

# *Ex Nihilo*

---

The Undergraduate Philosophy Journal of  
the University of Texas at Austin

---

The Undergraduate Philosophy Association  
Volume IV, 1999

---

# *Ex Nihilo*

---

**The Undergraduate Philosophy Journal of  
the University of Texas at Austin**

**Editor**

Christopher Furlong

**Selection and Editing Committee**

Jeffrey Casey

Jason Crickmer

Nirav Desai

Paul Domjan

Yve Golan

Justin Hall

Jason Mallory

Christopher Cade Mosley

Emily Nelson

Bryan Register

Sarah Register

Kimberly Ruyle

Jason Walker

**Faculty Reviewers**

R. L. Mackey

David Sosa

Mark Whitburn

R.J. Hankinson

**Administrative Advisor**

Cynthia Marshall

Volume IV, 1999

STATEMENT OF PURPOSE: *Ex Nihilo* is a publication that provides a forum for undergraduate students to explore and exchange ideas about philosophy. After 3 years as a departmental journal, this is the first volume of *Ex Nihilo* as a national undergraduate journal. The upgrade is a testament to the hard work of past editors and the commitment of the Philosophy Department at the University of Texas to the production of *Ex Nihilo*. We are confident that the high quality of the papers in this volume confirms that the Department's commitment was justified.

While University of Texas students receive no preference in the review process, 2 slots in the journal are reserved for the winners of the Matchette Essay Contest (a UT-Austin contest).

*Ex Nihilo* was produced by the Undergraduate Philosophy Association of the University of Texas at Austin under the auspices of the Department of Philosophy.

The opinions expressed in *Ex Nihilo* are those of the respective authors. They are not necessarily those of the Department of Philosophy, or the Undergraduate Philosophy Association.

CALL FOR PAPERS: *Ex Nihilo* will consider papers on any philosophical topic written by an undergraduate student. To submit, please send 3 paper copies and 1 disk copy (preferably in .rtf format) to:  
 University of Texas at Austin  
 Department of Philosophy  
 c/o Undergraduate Advisor  
 Austin, TX 78712-1180

Please do not include any identifying information on the papers. Instead, include a cover sheet with your name, address, school, phone number and email (if available). Include citations parenthetically; do not use annotated citation lists. Reserve footnotes for clarificatory and explanatory remarks only; do not use endnotes. Papers should not exceed 5500 words. The submission deadline for *Ex Nihilo* v.5 is January 28th, 2000. For additional information write to [ndesai@mail.utexas.edu](mailto:ndesai@mail.utexas.edu).

## Table of Contents

An Analysis of the Anthropic Principle and its Place in Modern Science <i>Steven G. Harms</i> .....	5
Frege's Ontology <i>Christopher Furlong</i> .....	20
Weakness of the Will: The Intelligibility of Immorality <i>E. Sonny Elizondo</i> .....	34
Lusting After the Absolute <i>Justin C. Hall</i> .....	43
Falsehood in Plato's Republic <i>Michael Keith Plocek</i> .....	53
Politics, Texts, Residues: Foucault's Philosophy and the Possibility of Social Activism <i>Rafeeq Hasan</i> .....	60

# An Analysis of the Anthropic Principle and its Place in Modern Science

Steven Harms

University of Texas

(Winner: Matchette Essay Contest)

## Introduction

In the beginning of the twentieth century, the academic and scientific communities faced a crisis in physics. Confronted with revolutionary implications from the empirical research of Edwin Hubble and theories of relativity of Einstein, the science of the day was forced to rapidly face one of the most significant paradigm shifts known in the modern age. In its wake, new sciences such as cosmology appeared and new methods of explanation such as quantum mechanics systematized the new knowledge.

Beginning in the mid-1970's, a subset of cosmologists broke away from the empirical tradition and devised a series of principles that attempted to explain the seemingly inexplicable properties of the universe. These principles, collectively known under the title, "The Anthropic Principle", are philosophical, not empirical, in nature. They offer a new and largely unexplored method of explaining the universe's properties. Cosmologists turned to logical conclusion-drawing versus trial experimentalism. In the Anthropic Principle, cosmologists proceed in an order of investigative steps exactly opposite to scientific tradition – from conclusion to premises. The question is: "Is this choice a change for the better?" This essay seeks to evaluate these principles by examining their meanings, pertinence, scientificity, and explanatory power. It aspires to conclude whether or not the Anthropic Principle has a valid place in cosmological research; and if so, what sort.

