for moral theory. I argue that Hobbesian self-love can account for both of Hume's objections. Further, I use an analysis of Hobbes' *Deliberation* to show, contra Hume, that self-love does not entail a lack of intention in moral action.

A Brief Overview Of Hume's Moral Philosophy

David Hume grounds his moral theory in the benevolent nature of humans, which he supports with an argument against the theory of self-love. Self-love, which can also be referred to as psychological egoism,¹ "accounts for every moral sentiment by the principle of self-love," or the ultimate concern with one's own happiness and preservation.² His moral theory is based on utility, which means that all social virtues are defined by their usefulness

1. Hume uses the term "self-love" in writing, but for the purposes of this paper I will use the synonymic term "psychological egoism" interchangeably with self-love.

2. David Hume, *An Enquiry Concerning the Principles of Morals*, Ed. Tom L. Beauchamp, (New York: Oxford University Press, 1998): 109.



How We Are Moral

to the individual or the society. Hume posited that morality is determined by man's naturally occurring sentiments rather than reason because pleasant sentiments indicate traits that are useful, such as prudence, courage, kindness, and honesty.³ These virtues inspire "our approbation and goodwill" because they "contribute to the happiness of society."⁴ However, our approval of these virtues, and in turn our concern with the welfare of society, is motivated by benevolence because the "principles of humanity and sympathy enter so deeply into all our sentiments."⁵ In other words, man values and praises what is beneficial to him and his fellow men because of his love for humanity. Hume's theory suggests that "desires and aversions themselves are the main motivating forces" of moral behavior because of their role in moral evaluation.⁶ However, unlike Hobbes, Hume posits that these desires and aversions are prompted by external events according to man's benevolence, including more specifically sentiments of sympathy, humanity, and natural concern for the welfare of others.

Hobbes' Moral Philosophy

Both Hume and Hobbes believe that moral virtues are based in utility and underscored by the passions. However, their views diverge when it comes to accounting for the mechanisms that motivate adherence to utility, i.e. how and why the passions are related to moral behavior. Although Hobbes does not make an explicit argument for self-love, his argument promoting the notion can be presented as follows: notions of desire and pleasure proceed from nature or experience, and desire and pleasure necessarily relate to what is good for the preservation and happiness of the individual; and since what we consider to be moral virtues are ultimately derived from notions of expected desire and pleasure, it must follow that moral virtues

3. *Ibid.*, 160.

4. *Ibid.*, 109.

5. *Ibid.*, 117.

6. Alan H. Goldman, *Moral Knowledge* (New York: Routledge, 1988): 55.

are ultimately rooted in what is good for the preservation and happiness of the individual, i.e. what is in the individual's self-interest. On a slightly different but equally important note, Hobbes' argument promoting self-love in moral action can be presented as follows: moral actions are voluntary actions, and "all voluntary actions aim at happiness or at some good to the agent."⁷ Voluntary actions necessarily arise from "Appetites to the thing propounded," as well as, "Aversion, or Feare of those consequences that follow the omission."⁸ When actions are motivated by appetites or aversions, the object of the action is "[an individual's] own Good."⁹ It follows then, that the object of all moral actions is personal good.

Like Hume, Hobbes posits that moral decisions rely on the passions. Hobbes explains that "whatsoever is the object of any man's Appetite or Desire" is that "which he for his part calleth Good."¹⁰ In other words, positive emotions and the desires associated with them lead man to a conception of what is moral, and likewise aversions to what is immoral, because values like "Good" and "Evil" are "not to be found in the objects themselves."¹¹ The passions are guides towards self-preservation in that "what men feel strongly about or desire strongly is what helps them to survive."¹² In this way, Hobbes can be seen as a subtle egoist. However, it should be noted that Hobbes is not advocating an image of man as guided blindly by his emotions; according to Goldman, "the interpretation of good as the object of rational desire is coherent with Hobbes' overall moral theory and psychology."¹³ Desires, or appetites, can be rational or irrational, and a rational desire is one that "[results] from a precedent Deliberation,"

7. Goldman, *Moral Knowledge*, 24

8. Thomas Hobbes, *Leviathan*, ed. Richard Tuck (New York: Cambridge University Press, 1991): 36.

9. *Ibid.*, 83.

10. *Ibid.*, 32.

11. Goldman, *Moral Knowledge*, 20.

12. Richard Tuck, "Hobbes on Skepticism and Moral Conflict," in Hobbes, *Leviathan*, ed. Tuck, 341.

13. Goldman, 30.



How We Are Moral

which is essentially a sequence of alternating appetites resulting in the act of willing.¹⁴

Hobbes and Hume agree that virtues are such because of their utility.¹⁵ For Hobbes, moral virtues are derived from the state of nature and are considered good, or moral, because “the practice of them conduces to peace, which every man must acknowledge to be good” by virtue of its usefulness to both the individual and other members of society.¹⁶ According to John Kemp, Hobbes did not seek to simply “derive my moral obligation to keep my promise from the purely selfish consideration that I shall be worse off if I do not.”¹⁷ Rather, in the context of Hobbes’ commonwealth, there is an “identity of interest” between the individual and the public insofar as the security of the individual depends on the security of others;¹⁸ likewise, the happiness of the individual is dependent on the happiness of others, and because happiness cannot be cultivated without security, it must sometimes be the case that an individual sacrifices immediate pleasure for the “Expectation [of happiness or security], that proceeds from foresight of the End.”¹⁹

Hume’s Argument Against Self-Love: Scope

Hume allows that self-love is a principle of human nature.²⁰ However, in Hume’s philosophy, self-love is not a moral principle, and in fact “competes with moral principles such as benevolence.”²¹ According to Hume, thinkers like Hobbes deduce the motive of self-love because they observe that virtues have “a tendency to encrease the happiness [of man-

14. Hume, *Enquiry*, 109.

15. John Kemp, *Ethical Naturalism: Hobbes and Hume* (New York: St. Martin’s Press, 1970):19.

16. *Ibid.*, 19.

17. *Ibid.*, 18.

18. *Ibid.*, 20.

19. Hobbes, *Leviathan*, 33.

20. Tom L. Beauchamp, “Introduction,” in Hume, *An Enquiry Concerning the Principles of Morals*, ed. Beauchamp, 24.

21. *Ibid.*, 24.

kind]”, whereas vices contribute to “the misery of mankind.”²² From these observations, egoists then discern the existence of a “union of interest” between the public and private as simply “modifications of self-love.”²³ Hume agrees with philosophers like Hobbes that, “the interest of the individual is, in general, so closely connected with that of the community,” but it does not follow from this that all interest for others is nothing but redirected self-love.²⁴ All humans have some degree of concern for public interest, and it is this aspect of human nature that Hume sees as problematic for an egoist account. Hume sees self-love as a theory that renders public interest, when separate from our own, “entirely indifferent to us,” which we can easily observe to be false.²⁵

It should be noted that Hume’s argument does not directly engage Hobbes, although his argument is likely a response to a theory of self-love like that presented by Hobbes. The principle of self-love that Hume argues against is not exactly the same principle that Hobbes advocates; however, the basic principle that all actions are motivated by self-interest is held in common. Hume’s argument against the version of self-love that he presents rings true in many ways. However, due to its narrow conception, by itself it does not seem to say much. Hobbes’ version of self-love is more robust than Hume’s, and when the two are pitted against one another, we can gain a greater understanding of the nature of moral action.

In essence, Hume’s argument against self-love can be divided into two main points: (a) the principle of self-love is not a moral principle because it is limited in its ability to account for moral behavior (and thus it is faulty); and (b) it is unnecessary for moral theory. First, Hume argues that self-love is limited in calling upon “the voice of nature and experience [which] seems to plainly oppose the selfish theory.”²⁶ For Hume, the

22. Hume, *Enquiry*, 108.

23. *Ibid.*, 108.

24. *Ibid.*, 108.

25. *Ibid.*, 109.

26. *Ibid.*, 106.



How We Are Moral

theory of self-love is faulty because (i) of instances in which the interest of the individual is either separate or wholly opposed to that of the public, “and yet we observe the moral sentiment to continue” in the interest of the public despite its opposition to private interest;²⁷ and (ii) spectators enjoy and praise the virtues or moral actions of others even though it has no effect whatsoever of the spectators themselves, and thus yields no benefit for the spectators.²⁸

Hume argues (i) with examples in which there is disunity between public and private interest. Self-love cannot account for a situation in which a mother, for example, sacrifices her own interest and well-being to take care of her sick child.²⁹ In this case, the private interest of the mother is at odds with the comparatively public interest of the child, and yet the mother’s moral sentiment leads her to attend to her child’s interest. Hume cites this as a case of benevolence divorced from self-love,³⁰ for it seems that compromising one’s health and wellbeing for that of someone else cannot be rooted in self-love.

In Hume’s second objection (ii) to the scope of self-love, he argues that “it can never be self-love which renders the prospect of [another’s virtuous character] agreeable to us, the spectators, and prompts our esteem and approbation,” because there is nothing in it for the spectator.³¹ While the virtuous agent enjoys the advantages of his character and the recipient of his actions enjoys their utility, the spectator gains nothing by means of utility or good moral character by his approbation of the virtuous agent’s actions. Rather, Hume claims that spectators enjoy moral behavior because of the “pleasing sentiment of sympathy and humanity” that the observation inspires due to man’s benevolent nature;³² it is this sentiment of sympathy that is necessary in order to qualify a “genuine moral evaluation.”³³

27. *Ibid.*, 109.

28. *Ibid.*, 119.

29. *Ibid.*, 167.

30. *Ibid.*

31. *Ibid.*, 119.

32. *Ibid.*, *Moral Knowledge*, 120.

33. Goldman, 63.

Although instances like the mother who sacrifices herself for her child in (i) ring true to human nature and provide a strong objection to self-love, when considered from a Hobbesian point of view, they can be related to self-love. For example, Hobbes claims that man endeavors to do what contributes to his good so far as he is able to anticipate.³⁴ The motivation of voluntary actions by appetites and aversions is a process described as follows: people develop an opinion “of the likelihood of attaining what they desire” through deliberation and act based on this likelihood.³⁵ In this case, the mother’s desire is the welfare of her child. The mother can be seen as acting out of self-love because “the satisfaction of desires for the welfare of others would count as contributing to an agent’s good on [Hobbes’] view.”³⁶ This is not to say that the mother is not acting morally, or that she lacks genuine concern for her child. In Hobbes’ view, benevolence is the “desire of good to another,” so the mother can indeed be said to have acted benevolently in this situation.³⁷

As for situations such as that described in (ii), which may seem to lie outside the realm of self-love, Hobbes’ self-love may account for them in terms of rational desires. For Hobbes, we approve of moral behavior even when we are not directly involved because we are always involved on some level. Our approval perpetuates the existence of virtues, which allow us as individuals to have security, and from there, happiness. Hobbes’ construction of moral rules, which might be compared to Hume’s virtues, holds that even people who lack a concern for others “should be rationally motivated to obey [moral rules]” regarding the welfare of others because it is in their best interest as members of a society.³⁸ Although approbation may not be a moral rule, the idea is that individuals are inclined to appreciate the moral behavior of others because it contributes to the welfare of

34. Hobbes, *Leviathan*, 83.

35. *Ibid.*, 33.

36. Goldman, *Moral Knowledge*, 25.

37. Hobbes, *Leviathan*, 33.

38. Goldman, *Moral Knowledge*, 25.



How We Are Moral

society as a whole, therein benefiting the individual as a member of that society.

Hume would likely object to this idea for its reduction and simplicity, and claim that it is contrary to the way people think: people do not consciously evaluate whether and how an outside event could possibly end up contributing to their own good before making a positive moral evaluation—they feel good about it immediately. Hobbes might respond to this objection by arguing that, as mentioned above, people desire and enjoy things that aim at their own good. This does not mean that an individual must be thinking about the way in which a given observation contributes to his personal good, but rather that it is his instinct to be attracted to it, and because a moral society is in an individual's best interest, the desire for a stranger to behave morally is rational.

Hume's Argument Against Self-Love: Necessity

In the second half of his argument, Hume claims that the theory of self-love is unacceptable because it is unnecessary for moral theory. For Hume, morality can be accounted for by benevolence and utility and it is useless and even detrimental to moral theory to “seek for abstruse and remote systems” to explain the motivation for moral behavior.³⁹ He claims that if the principle of self-love were true it would mean that “while all of us, at bottom, pursue only our private interest, we wear these fair disguises,” such as that of the friend, the lover, the helpful neighbor, and so on.⁴⁰ However, we do not need this kind of explanation, Hume argues, because benevolence is the “obvious and natural ... source of moral sentiment.”⁴¹

Here, Hume's conception of self-love is narrow, unbridled selfishness that necessitates indifference to the welfare of others. However, if we adopt a Hobbesian perspective of self-love, we can account for how we are

39. Hume, *Enquiry*, 164.

40. *Ibid.*, 164.

41. *Ibid.*, 109.

benevolent and why we care for and consider others. Hume “argues that passions and sentiments underlie our evaluations” and execution of moral behavior;⁴² Hobbes merely takes this idea a step further to argue that self-interest underlies the way in which the passions and sentiments drive us toward evaluation or executing a certain moral behavior the way we do. This step is necessary for moral theory because it provides an accurate explanation of the mechanism underlying moral behavior. Further, this step is important for Hume’s moral theory because it can account for the roles of utility and approbation where Hume’s Benevolence falls short.

According to Beauchamp’s reading of Hume, “benevolent acts are directed at promoting the good” of others as based in the sentiments of sympathy and humanity.⁴³ However, when this conception of human benevolence is considered in concert with utility and approbation, the moral waters are muddied. On one hand, humans are moral because they selflessly desire good for others; on the other hand, humans are moral because it is useful to others and yields approval, which is beneficial to the agent.

What, then, motivates moral behavior? If it is benevolence then the concepts of utility and approbation are unnecessary; if benevolence were the main principle of human nature that motivated moral behavior, moral theory would not need a concept of utility or approbation to perpetuate virtues because they would be automatic, and the moral agent would be indifferent to the approval of his peers because his end would simply be the good of others. However, if it is utility and approbation that motivate moral behavior, the mechanism that attracts the moral agent toward these must conduce to that of appetites and aversions. For if utility and approbation cannot be divorced from a conception of what it is to be moral, then it can be said that no moral act can be performed with utility and approbation in mind; and if all moral acts are performed with utility and approbation in mind, then all moral acts are performed with some idea of benefit

42. Goldman, *Moral Knowledge*, 53.

43. Beauchamp, “Introduction,” 35.



How We Are Moral

to the agent because of the obvious benefits of virtues. This does not mean that all moral acts must be performed with the consciously selfish motive of “good-for-me.” What it does mean is that humans desire good for others, and that satisfying this desire is good for the individual.

Here, it seems that Hume’s theory must give way to Hobbesian self-love. Even if the end of moral action “nowise affects us,” the approval or the expectation of approval associated with the action must ultimately become the end of a given moral action, and thus acting in the name of utility can only be motivated by self-interest.⁴⁴ Hume might respond by arguing that it would be false to reduce “all our concern for the public” to “a concern for our own happiness and preservation.”⁴⁵ However, what Hume is missing here is the all-pervasive idea of identity of interest. In Hobbes, utility and approval create the identity of interest between public and private sectors, so that any moral act performed in the interest of another is at bottom performed because it is also in the agent’s interest on some level.

Thus, it is not a coincidence that benevolent acts are met with approbation because we only know what virtues are by the public approval they elicit, so it cannot be said that we act virtuously without the expectation of approval. For all of these reasons, it seems that Hume’s principle of utility is more compatible with Hobbes’ self-love than Hume’s benevolence.

Rescuing Intention: The Role of Deliberation

The remaining objection Hume raises to self-love is that it removes intention from moral behavior. If self-love is the motive for all actions, then making a “moral” decision can be compared to seeking food when we have hunger or drinking water when we have thirst.⁴⁶ However, Hobbes’ self-love does not render all actions a result of crude, automatic instinct. First,

44. Hume, *Enquiry*, 109.

45. Hume, *Enquiry*, 108.

46. Hume, 168.

as discussed above, Hobbes holds that moral actions are voluntary actions, which implies the notion of purpose.⁴⁷ Second, Hobbes' concept of deliberation can account for the process involved in making moral decisions. Hobbes explains that deliberation is "the alternate succession of Appetites, Aversions, Hopes and Fears," and the consideration of the "good and evil consequences of the doing, or omitting the thing propounded."⁴⁸

Hobbes distinguishes between present pleasure or pain and expected pleasure or pain, placing much more importance on the expected for voluntary actions.⁴⁹ Weighing these sets of desires, which can be different or opposed to one another, is the work of deliberation, and is the activity that allows the agent to choose the rational desire, or that which is more closely connected with that agent's expected good.⁵⁰ It is true that, in some instances, self-love can compete with expressions of benevolence or other moral principles; in other words, an individual's selfish desires can interfere with or even directly work against the desires of others. However, for Hobbes, these are likely instances in which the individual chooses to follow his irrational desires, for choosing one's own selfish desires over or in spite of the desires of others is often detrimental to one's good. In situations like this, reputation, reciprocity, and other important social mechanisms that help to secure an individual's good are at risk, so the decision to act without concern for these is ultimately irrational.