## The Anthropic Principle Defined<sup>1</sup>

The Anthropic Principle is classified into two super-classifications: the Weak Anthropic Principle (WAP) and the Strong Anthropic Principle (SAP). The SAP is further divided into four sub-

---

<sup>1</sup> Barrow and Tipler's work (Barrow, John D. and Tipler, Frank J. The Anthropic Cosmological Principle. Clarendon Press, Oxford, 1986.) is the most definitive book on the subject of The Anthropic Principle. As such, I have taken their presentation of the anthropic principle and will be using it as the source for my definition of the AP. Particularities in interpretation are noted in the footnotes.

*Steven G. Harms is a 4th year MIS/Philosophy major at the University of Texas. His interests include, among others, the pre-Socratics, the history of science and the relationship of information technology and man..*

classes: the teleological-theological model, the Participatory Anthropic Principle, the Participatory Anthropic Principle - 2, and the Final Anthropic Principle.<sup>2</sup> These latter three are 'flavors' of their super-classification - the SAP.<sup>3</sup> The following are the direct citations from Barrow and Tipler's book and define the forms of the Anthropic Principle.<sup>4</sup>

**WAP:** The observed values of all physical and cosmological quantities are not equally probable but they take on values restricted by the requirement that there exist sites where carbon-based life can evolve and by the requirement that the Universe be old enough for it to have already done so.

**SAP:** The Universe must have those properties which allow life to develop within it at some stage in its history.

**PAP:** Observers are necessary to bring the Universe into being.

**Many Worlds (PAP2):** An ensemble of other different universes is necessary for the existence of our Universe.

**FAP:** Intelligent information processing must come into existence in the Universe and once it comes into existence it will never die out.

The ethos of the above definitions is readily apparent and of great importance. That is, "Knowing I (a carbon-based organism / a human) am existent in the universe, the Universe must have already had such properties that allowed a life form such as myself to develop."<sup>5</sup>

<sup>2</sup> I have divided the Participatory Anthropic Principle in two according to interpretation of Quantum Mechanics (Copenhagen Interpretation or Many Worlds). For a description of the interpretations see Casti, John L., Paradigms Lost: Images of Man in the Mirror of Science. Morrow, New York, 1989.). PAP2 is my own identification and will be used throughout this work. In Barrow and Tipler it is simply referred to as letter B of the SAP division.

<sup>3</sup> Barrow and Tipler present the FAP as something independent of the SAP. In my opinion, this is done for aesthetic reasons. The FAP uses the same terminology "...life must come into existence..."(Barrow and Tipler, p. 23) and thus is a variant of the SAP. I feel Barrow and Tipler gave it its own bracket as it assumes the SAP's truth and it possesses a highly speculative nature. As such, the FAP does not belong at home with the two "blasé" scientific categories. I have denied the FAP this distinction as its philosophical definition places it in the same category with the other subclasses of the SAP.

<sup>4</sup> The following citations should all be attributed to Barrow, John D. and Tipler, Frank J. The Anthropic Cosmological Principle. Clarendon Press, Oxford, 1986, pp. 16,21,22-23.

<sup>5</sup> For a more precise definition of 'properties' see the following section

This ethos will prove a vital criterion for considering the validity of various subcomponents of the AP.

Barrow and Tipler's presentation of the AP as quoted above lacks the rigor of language necessary for a philosophical inquiry. As such, the following section will be devoted to establishing a vocabulary such that we can present clarified versions of the aforementioned definitions. To this end, a small analysis of the method and vocabulary of the modern law-forming science follows.

### The Universe and its Vocabulary in Terms of Cosmological Research and Mathematical Modeling

The universe is the closed system in which humans and matter are located. Phenomena (the reactions between matter) in the universe are governed by *laws of nature*. The universe's *laws of nature* are quantified by human scientists into mathematical expressions called *laws*.<sup>6</sup>

#### What is a Law? How is a Law made?

*Laws* parallel real-world phenomena's relationships through quantified, mathematical representations (equations) of said relationships. Consider the example of the process of translation from phenomenon (and correlate *law of nature*) to *law*:

**Given:** The moon orbits the earth. It holds its orbit due to a gravitational attraction between the two objects.

- **Real-World Happening:** "The moon orbits the earth." This is assumed to follow from an immutable law of nature.
- **Law (in hypothetical phrasing) Formation:** "The gravitational force exerted between the two bodies is inversely proportional to the square of the distance between them."<sup>7</sup>

At this point experimental testing is performed and the hypothetical law is found to be upheld in comparison to phenomena in

<sup>6</sup> The distinction between *law of nature* and *law* is a semantic choice. A *law of nature* was written by The Author of The Universe and is, most likely, not expressed in mathematical terms. A *law* is written by men and parallels the phenomena which results from said Author's law. This work will be concerned with the latter sort of *law*.