Additionally, Hobbes admits that an individual can mistakenly choose the wrong course of action according to its "apparently good or evil consequences that one envisages in deliberating."⁵¹ These factors in mind, intention can be rescued from mere reduction to self-love because, in deliberation, the agent must consider present vs. future, self vs. others, rational vs. irrational, and so on. Moral decisions for Hobbesian self-

47. Goldman, *Moral Knowledge*, 24.

48. Hobbes, *Leviathan*, 35.

49. F.S. McNeilly, "Egoism in Hobbes," *The Philosophical Quarterly*, 16:64 (July, 1966): 197.

50. Hobbes, *Leviathan*, 36.

51. Goldman, *Moral Knowledge*, 20.



love cannot be accounted for by an automatic, self-serving instinct—they require many more dimensions. Thus, the concept of deliberation and its role in voluntary action allow for the importance of intention by placing evaluations and actions motivated by self-interest in the domain of moral consideration.

Concluding Thoughts

Hume and Hobbes are more compatible than what initially appears to be the case. When considered fairly, it seems that Hobbes' self-love is not a terribly shocking conclusion. However, Hobbes' theory is by no means flawless: can desire be divorced from self-interest? If it can, then Hume may be right and Hobbes' theory may find itself void of meaning. There are many more objections to psychological egoism that Hobbesian self-love may not be able to hold up against. However, Hobbes' self-love is only a non-normative theory, and even if self-interest is the motivation for human behavior, it does not necessarily imply that it ought to be. Hume's moral theory may be better suited as a normative theory that says that benevolence *should* be the basis of moral action. In any case, it seems that neither Hobbes nor Hume can be said to be completely right or completely wrong in their respective accounts of morality—neither philosopher helps us ascertain distinct moral rightness or wrongness. However, their theories do lend themselves to one another, and by finding a common ground between them, we may be getting close to a theory that encompasses both the accurate and the ideal in moral theory. ❖

Ruminations on Intermittent Existence

Anthony Adrian

ABSTRACT: Can objects exist, cease to exist, and then exist once more? I lay out three ways to think about intermittent existence (IE). The first section is based on intuitions. The second section will show that the intuitions are bolstered by the concept of supervenience. The final section will argue that the strongest way to think about IE, and about supervenience, is in terms of mereology, the theory of parts and wholes.

The Intuitive Appeal

I take my gun to a gunsmith for cleaning. Because this gunsmith is the best at what he does, he meticulously disassembles every part of the gun to clean. When finished, he reassembles the gun to return to me. This seemingly mundane example is just the kind of thing that gets metaphysicians foaming at the mouth. Some will say that this is a perfectly good example of a thing having a gap in its existence.¹ Others will say that the gun I receive back from the gunsmith is a different gun because it is not spatiotemporally continuous (STC).² Surely, if STC is the necessary and sufficient condition for identity over time, the firearm that the gunsmith returns to me is a different one.

Likewise, imagine a scenario where you possess a gun and there is a law that prohibits a person without a proper license to travel with an assembled firearm. In order to travel with the gun, she must disassemble

1. See Trenton Merricks, "There Are No Criteria of Identity Over Time," *Nous* 32.1 (1998): 106–124.

2. For the best formulation of this idea, see Peter van Inwagen, "Plantinga's Replacement Argument," in *Alvin Plantinga*, ed. Deane-Peter Baker (Cambridge: Cambridge University Press, 2007): 188-200.



Ruminations on Intermittent Existence

the firing mechanism for the duration of the trip and reassemble it once she has reached her destination. Upon reaching her destination and reassembling the gun, has she reassembled the same gun?

It would be singularly odd for this traveler to say, “Well I carried my gun to the car this morning, but we are never going to see that thing again. Instead, I have this new gun.” Intuitively, it is the very same gun before and after partial disassembly. Perhaps the tension with STC is not apparent because the disassembled parts are within proximity to each other, so instead imagine the gun being disassembled into five different parts and transported in five different cars, each driving five very different, and distant, routes to the same destination at which point the gun is reassembled.

It seems intuitively clear that the gun has intermittent existence. Compare the response to the gun scenario with the following similar examples:

- Our conversation was interrupted, but we resumed it the following week.
- The baseball game was called on account of the rain; it will be made up at a later date.

We would not say that a conversation or a baseball game existed while they were not taking place. Of course it might be challenged that the game and conversation are not material artifacts and so the analogy to the effect that the gun has a period of intermittent existence does not work. The idea is that events (like the baseball game) are just the kind of things that could be paused or interrupted with a period of intermittent existence. Material objects, on the other hand, simply go out of existence and stay out of existence; that is, objects can never come back from going out of existence. However, it is not obvious that the impossibility of intermittent existence is tied to the concept of “material object,” just as it is not obvious that the possibility of IE is tied up with the concept of “events.” These relationships

between concepts seem to be discovered from experience. At any rate, it seems clear that at an intuitive level we seem to believe in the idea of intermittent existence.

Though these all seem like fairly innocent intuitive responses, perhaps one may say that the gun never lacked existence. Rather, it existed for a while as a disassembled gun. At this point, intuitions cannot help us any longer since saying that the gun existed as a disassembled firearm is not absurd. The intuitional responses to this thought experiment prove inconclusive and we should seek guidance elsewhere. I find it in supervenience.

The Supervenience Story

Two paintings will be useful in illustrating supervenience. Suppose that two different paintings have precisely the same microphysical structure. Let us call the microphysical properties the paintings have the A-properties. Now when we ask, "Is one painting more beautiful than the other?" we are asking if the paintings differ in their aesthetic qualities which are macrophysical properties, or B-properties. Supervenience says that two things cannot differ in their B-properties without differing in their A-properties. Since the two paintings have the exact same microphysical structure, their aesthetic qualities will be identical. This is the nature of supervenient properties: larger, or macroscopic, qualities supervene on smaller, or microscopic objects arranged in a certain way.

It is easy to see that the B-properties supervene on the A-properties when they are arranged in a certain way. After all, we would hardly expect the two paintings to look exactly the same if they had the same kind of microscopic matter but arranged in a different way. Likewise, we would not say that two guns have the same B-properties such as texture, weight, engravings, etc. if the microscopic A-properties differed.

But, to return to the gun, it is going to be difficult to say that the gun exists as a disassembled gun for some time. (Recall that this is the objec-



Ruminations on Intermittent Existence

tion which led us to abandon our intuitional responses in the first place.) Consider a person who has yet to finish assembling her first hand-made gun. When a friend of hers sees her working on this gun and asks, "What have you got there?" it would not be unusual for the woman to reply, "This is my first gun." If this sort of reply were not wrong, it would mean that the B-properties of gunness (like the ones mentioned above, and, namely, the ability to fire a bullet) could supervene on unassembled parts of a gun. What are we to make of this consistency of an unassembled artifact with ordinary language? I am confident that this is just a convention of the English language; the woman does not honestly believe she has a gun in her hands. Rather, it is much more convenient for her to say she has a gun instead of saying the equally accurate statement, "These are the pieces to my first gun that I have been working on." I say the disassembled gun is not a gun because one of the features of artifacts is that it usually performs some particular sort of function.

An object is a gun if and only if it can function as a firearm. I take the "firing capacity" property to be essential for an object to qualify as a firearm. Note that the object does not have to be firing in order to be a gun. A historically important gun in a museum is not firing, but it does have the capacity to fire a bullet, even if we need to clean some rust off of it, and so it too qualifies as a gun. If an object lacks the capacity to fire a bullet, in what sense are we talking about a gun? Surely an object may resemble a gun, such as in the case of replicas and children's toys. However, people do not drive to a firing range to handle replica guns; they go to a firing range to practice using the gun's main feature, which is, of course, the ability to fire a bullet.

Note that the gun does not necessarily need to be made of a metal material. Plastic or carbon fiber, or gold or platinum, would do just as well if they were assembled to function as a gun. The fact that a gun could be made out of such different materials shows that what is absolutely essential to a gun is not the stuff it is made out of. Rather, the essence of a gun is

its ability to fire a bullet. Since I am arguing the function of the object is essential to its being a certain object, does a disassembled (or preassembled) gun really have the capacity to function as a firearm? It does when it is assembled, that much is clear, but in a state of disassembly there is nothing about the parts being within proximity to one another that lends the aggregate of gun parts the capacity to function as a firearm.

So my question to someone who holds that “gunness” can supervene on disassembled parts is this: what is it that the “firing capacity” property is supervening on? If a gun (or any other artifact) does supervene on its disassembled (or preassembled parts), when did the supervenience start? Surely there must have been a time when the gun did not exist. After all, the gun has not existed forever. When was it true that the gun did not exist? Perhaps the time when the gun failed to exist is when the smaller parts of the major parts of the gun (for example, the textured grips that are placed on the handle) were being assembled. But why? What is it about the larger, presumably more important, parts simply being close together that entails the existence of the gun?

One could insist that when a gun has been disassembled, the gun has not suddenly gone out of existence, but became a “disassembled gun” because the functional aspect of a gun is not essential to the identity of the object. But the gun and the disassembled gun are plainly not the same object. The gun has an additional property that the disassembled parts do not – the property of firing capability. And if two things do not have the same properties, by Leibniz’s Law (The Indiscernibility of Identicals), they cannot be identical.³

Ultimately, supervenience will only say something like, “The B-properties cannot differ without different A-properties.” Jaegwon Kim used to think that supervenience could explain the mind-body problem, but now he thinks that supervenience just is the problem stated.⁴ That is to

3. Wilhelm Gottfried Leibniz, *Discourse on Metaphysics* (1686): section 9.

4. Jaegwon Kim, *Supervenience and Mind*, (Cambridge: Cambridge University Press, 1993).



Ruminations on Intermittent Existence

say, the fact that psychological properties cannot differ without neurological properties differing is not a solution to the mind-body problem, it is the problem precisely. Why is it that this relation holds? That is the mind-body problem.

Likewise, when we say that a gun supervenes on its assembled parts, this does not amount to a metaphysically deep thesis. This, like the mind/body supervenience example above, is the phenomenon stated, not a solution. Since properties supervene on parts of an object when they are arranged a certain way, we need to closely examine the part-whole relationship.

Mereology

Mereology is the study of the relationship between parts and wholes.⁵ The notion of “part” is perhaps as familiar as any concept we use in ordinary language: the ham is part of the sandwich; the student is part of the class; this section is part of a paper on intermittent existence. The remainder of this paper will suppose the truth of some of the axioms laid out in classical extensional mereology (CEM) as presented in Leonard and Goodman’s *Calculus of Individuals*.⁶ Among the axioms pertinent to intermittent existence are the universality of composition, extensionality of parthood, and the like. For the sake of conciseness, I will not consider intermittent existence in terms of “alternative” mereologies that take their inspiration from Aristotle and Plato.⁷

Mereology takes the concept “part” as basic and unanalyzable. Parthood is a reflexive, antisymmetric, and transitive relation, as below:

5. For the most comprehensive look at the history and applications of mereology, see Peter Simons, *Parts: A Study in Ontology* (New York: Oxford University Press, 2000).

6. Henry S. Leonard and Nelson Goodman, “The Calculus of Individuals and Its Uses,” *The Journal of Symbolic Logic* 5.2 (1940): 45-55.

7. See Kathrin Koslicki, *The Structure of Objects*, (New York: Oxford University Press, 2008) and Verity Harte, *Plato on Parts and Wholes: The Metaphysics of Structure*, (New York: Oxford University Press, 2005).

Reflexivity P_{xx}

Antisymmetry $P_{xy} \wedge P_{yx} \longrightarrow x = y$

Transitivity $P_{xy} \wedge P_{yz} \longrightarrow P_{xz}$

Reflexivity states that anything whatsoever is a part of itself. This relationship is just like the identity relation that says that anything is identical with itself. (But is that the typical notion of parthood? More on that question below.) Antisymmetry states that no two things are parts of each other. For example, a banana is a part of a banana tree but the tree is not a part of the banana. As we can see, antisymmetry tells us that parthood is a one-way street. Lastly, suppose you have a Russian nesting doll. When the doll is opened you discover a slightly smaller doll much like the one just opened. This second doll is part of the first doll. Then, as with all Russian nesting dolls, when you open the second doll you will find yet another smaller doll. This third doll is a part of the second doll, and it is also a part of the first doll, and so on and so forth for subsequent smaller dolls inside the nesting dolls. This is the transitivity of parthood.

One of the main features of classical extensional mereology is that it is extensional. We will need to keep this principle in mind as we progress through this paper. The extensionality of parthood (EP) may be stated as such:

EP If x and y are composite objects with the same proper parts, then $x = y$.

Mereological universalism states that any plurality of objects is itself an aggregate or sum. A table, my left earlobe, and President Barak Obama form an aggregate. Mereological universalism entails that, no matter how scattered or gerrymandered the object in question is, as long as the parts exist then the aggregate exists. As such, the aggregate exists only when all the parts exist. If any of the parts cease to exist, then the aggregate is no more; mereological sums cannot survive the loss of a single part. Notice that the



Ruminations on Intermittent Existence

functional aspect of the aggregate is not needed, unlike the supervenience theory above.

Now let us return to the objection that the gun exists while disassembled. So the gun does exist while disassembled, says our interlocutor. To see this, note that the parts that compose the sum while disassembled are precisely the same parts that compose the gun after assembly. By extensionality of parthood, the sum of the parts and the gun are identical. Therefore, the same gun exists prior to assembly and after disassembly, and a gun while disassembled and once reassembled is not a case of intermittent existence.

But are the parts really identical with the gun? Here is an argument using Leibniz's Law (Indiscernibility of Identicals) that shows that they are not identical:

1. The sum, or aggregate, of the parts cannot lose any single part.
2. The gun can lose a single part,
3. Therefore, the gun is not a mereological sum

Propositions 1 and 2 describe different persistence conditions, thus, by Leibniz's Law, the gun and the aggregate are not identical. But did we not already concede the truth of EP? Yes, we have. So that means that the gun and the aggregate have the same parts, right? No. To see this we only need to invoke the standard mereological notions of proper and improper parthood:

Proper Parthood $\forall (x)\forall (y) (PPxy \longrightarrow (Pxy \wedge \neg Pyx))$

Which reads: if x is a proper part of y , then x is a part of y and y is not a part of x . This concept is sometimes put in terms of overlapping (part sharing). So if x is a proper part of y , then all parts of x overlap the parts of y but y has

at least one part that does not overlap with x . Proper parthood, then, is the typical, everyday conception of parthood that is only had by certain things.

Improper parthood is had by every object to itself. Any object, such as a spoon or a banana, is an improper part of itself. And so, the mereological sum of the parts of the gun is an improper part of itself. On the other hand, the mereological sum is a proper part of the gun. Therefore, the sum and the gun do not have the same proper parts. As mentioned above, the mereological sum has its parts essentially – it is not the same sum if even one microphysical particle is missing. Proper parthood allows us to keep the non-identity of a thing and its parts without having to countenance something extra or over and above the parts themselves.⁸

Now we can see where mereology can help with explaining the supervenient relationship of the A-properties and the B-properties of the gun. We may say that the B-properties (the firearm capacity of a gun, the uniform look of it, etc.) cannot be exemplified unless the A-properties (being a mereological sum of the smaller, perhaps microphysical, parts in this case) are in a subvenient relation, and to be in a subvenient relation is to be a proper part of something. As above, to be a proper part of something x is to be a part of an object y and for y not to be a part of x .

When the parts of the gun are not assembled, they merely form a mereological sum. The B-properties of a gun cannot supervene as of yet since the sum is not in an arrangement that would yield the parts as proper parts of the gun. Thus, “gunness” itself cannot supervene on the mereological sum of the parts, and, as such, when there is not a gun. So even mereology will not aid with the claim that the gun exists for a while as a disassembled gun. Let us return now to the case where the woman has disassembled her gun for transportation. Once she reaches her destination, she has her disassembled parts and she has her extensionality of parthood

8. This is sure to be controverted by traditional mereologists, but for more on how an artifact could have different proper parts than the mereological sum of its parts, see Achille Varzi, “The Extensionality of Parthood and Composition” *Philosophical Quarterly* 58.230 (2008) 108–133, especially pages 119–122.



Ruminations on Intermittent Existence

principle handy. Once reassembled, by EP, she has, once again, the same gun as before disassembly. Note that the reassembled gun has all the same proper parts as before the trip. And if we have the same proper parts, we have ourselves the same object.

Through mereology, we have found a case of intermittent existence. At the beginning of the trip, the woman takes her assembled gun, which is composed of parts $P_1 \dots P_n$, and disassembles it for transportation. During the trip, the disassembled parts do not compose a gun, only a mereological sum. So if the parts fail to compose a gun, a gun is not present while she takes her little trip. As above, once she reaches her destination, she takes the mereological sum and composes the same gun. By EP, she has her same gun back.

To be sure, this is not a conclusive argument. What I have said about the gun and mereology says nothing about mereological change. For instance, EP tells me that at the end of the story, the woman has her same gun that she had at the beginning of the story. But what about when she takes another trip, years later, to a gun shop because her gun needs a new part? The replacement of, say, the firing pin, may be a significant mereological change that would forever destroy the original gun. This would amount to the thesis of mereological essentialism. Artifacts, like sums, have their parts essentially and cannot survive change. This is not a popular thesis among mereologists by any means, though it does have a formidable endorser in Roderick Chisholm.⁹

Given the strength of mereological explanations, an argument for or against intermittent existence will have to address mereological change involving parthood at a time and even whether objects have temporal parts (4-Dimensionalism) or not (3-Dimensionalism). The mereological argument will also need to respond to paradoxes of composition such as the problem of the many. To be sure, the debate is complex. But at least we understand the debate would do well to carry on in mereological terms. ❖

9. See Roderick Chisholm, "Parts as Essential to their Wholes," *The Review of Metaphysics* 26.4 (1973): 581-603.

A Multi-Causal Approach To Synchronicity

Zachary Stinson

ABSTRACT: Synchronicity has long been described as an ‘acausal’ connecting principle. However, the use of this descriptor is not only misleading, but also outright false on any seriously considered picture of synchronicity due to admissions of multiple types of causes. Furthermore, previous attempts to clarify the ‘acausal’ label have served only to further muddy the waters of discussion. A ‘multi-causal’ conception of synchronicity is proposed to ease and encourage future discussion in many disciplines.

Synchronicity is a nefariously slippery topic. I contend that much of the confusion surrounding synchronicity stems from describing it as ‘acausal.’ The myriad of explanations and interpretations of this terminology muddy the waters of discussion. In my search to better understand this topic, I have arrived at the position that synchronicity should be described, instead, as ‘multi-causal.’ Gestures made to Aristotelian conceptions of causes favor the adoption of a ‘multi-causal’ description of synchronicity, and conflict with the current ‘acausal’ conception. More importantly, though, conceiving of synchronicity as ‘multi-causal’ opens discussion in many disciplines, whereas the term ‘acausal’ has immensely limited the conversation.



A Multi-Causal Approach to Synchronicity

Since I posit that synchronicity is better understood as ‘multi-causal,’ I shall attempt to explain synchronicity one cause at a time while addressing my concerns with the ‘acausal’ label and previous attempts to clarify it. I shall first explain the basics of synchronicity and the intimately related process of individuation, including the invocation of Aristotelian causes made by Jung and Mansfield. I shall then illustrate how these two thinkers, as well as Aziz, implicitly acknowledge material and formal causes. Finally, I shall address the substantive lack of efficient causal discussion, which I credit to the continued description of synchronicity as ‘acausal.’ When appropriate, I shall highlight pertinent philosophical points of interest.

Synchronicity, Individuation, ‘Acausal’, and Final Causes

Synchronicity is a term employed to describe a connection between two events.¹ These events are symbolically expressive of a need for compensation and integration of unconscious content into ego-consciousness.² One of the events, the objective event, is a public occurrence in the realm of normal sense perception. The other event, the subjective event, is only existent in the mind of the individual and only experienced by him, often as a dream. Taken together as a synchronistic experience, the paired events guide toward the ultimate goal of individuation.³ In the process of individuation, a person realizes unconscious facets of himself, and by integrating them into ego-consciousness he achieves a more complete and functional personality.⁴ The unconscious need for compensation, prior to its recognition and remediation, shall be referred to as the subjective state.

1. Aziz, C. G. *Jung's Psychology of Religion and Synchronicity* (Albany: State University of New York Press, 1990): 73; Carl Gustav Jung, “Synchronicity: An Acausal Connecting Principle,” trans. R. F. C. Hull in *The Structure and Dynamics of the Psyche*, vol. 8 of *Collected Works* (Princeton: Princeton University Press, 1978).

2. Aziz, *Jung's Psychology*, 97.

3. *Ibid.*

4. *Ibid.*

Although synchronicity can be described as the connection between the subjective and objective events, Jung introduces a further descriptor – ‘acausal.’⁵ In the decades since coining it, others, such as Aziz and Mansfield, continue using the term.⁶ However, the specificity of what this term signifies is not superficially evident. Superficially, it appears to denote an absence of causes.⁷ However, Jung remarks that synchronistic pairings of events can occur “...without there being any causal connection between them.”⁸ Likewise, when describing synchronistic pairings, Aziz states that the events are “not causally related.”⁹ He also remarks that there is “nothing to suggest any mutual dependency concerning the fact of their coming into being.”¹⁰

From these statements it would appear that in a synchronistic experience the events are not causally related and do not depend on each other (or do not mutually depend on something else) for their occurrence. This seems flawed from the outset, as the component events are each symbolically reflective of the individual’s subjective state and thus further the individuation process. Put differently, all synchronistic experiences (component events included) further the process of individuation and are thus related. Additionally, Jung himself explicitly endorses the operation of final causes in the occurrence of synchronistically paired events, as does Mansfield.¹¹ This seems to be a movement to include the long-neglected conception of various types of causes, extolled by Aristotle, into discussions of synchronicity.

5. Jung, “Synchronicity: An Acausal Connecting Principle.”

6. Aziz, *Jung’s Psychology*, 59; Victor Mansfield, *Synchronicity, Science, and Soulmaking: Understanding Jungian synchronicity through physics, Buddhism, and philosophy* (Peru: Open Court Pub. Co., 1995): 20.

7. The negating prefix ‘a-’ usually denotes some type of an absence or lacking, such as in the words asymptomatic, amorphous, or asexual.

8. Jung, “On the Nature of the Psyche,” in *The Structure and Dynamics of the Psyche*, 215.

9. Aziz, *Jung’s Psychology*, 64 & 73.

10. *Ibid.*, 73.

11. Jung, “Synchronicity,” 493; Mansfield, 20.



A Multi-Causal Approach to Synchronicity

Following Aristotle, there are four different types of causes: material, formal, efficient, and final.¹² The material cause of something is the substance that composes it, such as the Styrofoam used to make a cup. The formal cause is the form or organization of the material cause, which in the current example would be the shape of the cup. The efficient cause refers to the actions that brought about the particular arrangement of the material cause, such as the physical manufacturing of the cup or the process of molding Styrofoam. The final cause refers to the purpose or end for which something occurs, which in the current example would be to have a portable container that keeps a beverage at optimal temperature. Each type of cause provides a different component of a complete explanation of the cup.

Andrea Falcon notes that Aristotle maintains a rather stringent criterion for something to be qualified as an end, or telos.¹³ Aristotle contends that, “not every final stage has a claim to be called an end; only the best is an end.”¹⁴ Stating that something is an end or final cause, to some extent, involves attributing normative value to it. This is professed to be the case with synchronicity since its final cause is the process of individuation, which Jung says “all our highest and ultimate purposes seem to be striving towards.”¹⁵ The process of individuation is likewise understood by Mansfield to be the “highest good” in human life.¹⁶

This is a considerable ethical claim, a virtually universal declaration of the moral obligation of every individual. The issue becomes weightier when considering Aziz’s “synchronistic worldview,” wherein the individual benefits from individuation process, but also provides in return,

12. Aristotle, *Physics*, trans. Robin Waterfield (New York: Oxford University Press, 1999): 38-39; Aristotle, *Metaphysics*, trans. Hugh Lawson-Tancred (England: Penguin, 1998): 10 & 244.

13. Andrea Falcon, “Aristotle on Causality,” *Stanford Encyclopedia of Philosophy*, (28 April 2008), <http://plato.stanford.edu/entries/aristotle-causality/> (6 September 2010): §4 ¶4.

14. Aristotle, *Physics*, 38.

15. Jung, “The Relations Between the Ego and the Unconscious,” trans. R. F. C. Hull, in *Two Essays on Analytical Psychology*, vol. 7 of *Collected Works*, (Princeton: Princeton University Press, 1978): 215.

16. Mansfield, *Synchronicity, Science*, 17.

“something of genuine importance to the progressive unfoldment of the whole.”¹⁷ The process of dynamic integration plays out on both a microcosmic level and a macrocosmic level, and responsibility thus extends past the individual level. While the conception of final causes purported to be involved in synchronistic experiences does appear to accord well with Aristotle’s criteria and opens up robust ethical discussion, its use directly conflicts with a description of synchronicity as ‘acausal,’ both superficially and by exhibiting a type of causal relation between synchronistic events.

Material Causes: Psyche, Matter, and the *Unus Mundus*

An implicit admission of material causes is made when describing the substance of synchronistically paired events. The objective event is said to occur in the material world, or an arrangement of matter.¹⁸ The realm of the subjective event is often referred to as psyche.¹⁹ This is what is referred to by Aristotle as a material cause – the subject of change, the substratum or material that is “capable of receiving such and such a form.”²⁰ In Aristotelian terms, then, just as Styrofoam is the material cause of a cup, psyche and matter are the material cause of subjective and objective events, respectively.

The issue is accentuated when considering another aspect of discussions on synchronicity, that being the concept of the *unus mundus*, or the “unitary world.”²¹ Although the synchronistically paired events are described as occurring in the realms of both psyche and matter, these seemingly disparate substrates are actually conceptualized as being “two

17. Aziz, *Jung’s Psychology*, 214.

18. Aziz, *Jung’s Psychology* 56; Mansfield, *Synchronicity, Science*, 45.

19. *Ibid.*

20. Aristotle, *Physics*, 38; Aristotle, *On the Parts of Animals*, trans. W. Ogle. (London: Kegan, Paul, Trench & Co., 1882): 7.

21. Jung, “A Psychological View of Conscience,” *Civilization in Transition*, vol. 10 of *Collected Works*, trans. R. F. C. Hull (Princeton: Princeton University Press, 1978): 452.



A Multi-Causal Approach to Synchronicity

ways of perceiving the same thing.”²² Jung adopts the view of matter and psyche as ultimately identical, noting that we have to think of them as distinct “simply for the purpose of better understanding.”²³ A similar phenomenon has occurred in the realm of quantum physics where the bifurcation of particles and waves has been replaced with a conception of a wave-particle-duality.

Most of those involved in discussions of synchronicity adopt this type of dual-aspect view of the *unus mundus* as the one and only substance of synchronistic experiences. Wolfgang Pauli describes psyche and matter as “complementary aspects of the same reality.”²⁴ Both Aziz and Mansfield profess the “identity of” and “unity of” matter and psyche, respectively.²⁵ Viewed in this light, rather than being composed of different substances, the *unus mundus* is the substrate of both events, thus it is their mutual material cause.

At this point a robust metaphysical discussion could take place. One route would be to inquire if this is a monistic view. If so, is it a reductive or non-reductive monism? Is the nature of this underlying substance more akin to our conceptions of the material or the ideal? Does it share the properties of both, or is it something yet to be conceptualized? Indeed, a plethora of implications could stem from the assertion that the all of existence is, in fact, composed of one underlying substance, the *unus mundus*. Another question that could be raised is if this view of existence accords with other tenets of synchronicity theory. These questions aside, it can easily be seen that synchronistically paired events involve final and material causes, and that these causes are shared by both events.

22. Albert Einstein, *Albert Einstein, and His Human Side: New Glimpses from His Archives*, ed. H. Dukas and B. Hoffman (Princeton: Princeton University Press, 1981): 38.

23. Jung, “*Instinct and the Unconscious*,” *The Structure and Dynamics of the Psyche*, vol. 8 of *Collected Works*, trans. R. F. C. Hull (Princeton: Princeton University Press, 1978): 138.

24. Wolfgang Pauli, “The Influence of Archetypal Ideas on the Scientific Theories of Kepler,” in *C. G. Jung and W. Pauli: The Interpretation of the Nature of the Psyche* (New York: Pantheon Books Inc., 1995): 210.

25. Aziz, *Jung’s Psychology*, 56; Mansfield, *Synchronicity, Science*, 45.

Formal Causes: Archetypes and Images

Another instance of implicit admission to shared causes pertains to the archetypes, which I contend to be the formal cause of the subjective and objective events. In a synchronistic experience, the particular need for compensation is reflected by the symbolic meaning of both events and is experienced only indirectly through archetypal manifestations.²⁶ If the subjective state of an individual is a need for the integration of the nurturing and maternal aspects of his psyche, the synchronistically paired events would have to reflect this particular need. Both the dream and the public occurrence would then presumably manifest some image of the 'mother archetype' which could take the form of his mother, a maternal figure, the forest, the ocean, or a church, just to name but a few archetypal images associated with the maternal.²⁷ In this way, the archetypes are described as "structuring" or "informing" both matter and psyche, and are able to guide the individual to integrate the unconscious content they represent.²⁸

While there are a myriad of potential archetypal images for any given archetype, in a synchronistic experience the same archetype is manifested in both the subjective and objective events. In this sense, the formal cause is the same for both. It is only in this way that the events would be considered to be synchronistically paired, or meaningful. This is because if both events did not express the meaning of that particular archetype, then there would be no parallel or meaningful connection at all. In this sense, the formal cause must be the same for the subjective and objective events in order for the events to be considered synchronistically paired.

Mansfield contributes to additional confusion when he suggests that the archetypes are not a cause of the synchronistic pairing of events.²⁹

26. Aziz, *Jung's Psychology*, 55.

27. *Ibid.*, 20.

28. Mansfield, *Synchronicity, Science* 45, Richard Tarnas, *Cosmos and Psyche: Intimations of a New World View* (New York: Viking, 2006): 57 & 503.

29. Mansfield, *Synchronicity, Science*, 20.



A Multi-Causal Approach to Synchronicity

While he may intend this to mean that they are not an efficient cause of the events, in Aristotle's language, they are certainly the formal cause of synchronistic events. Just as the formal cause of the Styrofoam cup gives it its usefulness, it is through the archetypes that synchronistic events have their usefulness, in other words their ability to convey the meaning of the subjective state through their recognizable structure. It can then be understood that synchronistically paired events share the same formal cause and mutually depend on it for the expression of the need for compensation by the ego. Furthermore, these events mutually depend upon that which they express, the subjective state.

Efficient Causation

By now it should be evident that the term 'acausal' conflicts with elementary tenets of synchronicity theory. The inclusion of multiple types of causes and suggestions of mutual dependency should suffice as evidence enough of the superiority of a 'multi-causal' description. However, one further 'acausal' intricacy deserves elucidation. While the term 'acausal' superficially implies an absence of causes, and is also explained as signifying a lack of causal relation or mutual dependency between events, it is used to denote an absence of efficient causes as well.³⁰ Specifically, what is denied is direct efficient causal influence of one event on the occurrence of the other.³¹ Essentially, the dream does not cause the public occurrence. While I can offer no evidence to prove that this assertion is false, I am compelled to ask if it necessarily must be the case.

Assuming that psyche and matter are both the same substance, the monistic view of the *unus mundus*, one major metaphysical hurdle could possibly be overcome. It stands to reason that without the difficulty of the problem of interaction, one synchronistic event theoretically could act

30. Jung, "Synchronicity: An Acausal Connecting Principle," 445-6; Mansfield, 20, 19.

31. Aziz, *Jung's Psychology*, 73; Mansfield, *Synchronicity, Science*, 20.

upon the other through efficient causation.³² Stated differently, if there is no difference between what we conceive separately as thought and matter, then it may be logically possible for the seemingly immaterial to interact with something that appears to be material, as there is no true difference of substance. Also, it is equally possible that the component events of the experience could be parallel effects of a shared efficient cause. In this sense, one event does not bring about the occurrence of the other through efficient causation. Instead, both events occur due to the same, yet distinct, efficient cause, much like waves on opposite banks of a river caused by the outward rippling of the water after a stone has been tossed in. The wave on one bank is not identical with the wave on the other, though both are distant, yet parallel effects of a single, shared efficient cause.

Aside from metaphysical concerns, it appears that the focus on describing synchronicity as ‘acausal’ has served to obscure theorizing and investigation into possible indirect avenues for efficient causal influence. Due to the exclusion of efficient causation by the blanket term ‘acausal’ little, if any, attention has been given to the possibility of other avenues for efficient causation in discussions of synchronicity. Indeed, given the high value placed on efficient causal explanation in academic study, the ‘acausal’ aspect of synchronicity has distanced its study from that of modern empirical psychology. Admittedly, discussers of synchronicity (or at least those cited in this paper like Jung, Aziz, and Mansfield) refer frequently to the discipline of quantum physics, but not to neuroscience. No attempt is made to relate the theories presented to the ever-expanding body of knowledge of the neurological correlates of mental life.

It is not a far leap to assume that if empirical study revealed biological processes or brain states involved either the perception or interpretation of synchronistic events, there is the inevitable likelihood that these faculties would be compromised in certain individuals. Suppose an individual was rendered unable to remember dreams due to abnormalities in

32. Howard Robinson, “Dualism,” *Stanford Encyclopedia of Philosophy*, (2007). <http://plato.stanford.edu/entries/dualism/>. (1 February 2011).



A Multi-Causal Approach to Synchronicity

brain centers involved in memory storage or retrieval, or suffered damage to some part of the brain that left him unable to recognize the symbolic content of the synchronistically paired events? This could effectively cut him off from these guides to the process of individuation.

If neurological structures and processes are implicated in the recognition of archetypal or symbolic meaning or the conscious integration of unconscious content, then this may be the beginning of a form of the problem of evil for synchronicity.³³ Whether through genetics, environmental influences, or an interaction of the two, it is inevitable that some individuals would lack this hardware and functionality. This poses the question of why some would possess an unequal or insufficient capacity to partake in a process that is heralded as “our highest good.”³⁴ The detriment could extend past the individual, as the furthering of the process of individuation on the macrocosmic level may be stifled due to a lack of something of “genuine importance.”³⁵ However, if these deficits are somehow necessary for the individuation process, then this could spark discussion regarding a soul-making theodicy in relation to synchronicity and individuation.³⁶

Conclusion

I have shown that synchronicity is not an ‘acausal’ relationship. I have shown the additional confusion that can likely be traced back to the use of this term concerning the refutation of related, mutual, or shared causes. Furthermore, the focus on describing synchronicity as ‘acausal’ simply to denote an absence of direct efficient causal influence has served to obscure theorizing and investigation into possible indirect avenues for efficient causal influence.

Making sense of synchronicity is confounding, confusing work that often leaves one with more questions than answers. If the ethical and meta-

33. Michael Tooley, “The Problem of Evil,” *Stanford Encyclopedia of Philosophy*, (2009), <http://plato.stanford.edu/entries/evil/> (3 February 2011).

34. Mansfield, *Synchronicity, Science*, 17.

35. Aziz, *Jung’s Psychology*, 217.

36. Michael Tooley, “The Problem of Evil.”

physical claims made in discussions of synchronicity have any weight, this topic may deserve careful attention from individuals in the disciplines of empirical psychology, psychotherapy, philosophy, and religion. That being said, there is no reason to make the conversation any more confusing than it has to be.

A 'multi-causal' conception of synchronicity accounts for the explicit and implicit admissions of multiple types of causes and allows for a framework to discuss the possibility of causal relation or mutual dependence between synchronistic events. Furthermore, a 'multi-causal' description does not directly conflict with a view of synchronicity that denies efficient causal influence of one event upon the occurrence of the other. Most importantly, however, it provides a manageable method of explaining the convoluted subject matter it is employed to describe while opening up the doors for future investigation and discussion. ❖



A Critical Assessment of George Klosko's Version of the Principle of Fair Play

Leonardo Moauro

ABSTRACT: The nature of our obligation to obey the law has consistently been an important object of philosophical dispute. Fair play based theories of obligation purport to show that it is unfair for us to benefit from an organizational scheme (such as the state) without contributing our fair share to the provision of goods. George Klosko is a major proponent of this approach. I develop his particular version of the argument from fair play into a defensible theory of citizens' obligation to obey the laws of their state.

In this paper, I will critically assess George Klosko's argument for the grounding of political obligations in the principle of fair play. In particular, I will inspect the validity of his argument for extending this obligation from schemes that provide only what he calls "presumptively beneficial goods" to those that also offer "discretionarily beneficial goods."¹ I will delineate how he explains obligations from fair play in the case of presumptive goods, but I will assume the argument is sound. I will argue that Klosko's particular justification of political obligations is intuitively valid by addressing two possible objections, which I call the "multiplicity of schemes objection" and the "process of addition objection." As I will

1. George Klosko, "Presumptive Benefit, Fairness, and Political Obligation," *The Duty to Obey the Law: Selected Philosophical Readings*. ed. William A. Edmundson (Lanham, MD: Rowman & Littlefield, 1999): pp. 197-198.



A Critical Assessment of George Klosko

demonstrate, I believe the answer to the second objection obviates the first one, so that we may construct a theory of political obligation based on Klosko's position.

Klosko follows H.L.A. Hart and John Rawls in delineating a source of political obligation to a cooperative scheme grounded in the principle of fair play. In short, when we actively accept the benefits of the collaborative effort of a plurality of individuals, we may be obligated (to those individuals) to bear our fair share of the necessary work or costs. Klosko distinguishes between two types of goods that a scheme can offer: "excludable" goods can be provided to specific individuals and denied to others (such as electricity and water services); in contrast, "nonexcludable" goods cannot be kept from specific others.² Individuals cannot help but benefit from nonexcludable goods. For example, a state that concerns itself exclusively with national defense can be construed as a scheme that provides a single, nonexcludable good. While it is simple to apply the principle of fair play to cooperative schemes that provide excludable goods, as individuals must actively seek to become its participants in order to benefit from them, it is less clear when they supply nonexcludable goods. In the latter case, we must distinguish between one who actively seeks benefit and what Robert Nozick calls the "innocent bystander," who benefits from the goods offered despite her attempts to avoid them.³

This is where presumptive goods become relevant to the argument. According to Klosko, we can suppose by definition that everybody desires presumptive goods regardless of whatever else they want from a cooperative scheme. He argues that in the case of schemes providing presumptive goods the importance of the benefits provided overrides the prospective participant's normal right to choose whether she wishes to enter the scheme.⁴ For example, we can assume that all citizens of a state desire na-

2. Ibid., 197.

3. A. John Simmons. "The Principle of Fair Play," *Moral Principles and Political Obligations* (Princeton, NJ: Princeton UP, 1979): 120.

4. Klosko, "Presumptive Benefit, Fairness, and Political Obligation," 198.

tional defense whatever else lies in their interest. So even though we may not have accepted the benefit of national defense, the state (as a cooperative scheme) can assume that it is important enough that we all accept its provision. Klosko is thus able to avert a problem plaguing the idea of obligation to schemes that provide nonexcludable goods. We can now complete the definition of a cooperative scheme which, according to him, would bind everyone to participate from the principle of fair play. This scheme must provide nonexcludable benefits that are (i) worth their costs to the recipients, (ii) presumptively beneficial,⁵ and (iii) balanced justly with burdens and distributed fairly within the scheme.⁶ This is the explanation that I accept as sound for the sake of argument. The discussion becomes more complex when considering discretionary goods, because the argument of importance does not apply.

Klosko rightly notes that an obligation to schemes providing presumptive goods is not sufficient to describe political obligations, since governments offer a myriad of goods that do not fit this definition. Imagine a state that endeavored only to defend its territory from external invasion. As noted above, this would be a scheme that provides a nonexcludable presumptive good, but it is philosophically uninteresting since it does not in the least correspond to modern states. Modern states (for the most part) provide what Klosko calls discretionary goods, which may be desirable but “should not be viewed as essential to people’s well-being.”⁷ How can an obligation from fair play to participate in schemes that produce presumptive goods be extended to those that offer discretionary goods? In my example, Klosko has explained why we may be obligated to states that ensure border patrol, but not why we should pay taxes for public schools.

Klosko argues for such an extension. He seeks to extend the function of cooperative schemes that already produce presumptive goods to

5. *Ibid.*, 197.

6. *Ibid.*, 203.

7. *Ibid.*, 198.



A Critical Assessment of George Klosko

include discretionary goods. The idea is that there is a difference between demanding cooperation for a scheme that provides both presumptive and discretionary goods and requiring additional contributions of an individual to a scheme that already obliges him.⁸ Thus, individual A has a *prima facie* obligation to cooperate with a scheme providing presumptively beneficial goods that, afterwards, begins to offer discretionary goods as well. In my example, citizens who are already obliged to contribute to the costs of national defense may, in a second moment, also incur an obligation to pay taxes for public education. In this case, Klosko writes, it is up to A to show that, once the scheme has assumed the provision of discretionary goods in addition to the initial presumptive ones, it no longer satisfies, as a whole, the criteria required for fair play-based obligation. So scheme X that provides presumptive goods and discretionary goods a, b, and c must be shown by A herself to be (i) overly costly or (iii) unfair to A in order for her to justifiably deny her contribution.⁹

For the remainder of the paper, I will focus on two pivotal criticisms of this argument. The first one, what I have called the “multiplicity of schemes objection,” attacks the notion of benefit packages implicit in Klosko’s argument. Klosko seems to imply that when members of a scheme decide to take on additional forms of cooperation, they are incrementing the package of benefits that the schemes in which they participate provide, but perhaps this is not the correct way to think about it. Maybe the result of assuming new forms of cooperation is a division into multiple schemes with one providing the original presumptive good and the others offering new discretionary goods. This means individual A would be free to subscribe to the former scheme (e.g. defense of the borders) without being a member of the latter ones (e.g. public education). Klosko could rebut since the cooperative members that provide the presumptive and discretionary goods are identical, it would not make sense to say that there are two dif-

8. *Ibid.*, 205.

9. *Ibid.*, 206.

ferent schemes. But this is not implausible: a credit card company, for example, offers very distinct benefits to its more longstanding (or wealthy) members than it does to regular customers.

Thus, we have a problem. How can we tie the additional discretionary goods Klosko writes about to the original presumptive good in such a way that they can collectively be considered part of the same “package of benefits?” We could undertake to demonstrate that the provision of national defense depends necessarily on the concomitant offering of goods such as public education. However, the manner in which Klosko frames his argument seems to preclude such an approach. The presumptive goods are chronologically prior to the discretionary ones. A more tenable method becomes apparent when we consider the process by which members of a scheme should decide to cooperate in the provision of additional discretionary goods, which takes us to the second objection to the overall argument.

The second counterargument, which I called the “process of addition objection,” asks the question of how the members of a cooperative scheme are to decide which discretionary benefits the scheme should provide in addition to the presumptive ones. Consider state X that provides its citizens only with the presumptive goods of national defense and physical security. According to Klosko’s position, the functionaries of this state could be justified in adding any task to their office so long as the sum-total of the goods provided met the three conditions necessary to generate obligation, even if they may be individually detrimental. Suppose state X now decreed that all citizens must be subject to forced labor for three hours every day. The overall scheme (assuming it is still a single one) may very well meet Klosko’s three criteria, but some of the collaborative “goods” now provided are not beneficial at all. So the extension of obligation requires a legitimate process of addition of collaborative ventures at the least.



A Critical Assessment of George Klosko

In what manner, then, should members of a scheme elect to take on the provision of particular discretionary goods? It appears to me that the most tenable method would include a democratic process: the members of the scheme should vote upon which discretionary goods to collaboratively provide. Being an insider to a collaborative scheme gives individuals many stakes in the scheme itself; it would be unfair for any insiders not to have a voice regarding the direction in which their scheme is being lead. Thus, the principle of fairness would apply not only to the distribution of goods but also to the process whereby it is decided which discretionary goods a scheme is to provide. By my position, if we live in a state that provides only national defense, we can subsequently convene in the town hall and choose democratically what else we would like the state (as a scheme) to do for us. Notice that we now think of the goods as a single package of benefits: the “multiplicity of schemes” argument loses impetus. This process of addition would also ensure that no detrimental “goods” be added to the scheme as in the forced labor example, thus countering the second objection. In my view, then, a democratic process of addition is necessary for fair-play-based obligation.

There are counterarguments to my view in favor of incorporating democratic standards in collaborative schemes to save the principle of fair play, but I hope to dismiss them. In the first place, it might be said that the collaborative scheme becomes susceptible to the problematic situation of the tyranny of the majority. If a consistent and unchanging majority systematically votes against the minority, the objection follows that the minority is in a *de facto* state of subjection to power. However, I believe the intuitive force of the counterargument is misplaced in this context: added discretionary benefits must each be worth their costs to their recipients. The objection loses considerable force once this is remembered: a perennial minority would not be disrespected or exploited, because it would actually benefit from the ordinances of the majority, just not in the way it envisioned.

The second objection grants that incorporating a democratic ideal into the process of addition of discretionary benefits would constitute sufficient grounds for political obligation. However, it seems that we are now merely supplementing the principle of fair play with another source of obligation, as opposed to extending its influence. If we accept this counter-argument, we would have to admit that fair play fails to generate political obligation on its own, and must be buttressed by another theory. I believe we can surpass this conceptual hurdle by writing the fair process of addition into the very conditions a scheme must satisfy in order to generate political obligation. Thus, we could add a fourth (iv) criterion: there must be a fair method of decision regarding which discretionary benefits to add to the scheme. Notice that though this is a slight modification of the reasons why we are bound to schemes by a principle of fair play, it is not a source of obligation independent of fair play. We are not obliged to schemes because we have a right to vote, but because in certain conditions (the vote being one of them) conducive to the generation of nonexcludable benefits, it would be unfair of us not to do our part for the scheme providing those benefits.

In conclusion, we can ascertain that Klosko's extension of fair play obligation from schemes providing presumptive goods to those that additionally offer discretionary goods is intuitively justified. However, his argument encounters several difficulties that somewhat mutate its premises and conclusions. Klosko is mistaken, for example, in thinking that the three conditions required for the generation of fair play obligation apply only to schemes as a whole. I have shown that it is more plausible to think they are applicable also to the individual benefits themselves. Along with the addition of the democratic ideal condition, this is the only possible way to make the assumption that schemes provide a package of benefits rather than considering each good as produced by a distinct scheme. Finally, we can circumvent the criticism that the argument no longer resembles one



A Critical Assessment of George Klosko

of fair play if we write the voting practice for the addition of discretionary goods into the initial conditions required for the generation of political obligation. In sum, we may be politically obliged to a scheme that satisfies these four criteria: (i) it provides benefits that are worth their costs to their recipients, (ii) at least one of these benefits is presumptively beneficial, (iii) the burdens and benefits are distributed fairly within it, and (iv) the discretionary benefits provided are fairly decided upon. ❖

Gettiering Goldman

Kenneth Stalkfleet

ABSTRACT: This paper examines the causal theory of knowledge put forth by Alvin Goldman in his 1967 paper “A Causal Theory of Knowing.” Goldman contends that a justified, true belief is knowledge if and only if it is causally connected to the fact that makes it true. This paper provides examples, however, of justified, true beliefs with such causal connections that are clearly not knowledge. The paper further shows that attempts to salvage the causal theory are unsatisfactory.

I. Introduction

Identifying what instances of belief are knowledge has long been a problem of philosophy. From Plato until the 1960s, the traditional position was that knowledge was simply justified, true belief. Since Edmund Gettier showed in 1963 that justified, true belief is not sufficient for knowledge, various modifications to the traditional position have been suggested.¹ Chief among these suggestions was Alvin Goldman’s suggestion in 1967 that a justified, true belief is knowledge only if that belief is causally connected with the fact that makes it true.²

In this paper, I will show that Goldman’s causal condition is not, in combination with a justified, true belief, sufficient for knowledge. I will

1. Edmund L. Gettier, “Is Justified True Belief Knowledge?” *Analysis* 23.6 (1963): 121.

2. Alvin Goldman, “A Causal Theory of Knowing,” *The Journal of Philosophy* 64.12 (1967): 358.



Gettiering Goldman

do so first by giving an example in which a justified, true belief is causally connected only incidentally with the fact that makes it true. My second counter-example to Goldman will show that self-fulfilling prophecies present a problem for Goldman and the causal theory of knowledge. I close the paper with a suggested revision to the causal theory.

Edmund Gettier, in his 1963 paper "Is Justified True Belief Knowledge?," famously demonstrated that a true, justified belief is not necessarily knowledge.³ Gettier gave two cases of justified, true beliefs (JTBs) that are not knowledge.⁴ We may imagine that Smith has applied for a job. Smith is told by the president of the company that Jones will actually get the job. Smith also just saw Jones count the coins in his pocket, seeing him count to 10. Smith then forms the following justified belief: 'The man who will get the job has 10 coins in his pocket.' Smith, not Jones, actually gets the job. Coincidentally and unbeknownst to Smith, Smith also has 10 coins in his pocket, and so his belief 'The man who will get the job has 10 coins in his pocket' is justified and true. It clearly is not the case, however, that Smith knew 'The man who will get the job has 10 coins in his pocket' when he formed the belief.

Alvin Goldman, in his 1967 paper "A Causal Theory of Knowing," attempted to resolve the problem presented by the Gettier cases by amending the traditional definition of knowledge to include the requirement that a subject's justified true belief (JTB) of some fact (p) must be causally connected with that fact (p).⁵ Goldman offers the following interesting case where his causal theory triumphs over the traditional JTB analysis. Suppose there is some person, T, who intends to go downtown on Monday. T communicates the intention to S on Sunday, who forms the justified belief (p): 'T will go downtown on Monday.' T then decides not to go downtown, but is kidnapped and taken downtown, and so (p) is true. Under the tra-

3. Gettier, "Is Justified True Belief Knowledge?" 121.

4. Gettier, "Is Justified True Belief Knowledge?" 122.

5. Goldman, "A Causal Theory of Knowing," 358.

ditional, JTB account of knowledge, *S*'s belief of (*p*) is knowledge. According to Goldman's position, however, since there is no causal connection between *S*'s belief of (*p*) and the fact that (*p*) is true, *S*'s belief of (*p*) is not knowledge. Since, intuitively, *S*'s belief of (*p*) is not knowledge, Goldman's theory triumphs over the traditional theory.

I now hope to show that Goldman's JTB+C (that is: justified, true belief with causal connection) is not sufficient for knowledge. I claim that there are cases of justified true beliefs with causal connections to the facts making them true which are not proper instances of knowledge.

II. Complex Kidnappings

In this section, I will present an example of a case in which a justified, true, belief is causally connected to the fact that makes it true but is not knowledge.

Case 1: Jack decides, on Saturday, that he would like to have lunch downtown on Sunday, and sends an e-mail to Renee on Saturday telling her he will be downtown Sunday without giving her an explanation. Renee, knowing Jack to be honest and reliable, forms the justified belief that Jack will be downtown on Sunday. Unfortunately, the e-mail is intercepted by terrorists planning to kidnap the President, who is (unknownst to Jack) going to be downtown on Sunday. Since Jack is a highly trained government operative, the terrorists fear his interference and kidnap him Sunday morning, taking him to their base downtown.

It is convenient now to introduce notation for our causal diagrams. Let (*x*) represent a designated fact and *B*(*Y*,*x*) represent *Y*'s belief of *x*. As with Goldman's analysis, arrows represent causal connections. Unlike Goldman's diagrams, I will just use solid arrows and assume that solid arrows between beliefs indicate inferences.



Gettiering Goldman

Returning to our example, let (p)=Jack's being downtown Sunday, (j)=Jack's intending to go downtown Sunday, (k)=Jack's getting kidnapped, (t)=Jack's telling Renee he intends to go downtown on Sunday, R=Renee, and (u), (v) be auxiliary facts that provide justification for Renee's inference that Jack actually will go downtown, such as Jack's honest, reliability, etc. The causal diagram for this situation looks, roughly, like the following:

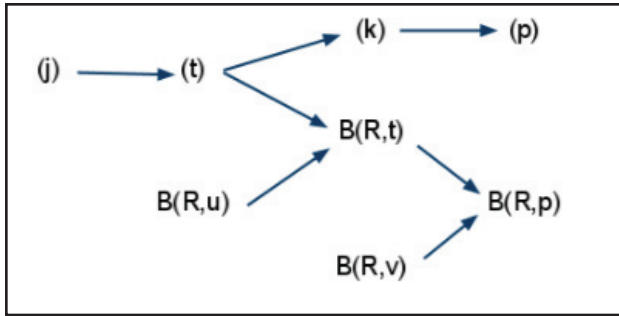


Figure 1. Causal diagram for case 1.

It is clear that there is a causal link between Jack's being downtown Sunday and Renee's believing Jack will be downtown Sunday, but intuitively, Renee does not know on Saturday that Jack will be downtown on Sunday.

The causal theorist must give an account of either where the causal chain is broken or how the theory can be adjusted to accommodate these cases. The causal theorist might just require that the believer have complete knowledge of the causal chain. This is obviously too strong a requirement on knowledge, however. For example, it might be the case that I discover a series of incorrect, hand-written multiplication tables in the library and correct them, thus allowing future readers to learn the multiplication tables. Future readers need not know that I corrected the tables in order for them to learn multiplication from them. Thus we see that for a belief to be justified the believer need not know the entire causal connection between the belief and the fact that makes it true.

The causal theorist might be uncomfortable with the role that intention plays in this example. He might hold that Renee does not just believe (p), she believes more strongly that Jack will go downtown of his own volition. Renee's belief that Jack will go downtown is contingent on her belief that his intention to go downtown of his own volition will be satisfied. If she had reason to believe that terrorists would kidnap Jack, she likely would not believe he would wind up downtown (unless of course she knew that's where the terrorists would take him). Thus the causal theorist might suggest that Renee's belief that Jack will go downtown voluntarily be included in the causal diagram. Since this belief is false, the causal chain fails and Renee does not have knowledge.

Since the best route of escape for the causal theorist is to show that what Renee really believes is false, I will present a pair of brief modifications of the original case to show that attempts to state Renee's belief in a way that works for the causal theory will either (i) be too complex, allowing us to construct a case where what Renee supposedly believes is false but in which she clearly has knowledge or (ii) be too simple, allowing us to construct a case where what Renee supposedly believes is true but insufficient for knowledge.

The causal theorist cannot get too stingy about what Renee thinks of Jack's intentions. To see this, consider the following:

Case 1a: Just like Case 1, except Jack escapes the terrorists and accomplishes his original goal of having lunch downtown.

In Case 1a, it would make sense to say that Renee's belief about Jack being downtown constituted knowledge (since the kidnapping was just an unexpected diversion that resulted in his being downtown), even though her belief that he would go downtown of his own volition was false (since,



Gettiering Goldman

although he remained downtown voluntarily, he went downtown involuntarily). Moreover, any of Renee's beliefs about how Jack was going to get downtown would be false. Once we take this into account, the causal theorist is left saying that Renee believes something like 'Jack will at some point on Sunday be downtown of his own volition'. This belief is not specific enough, however.

Case 1b: Just like Case 1 except Jack escapes the terrorists and instead of having lunch downtown he runs to warn the President about the terrorists.

In Case 1b, Renee's JTB 'Jack will at some point on Sunday be downtown of his own volition' would not constitute knowledge, even though it is still causally connected to the fact that makes it true (the e-mail that caused her belief caused Jack to be kidnapped which in turn caused him to be downtown of his own volition to warn the President of the attack). This forces the causal theorist to amend Renee's belief yet again to say something like 'Jack will at some point on Sunday be downtown in an effort to satisfy some expectation that prompted his e-mail.' It is clear that this is far more complicated than the belief that Renee intuitively has, namely just that Jack will be downtown on Sunday.

These complications suggest that complex beliefs should be avoided in the causal diagram in favor of their simpler constituents (that is, Renee actually has a collection of beliefs: {Jack will be downtown, Jack wants to go downtown, Jack will voluntarily get downtown...} some of which are true and some of which are false), otherwise one winds up with lengthy, convoluted beliefs that are intuitively very different from anything Renee actually believes. Notice that Goldman requires that enough of Renee's beliefs to be true to "ensure the existence of at least one causal connection."⁶ This being the case, if Renee's beliefs about Jack's going downtown are

6. Goldman, "A Causal Theory of Knowing," 368.

broken into their simplest possible constituents, it is clear that there is some causal connection of true beliefs that lead to Renee's belief that Jack will be downtown on Sunday, even if there are many false beliefs (about Jack's intention, method of getting downtown, etc.) that also support that conclusion.

The causal theorist might yet be convinced that Renee's beliefs about Jack's intentions are critical to her conclusion that he will be downtown. I will now present a case of JTB+C which requires no discussion of intention and yet presents a problem for the causal theory; this is the case of the self-fulfilling prophecy.

III. Self-fulfilling Prophecies

Case 2: Gene, who has an impeccable track record and is highly skilled at stock analysis, and who knows that company A is going to announce a new product this month, comes to the conclusion that the price of stock A will rise by \$100 over the next month. Gene tells some friends of this expectation and his prediction quickly becomes the talk of Wall Street. Knowing Gene's track record, traders everywhere start buying stock A at higher and higher prices until by the end of the month the stock has risen by over \$100. However, because the media is saturated with talk of Gene's report, when company A announces the new product, no one pays any attention. That is, the only factor that causes anyone to buy stock A that month is Gene's prediction. Did Gene know, at the start of the month, that the stock would rise by over \$100?

Let the following be our notation, G=Gene, P1 through Pn=people who invest in stock A, (a)=The fact that company A is going to make a product announcement this month, (m)=The fact that stock A goes up \$100 over the next month, (u)=Auxiliary beliefs about stock A and company A, (pi)=Person Pi buys stock A (where i goes from 1 to n for some large n), and A(Y,x)=the assertion of belief x by believer Y. Then the causal diagram for this case looks something like the following:



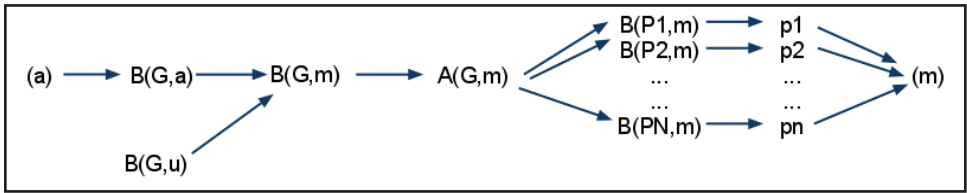


Figure 2. Causal diagram for Case 2.

Notice that the only causal role of (a) is to cause Gene’s belief of (a); (a) does not directly cause any other individuals to believe stock A is going to rise by \$100, those beliefs are all a result of Gene’s assertion that the stock will rise by \$100. Figure 2 illuminates the fact that self-fulfilling prophecies have exotic causal diagrams where a belief in some fact causes that fact to be true.

The causal theorist is now faced with a dilemma: if Gene has knowledge in this case, then it seems that all justified self-fulfilling prophecies should constitute knowledge, which is counter-intuitive. If for instance, the US government announces that they’re investigating company A, which would send the stock down were it not for Gene’s media-saturating prophecy, we would say that Gene did not really know that the stock would reach \$100. So long as Gene is justified in his beliefs about the strength of the company, however, the causal diagram would not show any difference between this case and Case 2.

On the other hand, if Gene’s belief of (m) is not knowledge, then the causal theorist has to explain where the causal chain faults, which he seemingly cannot do. The temptation is to claim that Gene has the false belief that the stock will go up as a result of how good the company is. This is not a false belief, however, as the strength of the company causes the prophecy which, in turn, causes the stock to rise. The causal theorist may simply disallow causal chains in which a belief is the cause of some fact which makes that belief true. This prohibition would be problematic first for cases where

a belief might be one of many causes of some fact that makes the belief true, and moreover for cases such as Case 2, where a completely sound prophecy is incidentally self-fulfilling.

Self-fulfilling prophecies thus present a clear problem for the causal theorist. If a belief of p is the cause of p , then clearly the belief and the fact are causally connected, but it seems a stretch to say that all self-fulfilling prophecies constitute knowledge, particularly those in which the prophecy certainly would not have come true were it not made. The consideration of what might have happened had the prophecy not been made motivates a possible way for the causal theorist to deal with the problem of self-fulfilling prophecies. A causal theorist might claim that a self-fulfilling prophecy constitutes knowledge if it is the case that, were the prophecy never made, the belief would have been justified, true, and causally linked to the fact that makes it true. Under this explanation, Case 2 seems to be a legitimate case of knowledge, assuming that without Gene's prediction saturating the media, the product announcement would cause the price of stock A to rise by \$100.

IV: Counterfactuals

Assume we accept the inclusion of counterfactuals into the causal theory, postulating something like: "A justified, true belief is knowledge if and only if it is causally connected to the fact that makes it true and only if a causal connection would be present had the belief never been asserted or acted upon."⁷ This definition of knowledge better handles cases of self-fulfilling prophecies, but is more difficult to use than the original causal theory. For example, to determine whether Gene had knowledge in Case 2, we must imagine a world in which Gene never makes the prophecy. Doing

7. At first glance, one might worry that if, for example, I do not act upon my belief that I will have a sandwich for lunch that I will never have said sandwich and thus cannot have knowledge about my future behavior. I do not take my belief that I will have a sandwich for lunch to be a cause of my having a sandwich for lunch, however, I rather take the belief and the fact to both be effects of my desire and ability to have a sandwich for lunch.



Gettiering Goldman

this, and examining the causal chain within the imagined world, is much more complicated and significantly less reliable than the empirical investigation required determining instances of knowledge under the original causal theory.

In order for the causal theory to be sufficiently repaired to handle self-fulfilling prophecies, its expedience must be sacrificed. Whereas the classic causal theory allowed for knowledge to be determined on an empirical basis, the amended causal theory requires an investigation of counterfactual situations, a much less reliable method for identifying knowledge. Ultimately, cases of the sort I have presented reveal serious defects in the causal theory that cannot be repaired without complicating the theory in suspect and unsatisfactory ways. ❖

A Rawlsian Revitalization of Gewirth's Normative Structure for Action

Bo Fox Pons

ABSTRACT: Alan Gewirth's *Reason and Morality* justifies certain fundamental moral principles and develops morality out of the basic structure of action. Contemporary literature exposes a critical flaw in the second stage of Gewirth's argument contending that Gewirth fails to create agent-neutral moral claims. In order to provide a transfer of interests between agents, the solution to Gewirth's problem, I argue that certain Rawlsian concepts buttress and are consistent with Gewirth's argument for the normative structure of action.

When looking to the history of moral philosophy, one can trace a search for moral rules that one must unconditionally accept. The two main branches of ethical theory highlight this search. The Categorical Imperative develops a moral framework from the principle of the good will. Utilitarianism maintains the absolute rule of maximizing happiness. Conflicting moral intuitions prevent resolution in judgment between these two theories. We are then led down the eventual path of ethical nihilism, as to abandon absolute norms would be the catalyst for losing morality to relativism.



A Rawlsian Revitalization of Gewirth's N.S.A.

The slippery slope emerges as relativism evolves into ethical subjectivism; once morality is reduced to opinion, the skeptic can question the purpose of morality. The nihilist, who would abandon morality's existence entirely, then follows.

In an attempt to defend the existence of moral rules, we should look to Alan Gewirth's *Reason and Morality*.¹ Gewirth's argument that action has an essential normative structure provides an argument for morality which surpasses other ethical doctrines. Even if one's moral intuitions disagree with Gewirth's moral theory, one is found, according to Gewirth's argument, in a logical contradiction. To defend Gewirth's argument and the ontological existence of morality, I shall address the criticism of Gewirth recently raised by Vaughn Huckfeldt. This project shall engage Gewirth's position, evaluate Huckfeldt's critique, and then show how Rawls' thought experiments can be used to defend Gewirth's supreme principle of morality.

Gewirth discusses the purposive nature of actions to introduce the normative structure of action. Gewirth highlights how agents act towards some goal or end. The purpose of acting or achieving that goal constitutes a reason for action. For example, if I desire a sip of soda pop, I connect that purpose with a specific action, such as moving the can to my mouth for a drink. Gewirth then connects an action's purpose to values. Gewirth writes, "[T]he agent necessarily regards his purposes as good and hence makes an implicit value judgment about them."² When looking at any purpose, even one as simple as taking a drink, the reason one acts is because he/she views his/her goals, such as quenching thirst, as good in some respect.

Gewirth then formulates his supreme principle of morality with three main steps. The first step evaluates the goodness of one's purposes and the necessary goodness of what helps one act on those purposes: free-

1. Alan Gewirth, *Reason and Morality* (Chicago: The University of Chicago Press, 1978).

2. *Ibid.*, 41.

dom and well-being. The second step transforms evaluative judgments about the necessary goods into deontic judgments forcing one to make right claims to freedom and well-being. The third step shows both how every agent must claim certain rights and how every agent must logically accept that “all prospective purposive agents have rights to freedom and well-being.”³ These steps will be the blueprint for providing further details to Gewirth’s argument for absolute rights.

The first concept of Gewirth’s to unpack is his dialectically necessary method (D.N.M.). Gewirth makes a distinction between two different types of methods, the dialectically contingent method and the D.N.M. The dialectically contingent method is described by Gewirth as, “[Beginning] from singular or general statements or judgments that reflect the variable beliefs, interests, or ideals of some person or group.”⁴ This method would allow one to determine what is contingently valuable to an individual. For example, the dialectically contingent method would say that a can of soda pop is necessary for me acting on the purpose of taking a drink of soda pop. Gewirth then defines the D.N.M. as, “[Beginning] from statements of judgments that are necessarily attributable to every agent because they derive from the generic features that constitute the necessary structure of action.”⁵ When one engages the D.N.M., he/she examines what it takes to act in general.

As one reflects on using the D.N.M., the agent would discover certain necessary goods needed for any agent to act on any purpose. Gewirth argues that because we necessarily see our purposes as good, we must then see how freedom and well-being, both required for acting, are necessary goods.⁶ Freedom is the procedural aspect of action as it allows agents to act on the purposes he/she views as good.⁷ When looking at the possibility of taking a drink, I would need to have the freedom to be able to move

3. *Ibid.*, 48.

4. *Ibid.*, 43.

5. *Ibid.*, 43-44.

6. *Ibid.*, 61.

7. *Ibid.*, 53.



A Rawlsian Revitalization of Gewirth's N.S.A.

around, as opposed to being tied in a chair, to get that beverage. Well-being is the substantial aspect of action which provides one with the ability to perform actions to achieve his/her goals.⁸ When I attempt to take a sip of soda pop, I need to be able-bodied to such a degree that I can physically achieve my goal.

After establishing these two necessary goods through the D.N.M., Gewirth enters the second stage of his argument: turning the generic goods of freedom and well-being into generic rights for an agent. Gewirth opens his argument with the following question: "If he regards these conditions as indeed necessary for the very possibility of his agency and for his chances of succeeding in his actions, then must he not hold that all other persons ought at least to refrain from interfering with the conditions?"⁹ Not only would rights be important for acting on one's purposes, but also rights would be in place to defend the value of the necessary goods. As freedom and well-being are necessary goods, Gewirth contends agents must believe that others should not interfere with the agent's own freedom and well-being. In other words, the agent is logically committed in claiming a right to freedom and well-being.¹⁰

Gewirth then gives his most explicit argument for why each agent must claim rights to freedom and well-being or else be caught in contradiction:

Suppose some agent were to deny or refuse to accept the judgment (1) 'I have rights to freedom and well-being.' Because of the equivalence between generic rights and strict 'ought,' this denial of (1) would entail the agent's denial of (2) 'All other persons ought at least to refrain from interfering with my freedom and

8. Ibid. A good article which provides further details about the hierarchy of well-being is Gewirth's "Ethical Universalism and Particularism," *The Journal of Philosophy* 85.6 (1988): 283-302.

9. Ibid., 63-64.

10. Ibid., 64.

well-being.’ By denying (2), the agent would have to accept (3) ‘It is not the case that all other persons ought at least refrain from interfering with my freedom and well-being.’ But how can any agent accept (3) and also accept (4) ‘My freedom and well-being are necessary goods’? That he must accept (4) we saw above; for by virtue of regarding his purposes as good the agent must also a fortiori value his freedom and well-being as required for achieving any of his purposes.... He must therefore accept, on pain of contradiction, that he has generic rights.

At this point in his theory, Gewirth has made a formal argument that because one must value freedom and well-being, one must accept he/she has certain generic rights with which others ought not to interfere.¹¹

In step three, Gewirth logically forces an agent into accepting similar right claims for all agents. Gewirth argues, “Now whatever the description under which or the sufficient reason for which it is claimed that a person has some right, the claimant must admit, on pain of contradiction, that this right also belongs to any other person to whom that description or sufficient reason applies.”¹² Gewirth uses the basic principle of universality of reasons to extend right claims to all agents. Gewirth now thinks he has justified that all agents must acknowledge that everyone has rights to freedom and well-being, the normative basis of morality.

Huckfeldt’s critique of Gewirth can now be examined. Huckfeldt first identifies the area with Gewirth’s argument where he discovers a problem. Huckfeldt claims that the success of the argument depends on “whether or not the necessary reason to pursue freedom and well-being require me to make agent-neutral claims (i.e. rights claims).”¹³ This assessment takes one back to the formal argument which forced one into logically

11. *Ibid.*, 80.

12. *Ibid.*, 104-105.

13. Vaughn E. Huckfeldt, “Categorical and Agent-neutral Reasons in Kantian Justifications of Morality,” *Philosophia* 35 (2007): 34.



A Rawlsian Revitalization of Gewirth's N.S.A.

accepting both the necessary goodness of freedom and well-being and the claim that other agents ought to refrain from interfering with one's necessary goods. If one could escape a logical contradiction and accept both claims (3) and (4), then the right claims preventing others from interfering with one's freedom and well-being cannot be made.

Huckfeldt goes on to explain where he sees the possibility of non-contradiction between claims (3) and (4). Huckfeldt argues, "[W]e notice that although it is required that I pursue my own possession of uninterfered with freedom and well being, and that others pursue their own, it is permissible, even according to my own judgment, for others to interfere with my freedom and well being. Although both myself and the other are in pursuit of *f*,¹⁴ neither of us is required to pursue the possession of *f* for anyone but ourselves."¹⁵ Huckfeldt's rejection of Gewirth is founded on the argument that one has no reason to think others should have any positive consideration for another's necessary goods. Though Gewirth does attempt to universalize this principle during the third stage of his argument, the argument cannot follow when the *reductio* Gewirth provides during the second stage does not establish a reason for consideration of the non-interference of others' necessary goods.

Fortunately, Huckfeldt presents both a problem in Gewirth's argument and what is needed to fix Gewirth's justification. Huckfeldt acknowledges, "For his argument to work, Gewirth would need a principle entailing a transfer of interests between people."¹⁶ To rebuild Gewirth's justification for morality, one would need to establish a principle within his argumentative structure that transfers interests between agents without simply assuming said moral principle. The remainder of this paper will attempt to describe a possible solution to Huckfeldt's critique by connecting Gewirth's D.N.M. with certain Rawlsian concepts. The main argument

14. The letter 'f' is standing in for 'uninterfered with freedom and well-being' within Huckfeldt's argument.

15. Huckfeldt: 35.

16. Ibid.

I will defend is that certain aspects of Rawls' philosophy can be adopted by an agent engaging the D.N.M. When these two philosophical principles work together, a new D.N.M. will justify one's interest in the non-interference of other agents' freedom and well-being.

When looking for compatibility between these two philosophies, we see that Rawls' thought experiments are quite similar to the way Gewirth frames the D.N.M. Gewirth frames the D.N.M. so that any agent acting on a purpose must acknowledge certain necessary goods for generic action. Rawls' original position adopts a similar outlook. Rawls believes the original position is one that the same results are produced when it is adopted by anyone at anytime.¹⁷ It follows that both the identification of certain necessary goods reached through the D.N.M. and the principles chosen in the original position will always be the same for any agent entering either reflective state.¹⁸

Another similarity is found when Gewirth adopts the D.N.M. rather than the dialectically contingent method. The dialectically contingent method looks for goods in relation to contingent and variable purposes. This is parallel to the specific inequalities Rawls wants to avoid. Rawls contends that the principles behind the veil of ignorance would be those chosen by interest-advancing rational agents "when none are known to be advantaged or disadvantaged by social and natural contingencies."¹⁹ Both Gewirth and Rawls desire to avoid contingent matters of morality and justice, respectively, and focus on near objective and necessary principles.

The advantages of using the veil of ignorance and the original position with Gewirth's argument, namely that adopting the D.N.M. under Rawls' veil of ignorance, will force a contradiction in Gewirth's *reductio*, can now be explored as we have seen how the two theories are similar and compatible. The first argument stems from the result of social cooperation.

17. John Rawls, *A Theory of Justice*. (Cambridge: Harvard University Press, 1971): 139.

18. This is not to argue that the principles of Rawls are equivalent to freedom and well-being, what Gewirth identifies as necessary goods.

19. *Ibid.*, 19.



A Rawlsian Revitalization of Gewirth's N.S.A.

Rawls contends that "There is an identity of interests since social cooperation makes possible a better life for all than any would have if each were to try to live solely by his own efforts."²⁰ When one is logically committed to viewing freedom and well-being as necessary goods, one is also logically committed to pursuing and maintaining them. To secure one's necessary goods, it is seen that one can better pursue his/her freedom and well-being in conjunction with others rather than on one's own. One is then logically committed to promoting each other's right claims to freedom and well-being, under the veil of ignorance, as one is always logically committed to the pursuit and maintenance of one's own freedom and well-being. This new conclusion from the D.N.M. would be the direct result of our Rawlsian addition. Furthermore, Rawls writes, "There is no inconsistency, then, in supposing that once the veil of ignorance is removed, the parties find that they have ties of sentiment and affection, and want to advance the interests of others and to see their ends attained."²¹ The advantages Rawls is able to produce, namely social cooperation, are remedies to the exact problem Huckfeldt addresses. The ties of sentiment and advancement of others' interests, as Rawls describes, is the principle allowing a transfer of interest needed to force a contradiction between the claims (3) and (4).

The second argument relates to how the agent would not know which agent he/she will be outside the veil of ignorance. If an agent reflects on one's necessary goods under the D.N.M., the agent recognizes that freedom and well-being are essential to any action, but the agent's ability to pursue and obtain those necessary goods in the real world is unclear. When the agent does a similar reflection, this time with a veil of ignorance, the agent would not be able to determine the practical ability of pursuing freedom and well-being at all as knowledge of contingent features are removed. Therefore, that agent, before he/she emerges out of the

20. *Ibid.*, 126.

21. *Ibid.*, 129. While Rawls contends that the ties of sentiment emerge outside the veil of ignorance, the ties of sentiment still need the veil of ignorance to provide a justification in taking impartial action for the promotion of others.

veil of ignorance after using the D.N.M., would contend that an interest in other's non-interference of freedom and well-being would be a necessary moral principle guaranteeing his/her own pursuit of necessary goods as knowledge of which identifiable agent one will be is unknowable.

This argument is the key to how an appropriation of Rawls can force one to be logically committed to the positive consideration of others' freedom and well-being. It is so because if one did not add the moral principle of interest in others' non-interference with ones own necessary goods under the veil of ignorance, then one would not be consistent with his/her logical commitment to necessary goods. Because the veil of ignorance removes contingent features of an agent, yet lets the agent keep his/her essential agency, the agent reflecting on the D.N.M. would not know which agent he/she will be when the veil of ignorance is removed. Thus, to secure and allow the pursuit of necessary goods outside of the veil of ignorance, that agent, and every agent through the nature of the original position, must be committed to the interests of others in order to guarantee the promotion of his/her own necessary goods. As a rational agent, the only course to guarantee the maintenance of one's necessary goods is to have a well-ordered society, with extensive bonds between citizens,²² where everyone refrains from interfering with the freedom and well-being of all other agents—a society which agents respect the necessary goods of other agents.²³

Having explained how the veil of ignorance enhanced D.N.M. avoids Huckfeldt's critique, it needs be shown how Gewirth's critical remarks towards Rawls do not affect our defense of morality. Gewirth criticizes Rawls' for his contention that principles of justice need to be determined with an abandonment of particular qualities. "Since the assumption

22. *Ibid.*, 500.

23. This argument is buttressed by the fact that Rawls feels a rational agent would have characteristics of risk-aversion. *Ibid.*, 144. Even if complete risk-aversion is irrational, it should not be seen irrational to be risk averse when considering necessary goods. If one were to be risk averse towards anything, then it would be that which is fundamentally necessary for agency, namely freedom and well-being.



A Rawlsian Revitalization of Gewirth's N.S.A.

of persons' total ignorance of their particular qualities, being factually false, is hardly rational, how can it be rationally justifiable to rational persons in the real world to which Rawls intends that his principles of justice be applied?"²⁴ Gewirth contends that in the real world there are factual claims about the inequality of power and ability between people.²⁵ Gewirth can justify the abandonment of contingencies to the rational agent because he is searching for generic goods, which apply to purposes beyond particulars, a generalization relevant to all agents. Rawls, on the other hand, abandons particulars in relation to how one should act outside the veil of ignorance, something specific and not general to all agents. While Rawls acknowledges his argument's hypothetical nature,²⁶ he also shows the value which can be derived from such a thought experiment. "The conception of the original position is not intended to explain human conduct except insofar as it tries to account for our moral judgments and helps to explain our having a sense of justice."²⁷ With this perspective, Rawls is not intending to justify the original position to rational agents, but rather follow why we make normative judgments.

It still must be held that a rational agent is justified in believing a transfer of interests, either as a matter of social cooperation or risk aversion, can be a principle of morality even when the concept of losing all knowledge about particular abilities and power is not rational itself. After all, Thomas Hobbes provides a justified worry for even the most able and powerful of agents. Hobbes points out the possibility of people living without security, being forced to live on their own strength, of everyone being in conflict with everyone else, and living lives which are "solitary, poore, nasty, brutish, and short."²⁸ Every agent, regardless of his/her ability or

24. Gewirth, *Reason and Morality*, 108.

25. *Ibid.*

26. Also see John Rawls, "Justice as Fairness: Political not Metaphysical," *Philosophy and Public Affairs* 14.3 (1985): 223-251, for how Rawls claims his philosophy develops a political framework, not moral metaphysics.

27. John Rawls, *A Theory of Justice* (Cambridge: Harvard University Press, 1971): 120.

28. Thomas Hobbes, *Leviathan* (Cambridge: Cambridge University Press, 1651): 84.

power, has a justified reason in wanting to further secure his/her necessary goods of freedom and well-being as it is possible that even the strongest agent can be outclassed.

This paper evaluated Gewirth's justification of morality, from *Reason and Morality*, in an attempt to defend morality as an absolute and universal concept. After providing a thorough explanation of Gewirth's argument, Huckfeldt's critique of Gewirth was entertained. Huckfeldt's article exposes how the critical turn in Gewirth's argument could prevent an agent from making agent-neutral right claims. Huckfeldt's remedy of providing a transfer of interests between agents was the challenge the rest of the paper took up in the form of using Rawls' thought experiments of the original position and the veil of ignorance. Though Gewirth is critical of Rawls' principles, it was shown how Gewirth's philosophy can accept this paper's appropriation of Rawls' philosophy to justify morality as a normative feature of action by identifying aspects of social cooperation and risk aversion. Gewirth and Rawls are not only consistent, but also, when properly combined, provide an excellent justification of morality. ❖



A New Approach to the Paradox of Fiction

Pete Faulconbridge

ABSTRACT: It seems that an intuitive characterization of our emotional engagement with fiction contains a paradox, which has been labelled the 'Paradox of Fiction'. Using insights into the nature of mental content gained from the disjunctive theory of perception I propose a novel solution to the Paradox, explained and motivated by reference to Kendall Walton's influential account of fictionality. Using this insight I suggest that we can take the phenomenology of fictional engagement seriously in a way not allowed by Walton.

The 'Paradox of Fiction'

The following paradox underlies recent analytic philosophy on the topic of fictionality:

- (a) Readers or audiences often experience emotions such as fear, pity, desire, and admiration toward objects they know to be fictional, e.g. fictional characters
- (b) A necessary condition for experiencing emotions such as fear, pity, desire etc. is that those experiencing them believe the objects of their emotions to exist.
- (c) Readers or audiences who know that the objects are fictional do not believe that these objects exist.¹

1. Peter Lamarque and S. H. Olsen, ed., *Aesthetics and the Philosophy of Art* (Oxford: Blackwell, 2007): 298.



A New Approach to the Paradox of Fiction

These three statements have a certain intuitive appeal; each seems like a plausible part of the characterization of fictional experience. However, the three statements are clearly mutually incompatible. Any attempt at a theoretical solution to the paradox has thus tended to take the form of choosing which statement to reject, and giving an account of why it is dispensable. Each statement in turn has been rejected by at least one recent theorist, and a variety of compensatory moves have been suggested. My approach will be introduced and motivated by a discussion of Kendall Walton's solution, from which I will then move, appropriating an insight from the 'disjunctive theory of perception,' in order to propose a novel solution to the paradox². In brief, my suggestion is that we take as seriously as possible the phenomenology of fictional experience as commonly reported, in particular the fact that fictional emotions tend to feel both real and relational. I will show in what follows that we can do justice to this aspect of fictional experience without implying the existence of fictional objects, by appealing to the idea that we are not always authoritative about 'how things are for us.'

In brief, Walton retains statements (b) and (c) and rejects (a). Walton claims that, rather than it actually being the case that we are afraid, or envious, or pitying of fictional characters and objects, it is 'make-believe' that we are afraid, envious, pitying etc.

For Walton, when his imaginary film-goer Charles is watching a film about a ball of slime terrorizing people and says 'I am afraid of the slime,' this should not be taken as evidence that he is indeed afraid of the slime. Walton claims he cannot be afraid of the slime because the slime does not exist and Charles knows this. Instead we should interpret Charles' utterance as part of the game of 'make-believe' which constitutes Charles' artistic or imaginative engagement with the film. Charles does have "certain phenomenological experiences" which normally arise "as a result of

2. In discussing Walton's work I will draw primarily on his *Mimesis as Make-believe: On the Foundations of the Representational Arts* (Harvard University Press, 1990). cf. also 'Fearing Fictions', *Aesthetics and the Philosophy of Art*: 307-320.

knowing or believing that one is endangered.”³ These experiences, termed by Walton as ‘quasi-fear,’ constitute fear within the make-believe, meaning Charles does not feel genuine fear of the slime.⁴ This analysis is not in line with how Charles takes his experience to be, and so Walton’s solution should be considered as revisionist with respect to the subject’s self-report of his or her experience of fiction.

An Alternate Approach to the Paradox

I propose that we attempt to do greater justice to the experience of the ordinary film-goer or book-reader. This includes the fact that most take themselves to experience genuine emotions for fictional objects and characters. More specifically, I aim to do justice to the intuition that what makes Charles’ exclamations of fear correct is something other than their being true in a relevant game of make-believe.

Let us examine (a). Why do some theorists deny this statement? Walton frames this claim in terms of belief in the object of one’s emotions, but it seems that more fundamental than positing that this is the actual existence of the objects themselves. That is to say that fear, pity, envy, admiration, all of the commonly used examples in this field, are relational psychological states. One would hardly deny that one could become sad or melancholy after reading a book, or indeed that a book made one sad. That is because sadness is a mood; it does not have an object as such, though it

3. Kendall Walton, *Mimesis as Make-believe: On the Foundations of the Representational Arts* (Harvard University Press, 1990): 244.

4. Make-believe truths are generated according to the rules of the ‘game,’ which obviously varies, though there will likely be a set of standard rules, based on conventions, for each artistic medium, say; the work of art serves as a ‘prop’ in the game. For example, the slime moves towards the camera which generates the fictional truth that the slime is advancing towards the audience, and thus that Charles is threatened by the slime. This fictional danger gives rise in Charles to feelings of ‘quasi-fear’ which, combined with the fictional truth that he is in danger, make it fictional that he is afraid. What makes it fictional that he is afraid of the slime, is that fictionally he is threatened by the slime, and that this is what gives rise to his feelings of quasi-fear. Whilst it may sound somewhat convoluted, Walton suggests that this process is parallel to the case of ‘normal’ fear. (*Mimesis as Make-believe*, 249-5).



A New Approach to the Paradox of Fiction

may have an explicit cause. Pity, however, is pity of or for an object. This I shall take as an assumption in what follows. The worry motivating denials of (a), then, is that if states such as 'pity for x' are relational, then they cannot be found to be objectless. Thus if the apparent object of pity does not exist then either it is not pity or its object is not what it appeared to be.

If this is the case, then such let us consider self reports as Charles'

(1) I am afraid of the slime.

According to the above assumptions, (1) reports Charles as being in a relational state, which requires the existence of its object for its own existence. Therefore (1) can only be accepted by us as theoreticians on pain of implying

(2) The slime exists.

This, I will assume, is something we wish to avoid.⁵ The desire to preserve the relationality of such emotions as fear, without being committed to the inference from (1) to (2) has motivated Walton to provide an alternative analysis of self-reports such as (1). Walton's analysis has (1) mean

5. Having said this, to explain exactly how we wish to deny the slime's existence is somewhat tricky. Nathan Salmon for example, has suggested that fictional characters do exist as 'abstract entities' created by their authors ('Nonexistence,' *Noûs*, 32.3, 1998: 295). It seems, then, that we could claim that Charles is really afraid of the abstract entity which is the slime; a similar position is advocated by Peter Lamarque who claims that we are afraid of fictional characters in the sense of their being the 'intentional object' of the thought of the character ('How can we fear and pity fictions?' *Aesthetics and the Philosophy of Art*, 294). There is the problem here, however, that just as Charles does not feel himself to be make-believedly but really afraid of the slime, so he does not feel that he is afraid of the sense of a thought, or an abstract entity. His experience, we are claiming, feels like 'fear-of-the-slime,' where the slime very much resembles the monster he sees on-screen. Thus even if, with Salmon, we wish to say that fictional characters do exist, it is a much greater stretch to claim that they exist in a sense strong enough to claim that Charles is afraid of the slime which does justice both to how he feels and to our assumed relationality of fear. It is difficult to specify a sense in which fictional characters exist which does not succumb to revisionism with respect to self-reports.

(1') Make-believedly I am afraid of the slime.

This, to be sure, implies

(2') Make-believedly the slime exists.

But this is not a problem, because make-believedly the slime does exist. It is one of the simplicities of Walton's theory that he allows all interaction between the real and the fictional to occur in the same 'world,' that of the game of make-believe where both Charles and the slime exist; neither (1') nor (2') imply (2).

We can see, then, that Walton's account copes admirably with this shift of focus on the paradox. However, as I have stressed, it remains revisionist. I have not yet proposed an alternative, but I hope that, having re-oriented our perspective on the central problem of the paradox, we can now make some progress towards an alternative solution.

The Disjunctivist Insight

In this section I will attempt to sketch an alternative approach to the characterization of Charles' situation by appropriating an insight from a position in the philosophy of mind and perception known as 'disjunctivism.'⁶ It has been an implicit assumption of the above discussion that if Charles was in a state which he could not tell from fear of the slime, then this would imply that he was indeed in a state of fear of the slime. I will call this assumption the assumption of the 'transparency of phenomenology.' The above attempted to show that this is to be considered problematic, given

6. cf. Matthew Soteriou, "The Disjunctive Theory of Perception," *The Stanford Encyclopedia of Philosophy* (Winter 2010 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2010/entries/perception-disjunctive/>> I hope that at least in characterizing the 'disjunctivist insight,' I will have stuck to relatively uncontroversial ground within this position. For an overview of the development of this position, along with a discussion of some of the controversies within it, see the 'Introduction' to *Disjunctivism: Contemporary Readings*, ed. Byrne and Logue (London: MIT Press, 2009).



A New Approach to the Paradox of Fiction

the assumption that fear is an object-directed state. It is the apparent dilemma we face in characterizing Charles' experience that we either need to reject Charles' self-reports or reject our very plausible model of fear, on pain of having to accept (2). I would suggest that this apparent dilemma is what allows Walton and others to propose their accounts without much regard to Charles' phenomenology and self-reports; after all, if we acknowledge that he does feel exactly like he is afraid, then he must be, according to the assumption about phenomenology.

Our commitment to anti-revisionism leads us to accept that (1) is the natural and proper thing for Charles to say. We should try to take his report at face value. How then do we stop the inference to (2)? The reason for accepting (1) is that whatever the cause of Charles' state, it seems to him that he is in fear of the slime.⁷ Given the transparency assumption, to say that it seems to a subject that he is in fear of the slime is to report on a state of his phenomenology that is common between the veridical case in which

(3) I am experiencing fear with relation to the slime

and the illusory case in which

(4) I am subject to an experience indistinguishable from relational fear of the slime.

These two different cases share a 'highest common factor' and it is that highest common factor that makes it the case that (1) is true.⁸ The highest common factor is normally conceived in a mind-dependent fash-

7. In this respect, the treatment of (1) follows what many say about self-reports of perceptual experiences. Compare the case where someone says, when told that there is no lemon in front of them, "Well it certainly seems that there's a lemon, whether there is or not!"

8. John McDowell, "Criteria, Defeasibility and Knowledge," *Studies in the Philosophy of Logic and Knowledge* (Oxford: Oxford University Press, 2004): 22.

ion, such as a 'sense-data' or 'representational content.' For now, we can characterize it neutrally as

(1'') It is to me that I am fearful of the slime.

If this common factor is the core of the experience reportable by (1), then it implies (2). Accepting the transparency of phenomenology, (1'') means that (1) and so it seems to Charles that he is afraid (a state necessitated by our anti-revisionism), which means that he is afraid. And, in the case at hand, states of fearfulness are relational. So, without some alternative account of what it is to be fearful, (1'') still entails (2).

The 'disjunctivist insight,' which I wish to discuss here rejects the assumption of the transparency of phenomenology, and so denies that we can move from the fact that two states are subjectively indistinguishable to the claim that they share a highest common factor. Put another way: we should not individuate psychological states by reference to their subjective indistinguishability.⁹ This denial of the common factor between subjectively indistinguishable experiences allows for the genuine object-dependence of perceptual states even in the face of arguments from illusion or hallucination – just because I cannot tell the difference between veridical and hallucinatory cases, it does not mean that the veridical perception of an apple does not essentially involve that apple. If, then, (1) can be truthfully reported in the circumstance of (3) and (4), and we deny that (3) and (4) share any relevant common factor, then the claim (1'') that makes (1) true cannot be construed as a univocal state common between (3) and (4). It must be taken disjunctively:

(1'') = either (3) or (4)

9. cf. J.M. Hinton, "Visual Experiences," *Mind, New Series*, 76.302, (1967): 226 and M.G.F. Martin, 'The Limits of Self-Awareness,' *Philosophical Studies*, 120.1-3, (2004):37-89.



A New Approach to the Paradox of Fiction

And as it is (1'') that makes (1) true, the same applies for (1). The move from (1) to (2) is then blocked because (4), according to the disjunctivist insight, does not describe an object-dependent state, and so does not entail the existence of the slime. In short, when Charles truthfully self-reports (1), either he is reporting on a relational state involving slime, or he is reporting on something else, but the difference between the two is not transparent to him; for all he knows, it could be one or the other. There is no common state between the two.

If, then, we take this plausible account of the meaning of Charles' utterance as an expression of how he takes himself to feel, but persist with the disjunctivist insight that how Charles takes himself to feel is to be explained by appeal to one or another of two distinct types of mental state, then we can accept Charles' utterance of (1) without being committed to (2). Whereas the traditional (a)-denialist approach sought to prove that Charles does not seriously or genuinely take himself to be afraid of the slime because that would entail the existence of the slime, we can escape this revisionism with our disjunctive analysis.

It might be objected that we have not saved how Charles takes his experience to be – after all, he thinks he genuinely is afraid of the slime, which we deny – but what we have preserved is how it is for him in terms of what he thinks he is doing when he utters sentences such as (1). We then ultimately deny (a)-the claim that we have genuine emotional reactions to fictions - as we must if we wish to preserve the assumption of the relationality of fear whilst denying the relevant existence of fictional objects, but we have done so without the same revisionism with respect to Charles' self-reports. We achieve this, in part, by saying that things can be, from Charles' point of view, exactly as they would be if he were genuinely afraid of the slime without implying that the slime exists.

Though our analysis of (1) may seem counter-intuitive, it is not an analysis which is unique to the problem of fictionality, and if we accept disjunctivism about perceptual states then we will be committed to many

other such analyses. Furthermore, representationalist or sense-datum theorist analyses are hardly famed for their intuitive appeal. Regardless, my aim here is merely to show that, given the disjunctivist insight, we can provide an alternative approach to the paradox.

Conclusion

We have established that an alternative analysis of the paradox is possible on the disjunctivist model, which blocks the problematic inference from (1) to (2). We have thus removed a key motivation for solutions such as Walton's. The ultimate value of this account will rely upon a number of questions which have not been fully addressed here, such as the plausibility of the disjunctivist position with respect to subjective reports, and the plausibility of the analogy between perceptual states and certain emotional states. As my purpose here is primarily to propose an alternative approach to the paradox, I will not go into such matters here. I do wish, however, to anticipate two criticisms likely to emerge from the literature on the paradox.

First, although we have avoided the mutual incompatibility of the paradox's claims by ultimately rejecting (a), our account itself might be incompatible with (b) – the claim that belief in the object is a necessary condition for emotion. It would certainly render our solution inefficient if we found ourselves rejecting two of the statements of the paradox. As it is, we can show that (b) does not have any claim upon (4) and thus no necessary impact upon our analysis of (1) as (1''): simply put, (b) is a claim about real emotions, (4) is not. To elaborate, let us imagine a mental state called 'counterfeit-fear.' The minimal definition of counterfeit-fear is that it is a mental or psychological state which is subjectively indiscriminable from genuine fear, but which is not dependent on its 'object' for its existence; it is the sort of thing which is appropriately reportable by statements such as (4). My claim here is that (b), in its current formulation, has nothing to say about counterfeit-fear, and so does not pose a problem for anyone wishing



A New Approach to the Paradox of Fiction

to interpret Charles' utterance as (1''). To draw the analogy with the perceptual case, a veridical visual experience of a lemon requires the relevant existence of a lemon; an illusory experience of a lemon does not. On the disjunctivist model, to say that 'seeing a lemon' requires a lemon obviously does not make a claim about hallucinating a lemon.

Second, having introduced the term 'counterfeit-fear' might cause one to wonder more specifically what such a mental phenomenon might be like, and particularly whether or not there is anything to distinguish it from Walton's quasi-fear.¹⁰ The answer is that there is nothing explicit at this point neither in Walton's nor our present discussion to preclude the identification of counterfeit and quasi-emotions, and this may lead to worries.

However, quasi-fear alone is a rather formless feeling in Walton. It lacks intentionality and should be seen more as an ingredient of the experience of fear than as its manifestation. Walton explains quasi-fear with reference to its role in an experience of real fear:

"To be (really) afraid of a tornado, for instance, is to have certain phenomenological experiences (quasi-fear) as a result of knowing that one is endangered by the tornado. What makes the state one of fear rather than anger or excitement is the belief that one is in danger, and what makes the tornado its object is the fact that it is the tornado that one takes to be dangerous."¹¹

The role of quasi-fear here is that, in combination with beliefs or facts about its causes, it forms part of the experience of fear-of-x. It is not open to us to allow counterfeit-fear this sort of role, because by hypothesis

10. Walton says in direct characterization of quasi-fear is that it is a set of "certain phenomenological experiences" which arise "as a result of knowing or believing that one is endangered," and which make it make-believe that one feels fear (Walton, *Mimesis as Make-believe*, 244). Elsewhere it is suggested that quasi-fear in Charles' case is the relevant aspects of his 'physiological-psychological state' as he watches the film (Walton, *Mimesis as Make-believe*, 196). Both of these definitions are broad enough to encompass all we require of counterfeit-emotions.

11. Walton, *Mimesis as Make-believe*, 245.

it should be subjectively indistinguishable by itself from fear-of-x. What this boils down to is saying that counterfeit-emotions must be considered as having an intentionality which Walton denies (at least provisionally) to quasi-emotions. The challenge is thus to suggest a model upon which counterfeit-emotions could have some feeling of intentionality, though obviously they will not be 'about' their 'object,' as it is their object which we wish to deny. Again, having recourse to the perceptual case, a hallucinatory experience of a lemon will certainly feel exactly as though it is 'about' a lemon, even though there is no lemon. Either way, I find it quite implausible to suggest that one's psychological or phenomenological state when, say, afraid of a dog, does not at the very least include a disposition to react in a certain way to dog-like stimuli, and I believe the case could be made in much greater detail. The above does not constitute a full defence of the proposed solution, and is not intended as a full dismissal of Walton's account. I hope, however, to have shown that there is a plausible position which is capable of defence.

We have seen that there are problems with Walton's account, at least with respect to his revisionism regarding the subject's self-reports. Further, it has been shown that if we accept that there can be non-relational psychological states which are subjectively indistinguishable from genuine emotions, then the analysis of statements such as (1) as (1'') allows us to accept utterances of (1) made in response to fiction without falling into such revision. This move is one which has had great success in recent philosophy of perception, and it has been shown that it has promise in helping us to understand the logic and quality of fictional engagement, suggesting that Walton's Charles is best understood not as making-believe that he is afraid, but as mistaken about the nature of his own experience.¹² ❖

12. I am indebted to Prof. Michael Luntley, firstly for acting as supervisor for the University of Warwick URSS project which funded this work and secondly for his sustained and invaluable contribution and support in countless conversations going far beyond the requirements of this role. I would also like to thank Prof. Greg Currie and Dr. Eileen John for their illuminating comments on earlier drafts of this paper which led to a number of important improvements.



The Universe Began to Exist? Craig's Philosophical Arguments For A Finite Past

Blake McAllister

ABSTRACT: William Lane Craig offers two philosophical arguments for the conclusion that the universe began to exist. To be compelling, these arguments must not only be sound—we must also have reasons to believe that they are sound. I determine that these arguments do not provide such reasons to many individuals. The arguments ultimately rely on supposedly intuitively obvious absurdities. However, if one fails to see these ostensible absurdities—as many philosophers do—then for her, Craig's arguments lack all epistemic force.

The Kalām Cosmological Argument¹

1. Whatever begins to exist has a cause.
2. The universe began to exist.
3. Therefore the universe has a cause.

The Kalām cosmological argument has received considerable attention since William Lane Craig formulated its modern articulation. Interest in this argument has only increased with time, and understandably so. The Kalām has distinct advantages over other formulations of the cosmological argument. Primarily, the strength of the Kalām lies in the modesty

1. William Lane Craig, *The Kalam Cosmological Argument* (New York: Harper & Row Publishers, 1979): 63.



The Universe Began to Exist?

of certain metaphysical principles that underlie the first premise of the argument. While traditional cosmological arguments and arguments from contingency must depend on exceedingly strong and, hence, contentious formulations of the principle of sufficient reason, a much more effacing principle is expressed in (1). Though the adoption of such a principle alleviates the amount of work necessary to defend the first premise, it shifts the evidential burden onto the second.

Accordingly, Craig offers four arguments, two scientific and two philosophical, in support of (2). In this article, I will limit my examination to the two philosophical arguments—specifically the persuasiveness of these arguments. It seems safe to assume that Craig’s presentation of the Kalām is not solely for the purpose of formulating a sound argument; ideally, he wants to give reasons in support of his argument that should be, at the very least, minimally forceful to everyone who understands them.² Thus, it is both fair and worthwhile to evaluate whether Craig achieves this purpose. Keep in mind that an effective persuasive argument need not demonstrate the truth of its conclusion beyond all reasonable doubt, but merely show why its premises, and thus its conclusion, are more reasonable to believe than their denials. Therefore, if Craig’s two philosophical arguments are to succeed, he must provide reasons that philosophically obligate all evaluators who understand the reasons to accept them as forceful to at least a minimal degree. I will argue that neither of Craig’s two philosophical arguments in support of (2) meets this standard. While Craig provides reasons that should persuade some individuals, they are not of such strength that they are compelling for all reasonable individuals. In other words, certain evaluators are rationally justified in denying that Craig’s arguments have any epistemic force.

I will begin by examining the argument from the impossibility of the formation of an actual infinite by successive addition. The argument is as follows:³

2. If all people should find some reason R forceful, then no rational and honest person can properly understand R and proceed to reject R as having no epistemic force.

3. Craig, *Kalam*, 103.

4. The temporal series of events is a collection formed by successive addition.
5. A collection formed by successive addition cannot be an actual infinite.
6. Therefore the temporal series of events cannot be an actual infinite.

In order to understand this argument one must first comprehend the difference between potential and actual infinities. A potential infinite is a collection that increases in number indefinitely, always approaching infinity but never reaching it. This can be represented by a curve getting ever closer to an asymptote but never touching it. A potential infinite is in the process of becoming, moving higher and higher on the scale of natural numbers (1, 2, 3, ...), while an actual infinite is a completed totality equal in number to the entire set of natural numbers. If one requires a simple way to differentiate the concept of a potential infinity from the concept of an actual infinity, then just remember this: a potential infinity is merely indefinite, whereas an actual infinity is truly infinite. This distinction is important because the infinities most often discussed in mathematics are only potential infinities (∞) whereas Craig's argument deals with actual infinities (\aleph). Note also that there is no highest natural number, for no matter what natural number (x) you may consider, it is always possible to generate a higher number ($x + 1$). Thus, while the set of natural numbers in its entirety is an actual infinite collection, no natural number is the immediate predecessor of actual infinity.

Bearing this distinction in mind, I will briefly outline the argument for (5). If you form a collection by adding one member after another, each addition increases the number of members in the set by a finite amount. In other words, the number of members in the set progresses higher on the scale of natural numbers with each addition. Since no natural number is



The Universe Began to Exist?

the immediate predecessor of actual infinity, a set formed in this manner can never reach the point when the number of members in the set is equal to actual infinity. Hence, a collection formed by successive addition, even one progressing indefinitely into the future, would merely be a potentially infinite set. The principle that underlies this reasoning is often called the impossibility of traversing the infinite—it is impossible to progress from a finite set to an actually infinite set through successive addition.

This principle is certainly sound, but it is important to note that it only applies to finite sets. Thus, if this principle is going to serve as a part of a valid argument in support of (5)—a collection formed by successive addition cannot be an actual infinite—then Craig must make the additional assumption that

7. All collections formed by successive addition are finite at some point.

This assumption seems, at the very least, contestable. It is not immediately apparent why all collections formed by successive addition must be finite at one point, and, as far as I can tell, Craig offers no explicit reason to defend this assumption. We can certainly conceive of a collection formed by successive addition that was at no point finite—consider an actual infinite collection that has always been an actual infinite and is being added to successively.

In fact, Paul Draper points out that if the universe is eternal, then the temporal series of events in time would be such a collection. Draper is worth quoting at length:

If the temporal regress of events is infinite, then the universe has never had a finite number of past events. Rather, it has always been the case that the collection of past events is infinite. Thus, if the temporal regress of events is infinite, then the temporal se-

ries of events is not an infinite collection formed by successively adding to a finite collection. Rather, it is a collection formed by successively adding to an infinite collection. And surely it is not impossible to form an infinite collection by successively adding to an already infinite collection.⁴

This objection undermines support for (5) and consequently, the argument as a whole. Draper's objection does not show that Craig's argument is unsound, but in the absence of some independent reason for (7), it does prevent us from saying that it is more reasonable to accept the argument than to deny it.

Just because Craig does not offer a reason in support of his assumption does not mean that he cannot produce such a reason. So what might Craig say in defense of (7)? At first blush, it seems as if Craig might be tempted to respond by appealing to the word "formed." He might argue that if a set has always existed, then it cannot be formed in any relevant sense. However, this response fails. If any collection that is formed must have begun to exist, then proponents of an eternal universe would have no reason to accept (4)—the temporal series of events is a collection formed by successive addition. They would simply insist that the temporal series of events is not formed by successive addition. Events are being successively added to the temporal series, but the series itself is not formed. Anyone who did not already believe the universe to have a finite past would have no reason to accept (4), undermining the strength of the argument.

The only other response immediately apparent is to argue that it is impossible for any actually infinite set to exist at all. This, however, is Craig's next philosophical argument. If Craig does not use this line of reasoning in support of (7) and the former philosophical argument, then it seems as if we have been given no good reason why we should accept the

4. Paul Draper, "A Critique of the Kalam Cosmological Argument," in *Philosophy of Religion: An Anthology*, ed. Louis P. Pojman and Michael Rea (Belmont, CA: Thomson Wadsworth, 2008): 47.



The Universe Began to Exist?

argument as sound. On the other hand, if Craig does appeal to the latter philosophical argument to support the former, then these arguments are no longer logically independent as Craig claims that they are.⁵ In either case, it seems as if whether Craig succeeds in proving that the universe began to exist will be determined by the success of his next philosophical argument.

As mentioned previously, this argument is based upon the impossibility of an actual infinite set existing in the real world. Craig formulates it as follows:

8. An actual infinite cannot exist.
9. An infinite temporal regress of events is an actual infinite.
10. Therefore, an infinite temporal regress of events can not exist.

(8) is clearly the key premise of the argument, so I will give a concise presentation of Craig's argument in support of it. Craig defends (8) by offering a multitude of thought experiments. These thought experiments serve as *reductio ad absurdums*. They are meant to show the absurdities that would arise if an actual infinite existed in the real world. Craig sets up these scenarios, demonstrates certain logical implications, and then assumes the absurdities to be intuitively obvious.

Craig's favorite thought experiment is that of Hilbert's Hotel. In this experiment, Craig describes a hotel with an actual infinite number of rooms filled with an actual infinite number of guests. He then proceeds to demonstrate the absurdities that would arise if such a hotel were to exist. For instance, if all of the guests in the odd numbered rooms leave and all of the remaining guests move to the room number that is half of their current room number, then all of the rooms would be filled despite the fact that an

5. William Lane Craig, *Reasonable Faith* (Wheaton, IL: Crossway Books, 2008): 120.

infinite number of guests had checked out ($\aleph - \aleph = \aleph$). To further complicate the situation, if all of the guests from rooms four upwards checked out, then only three guests would remain ($\aleph - \aleph = 3$).⁶ This, Craig argues, is absurd.

At first glance, one might question whether these absurdities, even if genuine, are germane to the impossibility of an infinite temporal regress. It could be argued that Hilbert's Hotel demonstrates the absurdity of an actual infinite set whose members *coexist* in reality, but an infinite temporal regress is an actual infinite set whose members exist *successively*. To grasp this objection, we must first understand that the Kalām presupposes an A-theory of time.⁷ Craig explains that in an A-theory of time, "things/events in time are not all equally real: the future does not yet exist and the past no longer exists; only things which are present are real."⁸ Consequently, even if the temporal regress of events in time were an actual infinite, at no time would an actual infinite number of events coexist. This is certainly a marked difference between the actual infinite sets involved in Hilbert's Hotel and an actual infinite set of events in time; however, it remains to be seen whether this is a relevant difference.

There are initial reasons to think that this may, in fact, be a relevant difference. Most of the absurdities generated in Hilbert's Hotel are the result of inverse operations such as subtraction and division. Craig explains, "In trans-finite arithmetic, inverse operations of subtraction and division are prohibited because they lead to contradictions; but in reality, one cannot stop people from checking out of the hotel if they so desire."⁹ Notice, however, that these trans-finite, inverse operations are only applicable to actual infinities whose members coexist. If the members of an actual infinite set exist successively, then such operations are impossible, for no one can "take away" events that no longer exist. Thus, we might be tempted to

6. *Ibid.*, 118-119.

7. *Ibid.*, 121.

8. *Ibid.*, 121.

9. *Ibid.*, 120.



The Universe Began to Exist?

think that the supposed absurdities demonstrated by Hilbert's Hotel have no bearing on the possibility of an infinite temporal regress.

This line of reasoning, however, is off base. The originator of the Kalām, al-Ghazali, developed an argument to demonstrate that an actual infinite set of events in time entails the possibility of an actual infinite set whose members coexist.¹⁰ Imagine that every day God creates an immortal human being. If the universe has existed for an actual infinite number of days, then there would also be an actual infinite number of human beings coexisting in reality. Therefore, if it is impossible for an actual infinite set to coexist, it is also impossible for an infinite temporal regress to exist.

For the sake of clarity, however, I will introduce an additional thought experiment created by al-Ghazali that works directly with sets whose members exist successively. Imagine two planets that have been eternally orbiting the sun. The first planet requires only one year to complete a full rotation, while the second planet completes a single rotation every thousand years. If these planets have been orbiting from eternity past, then they have both completed an actual infinite number of rotations or, in other words, the same number of orbits, despite the fact that every thousand years the first planet completes one thousand times as many rotations as the second planet.¹¹ This, Craig claims, is obviously absurd. Many philosophers, however, simply do not agree.

In fact, a common rejoinder to such *reductios* has been to deny the absurdity of their conclusions¹²—a strategy Graham Oppy (humorously, I suppose) labels “outsmarting” one's opponent.¹³ In regards to al-Ghazali's orbiting planet, Oppy is quite content to embrace the ostensibly absurd conclusion. The planets have indeed completed the same number of rota-

10. I am thankful to Alexander Pruss for bringing this argument to my attention.

11. Craig, *Kalam*, 98.

12. J.L. Mackie, *The Miracle of Theism* (Oxford: Clarendon Press, 1982): 93; Jordan Howard Sobel, *Logic and Theism: Arguments for and Against Beliefs in God* (Cambridge: Cambridge University Press, 2004): 186-187; Wes Morriston, “Craig on the Actual Infinite,” in *Religious Studies* 38 (2002): 147-155.

13. Graham Oppy, *Philosophical Perspectives on Infinity* (Cambridge: Cambridge University Press, 2006): 48.

tions, but the set of all rotations completed by the first planet has a cardinality that is one thousand times greater than the cardinality of the set of all rotations completed by the second planet.¹⁴ According to Oppy, there is nothing absurd about this. Craig points out that a strategy of outsmarting one's opponent can be highly problematic since any position, no matter how obviously absurd, could be defended as long as its proponent is willing to bite the bullet.¹⁵ Therefore, we must determine whether the implications of al-Ghazali's orbiting planets, as well as other relevant thought experiments, are such intuitively obvious absurdities that those who deny them are either intellectually dishonest or significantly out of touch with reality.¹⁶

To gain a clearer understanding, let us examine exactly how these supposed absurdities are generated. The following discussion involves some basic concepts in set theory including one-to-one correspondence and proper subsets. One-to-one correspondence exists between sets A and B if and only if [iff] every member of Set A has one and only one corresponding member in Set B. Further, Set A is a proper subset of Set B iff every member of Set A is also in Set B and Set A is not identical to Set B. Craig attempts to explain the absurdities in his thought experiments by defining two principles.¹⁷

- i. Cantor's Principle of Correspondence. If one-to-one correspondence exists between two sets, then the number of members in each set is equal.
- ii. Euclid's Maxim. The number of members in a set is always larger than the number of members in any of its proper subsets.

14. In a very rough sense, the cardinality of a set is a measure of how large the set is. Oppy, 49-51.

15. Craig, *Reasonable Faith*, 119.

16. By 'out of touch with reality' I do not mean insane; rather, I refer to situations in which extensive isolation in the world of academia has greatly diminished the richness of one's intuitions such that he or she has lost even the most evident intuitions.

17. William Lane Craig and Quentin Smith, *Theism, Atheism, and Big Bang Cosmology* (New York: Oxford University Press, 1995): 23.



The Universe Began to Exist?

These principles certainly seem obvious, and they are constantly confirmed in our experience. We are not able, however, to endorse both of these principles at the same time when dealing with actual infinite sets. In other words, (i), (ii), and (iii) - There are actual infinite sets - form an inconsistent triad, such that endorsing all three at the same time entails a contradiction.¹⁸

Let us apply this analysis to al-Ghazali's planets. Examine the actual infinite sets of completed rotations for Planet 1 and Planet 2.

Planet 1: [1, 2, 3, 4, ...]

Planet 2: [1000, 2000, 3000, 4000, ...]

When comparing the two actual infinite sets of completed rotations, it is clear that the members of these sets can be placed in one-to-one correspondence with each other - for every member in the first set there is one and only one corresponding member in the second set.

Planet 1: [1, 2, 3, 4, ...]

↑↓ ↑↓ ↑↓ ↑↓

Planet 2: [1000, 2000, 3000, 4000, ...]

According to Cantor's Principle of Correspondence, these sets must be equal in number. Set 2, however, is a proper subset of Set 1, meaning that Set 1 will contain each of the members in Set 2 and many additional members. Euclid's Maxim dictates that the number of members in Set 1 is larger than the number of members in Set 2. Here we see the contradiction arise. The number of members in each set cannot be both equal and unequal. A contradiction of this nature will be generated anytime that (i), (ii), and (iii) are simultaneously endorsed. Thus, we must reject either (i) or (ii) when dealing with actual infinities.

18. Similar discussions of this triad can be found in Draper, 48, and Morrision, 154.

Craig argues that in the real world, Cantor's Principle of Correspondence and Euclid's Maxim cannot be reasonably rejected. We may be able to conceive of what it would be like to reject them in mathematical discourse, but when it comes to what is actually instantiated in reality, these principles cannot be denied. Hence, in order to avoid contradiction we must dismiss the possibility of actual infinite sets existing in our world. This is the very point on which many philosophers have challenged Craig. It is obvious and uncontested that (i) and (ii) hold for finite sets, with which we interact continuously in our lives, but why think that it is impossible that one of these principles be denied? What reason can Craig give to convince us that the Principle of Correspondence and Euclid's Maxim must hold for all sets in the real world?

Wes Morriston responds to this point by saying, "Craig's stock answer is to point once again to the intuitive 'absurdity' of infinite libraries and hotels and the like."¹⁹ Ultimately, Craig's claim will rest on intuition. I, for one, do not find arguing in this fashion to be inherently problematic—in fact it seems that virtually all arguments will come to rest on premises we take to be intuitively obvious; however, in such cases the reach of the argument only extends as far as the intuitions supporting it. If Craig's thought experiments do not seem intuitively absurd to an individual, as seems to be the case with many philosophers, then he has not offered any independent reason for why that individual should believe the situation to be absurd. Presumably these absurdities are not so evident that one would have to be intentionally deceitful or significantly out of touch in order to lack the necessary intuitions. It seems reasonable to assume that someone familiar with the branch of trans-finite mathematics could view Craig's thought experiments as merely drawing out intriguing implications of actual infinities in the real world.

19. Morriston, "Craig on Actual Infinite," 154.



The Universe Began to Exist?

It would seem, then, that Craig does not provide reasons in support of his argument that are minimally forceful to all rational observers. The force of his argument ultimately relies on an intuitive appeal. This in itself is innocuous, but it is problematic when it becomes clear that many individuals seem to reasonably lack the intuitions to make the appeal effective. The argument issues no epistemic obligation to those for which the “absurdities” are not intuitively evident and it appears that the number of people to which this applies is significantly higher than Craig would like.

In this article I have not tried to evaluate whether Craig’s two philosophical arguments are sound; rather I have argued that, for many, Craig does not provide strong enough reasons to think that they are sound. In the end, this is not a devastating conclusion for the Kalām. I tend to agree with Michael Bergmann in thinking that disagreement between two individuals, even radical disagreement, cannot always be traced back to irrationality or the use of an impermissible philosophical move.²⁰ It is vain hope to think that there are always going to be reasons that should be forceful for all rational evaluators. Still, anyone who does see the absurdities as intuitively obvious is obligated to affirm Craig’s argument as more reasonable to accept than to reject. As for those who do not possess such intuitions, Craig must provide some independent reason to support his claim before it will be reasonable for them to accept that the universe began to exist on the basis of these arguments. ❖

20. Michael Bergmann, *Justification Without Awareness* (Oxford: Clarendon Press, 2006): 231.

Book Review:

***Why Some Things Should Not Be for Sale* by Debra Satz**

Brittney Sovik

What is wrong with selling a kidney for some extra cash or putting a price on a woman's sexual services? Is it unethical to let children work or allow couples to hire a surrogate to carry their child? The rapid evolution of an expansive market system brings with it questions about the appropriate parameters of such a structure. Ought we restrict markets? In her new book *Why Some Things Should Not Be for Sale*, Debra Satz considers the ethical limitations of markets, and she looks closely at potential intuitive defenses of the sale of controversial products such as child labor, organs, reproductive services, and more.¹

At the heart of Satz's book is a critique of both contemporary economists and egalitarian political philosophers. Labeling them heterogeneous and unequal, Satz argues that markets should be treated asymmetrically.² A market in life-saving medicines differs from a market in bananas, and therefore they should not be treated alike. While contemporary economists tend to evaluate exchanges based only on efficiency, Satz considers the social context of individual practices and preferences. Satz also criti-

1. Satz, Debra. *Why Some Things Should Not Be for Sale* (New York: Oxford University, 2010): 4.

2. *Ibid.*, 93.



Book Review: Why Some Things Should Not Be For Sale

cizes the egalitarian argument that all problems in the market system can be solved through a redistribution of wealth.³ Many egalitarians reject market restrictions because they believe targeted action is less efficient than redistribution and limitations on free choice are affronts to personal liberty. Specifically, Satz finds philosopher Ronald Dworkin's hypothetical ideal market, wherein each citizen is given equal purchasing power and then "bids" on the resources s/he prefers, is inadequate when dealing with persons of disability, female caregivers, and individuals who make risky choices.⁴ In each of these cases, she argues that even under perfectly egalitarian circumstances the market can still produce systematic inequalities.

For Satz, a market is noxious when it creates or perpetuates inequalities between citizens.⁵ Satz identifies four characteristics that qualify a market as noxious: vulnerability, weak agency, extremely harmful outcomes for individuals, and extremely harmful outcomes for society.⁶ She uses these guidelines to demonstrate how markets in women's reproductive labor, women's sexual labor, child labor, voluntary slavery, and human kidneys are noxious and require regulation.⁷ Her treatment of each market varies, but in each case Satz works to dissect our intuitions that these markets are unacceptable. Many times she argues that our negative reactions are not a result of any essential feature of such markets;⁸ rather, they are unethical because of the social circumstances in which they operate. For example, in chapter five, Satz approaches the market in women's reproductive labor from a feminist perspective. As elsewhere, she argues for the asymmetry thesis, the view that markets in reproductive labor may be different from other markets.⁹ Satz claims that women's reproductive labor is not an intrinsically different form of manual labor, and yet the social

3. *Ibid.*, 63.

4. *Ibid.*, 70-1.

5. *Ibid.*, 94.

6. *Ibid.*, 9.

7. *Ibid.*, 99.

8. *Ibid.*, 94.

9. *Ibid.*, 115

context of the reproductive labor market expresses and reinforces a particularly pernicious form of gender inequality.¹⁰ In chapter six, Satz offers a parallel argument about markets in women's sexual labor. Prostitution is not intrinsically wrong, but in our current cultural context prostitution is a performance of female sexual servitude to men, and its legalization would have a negative impact on the perceptions and expectations of women as a class.¹¹

Satz's "social contextualization" approach to evaluating markets makes a valuable contribution to the debate regarding the ethics of markets. However, I do not think Satz's treatment of noxious markets sufficiently allows for the possibility of social change. Consider her argument against contract pregnancy. According to Satz, contract pregnancy is a pernicious market for three reasons: (1) "Contract pregnancy gives others increased access to and control over women's bodies and sexuality," (2) "contract pregnancy contributes to gender inequality by reinforcing negative stereotypes about women as 'baby machines'," and (3) contract pregnancy raises the danger that in contested cases of parental rights, motherhood will be defined in terms of genetic material in the same way as fatherhood, failing to recognize the unequal contributions of men and women to the birthing process (where women's gestational labor is not equivalent to a man's genetic contribution).¹² In response to these three concerns, I suggest that (1) contract pregnancy gives women a specialized medium for reclaiming control over their own bodies and its reproductive abilities despite social gender inequalities, (2) reproductive labor in certain forms reinforces negative stereotypes of men as sperm donors rather than active parental figures, and (3) all cases of contested parenthood, for both fathers and mothers, should consider more than just the genetic relationship to the child.

Specifically, when referring to the perpetuation of gender inequality in the reproductive labor market, Satz cites the unequal burden of men

10. *Ibid.*, 117.

11. *Ibid.*, 147.

12. *Ibid.*, 128-31.



Book Review: Why Some Things Should Not Be For Sale

and women involved in reproductive labor where a man's commitment in donating sperm is radically less than a woman's commitment to gestation and labor.¹³ However, the market recognizes this inequality by rewarding a female reproductive contract with a much larger sum of money than a man's less involved and less time consuming sperm donation. In this way, the market recognizes and rewards the biological differences in the reproductive labor of men and women. For Satz, surrogate pregnancy is unacceptable because it reinforces gender inequalities; if men and women held equal social positions, then surrogacy would be acceptable. But what if surrogacy could help improve the social position of women? Satz appears to underestimate and dismiss the possibility that markets sometimes promote social change.

Overall *Why Some Things Should Not Be for Sale* offers provocative arguments and fresh insights to discussions of the morality of the marketplace. Her objections to an unrestricted market system echo early liberal thinkers such as Locke and Hobbes who understood that in order to defend basic human rights, an individual must surrender some of her own freedoms to an authoritative power. For Satz, this means protection from noxious markets despite the infringement on an individual's liberty. Those interested in contemporary political philosophy will find Satz's book helpful in responding to a libertarian approach to the market system, and proponents of free market capitalism will be confronted with challenging arguments supporting the view that not everything should be commodified. ❖

13. *Ibid.*, 131.