<sup>7</sup> Verbal definition of The Inverse Square Law taken from - Goodman and Russell. The Rise of Scientific Europe 1500-1800. Hodder and Stoughton, England, 1991, p. 254.

the real world. At this point, the result of the *law of nature* is translated into the language of mathematics and becomes a *law*.

**Mathematical Representation (law status):**  $F = k (m1) (m2) r^2$

### Components of an Equation

To aid understanding of the following distinctions, an example for consideration will be introduced. The following formula represents Newton's Inverse Square Law.<sup>8</sup>

**Ex. 1:**  $F = k (m1) (m2) r^2$

Laws are composed of constants and variables

### Constants<sup>9</sup>

There exists a finite, albeit possibly inconceivable, number of *constants*. The number of constants and their values are determined by whatever force or being created the universe. *Constants* possess values. They are required in the construction of *laws*. In the Universe, the assigned value for a constant is permanently fixed; examples are  $\pi$ ,  $k$ , or  $G$ . In cosmological modeling there is no difference between *constants* and *variables* - the cosmologist performing the modeling operation can freely assign whatever *values* he chooses.

### Values

*Values* are represented by a number on the number line from negative infinity to positive infinity. In the example,  $k$  is a constant that possesses a numerical *value* located in the domain of  $-\infty$  to  $+\infty$ . Neither *constants* nor *variables* have an *a priori* reason for preferring any given value.

### Variables

*Variables* are "temporary-constants." All *variables* possess *values*. In the example, variables are  $F$ ,  $m1$ , and  $m2$ . Unlike constants, they have no permanently assigned value in the "real world." In addition, they are equally susceptible to the value assignment whims of cosmologists. This lack of permanent value affixation is also true in mathematical representations (the equations) of *laws*. It is precisely this impermanence of value assignment and interdependence of some variables on other constants and variables in an equation-law that allows the effects of the *laws* and their constituent *constants* to be seen. For

<sup>8</sup> Formula taken from - Goodman and Russell. The Rise of Scientific Europe 1500-1800. Hodder and Stoughton, England, 1991, p. 254.

<sup>9</sup> Cosmologists occasionally differentiate between physical and cosmological constants. This difference will not be of note in this work.

the researcher, the possibility of changing a variable and seeing a quantifiable change in the *dependent* variable (the value of  $F$ ) is where the explanatory power of a *law* lies. Only a *dependent variable* is able to reflect, through changes of value, the phenomenon the *law* seeks to mirror. As such, all laws necessarily contain *variables*. By studying the mathematical relationship between constants and variables in the *law*, the researcher learns how the *law of nature* and the phenomenon relate to each other.

### How do variables receive values?

Some *variables* can be assigned by the person using the formula (the cosmologist doing the modeling). In modeling, some variables are assigned values by the cosmologist, while other values, have their values assigned (*dependent variables*) to them via the solution of the *law's* correlate equation. Using the example, the values of  $m1$  or  $m2$  may be changed according to the will of the person using the formula. The value of  $F$  is mathematically dependent and assigned as a result of the chosen values for  $m1$  and  $m2$ .

### What is the Role of the Cosmologist?

The principle work of the cosmologist is modeling. He builds representative models of the universe with the aid of computers. The models assume the truth (and are comprised of) of the *laws* which describe the interaction of matter within the Universe. These *laws form* the foundation in his modeling. The cosmologist is free to set the values of variables and constants as he pleases. The cosmologist could, in theory, assign  $\pi=5.5$ . All things are possible in modeling. Such silly assignments, of course, go against the cosmologist's best interest as he attempts to model the universe in which he finds himself.

With this, I retire the discussion of the manner of procession of the modern law-forming science, and return to the subject of the inquest: the evaluation of the APs.

### Translation of Barrow and Tipler's Presentation of AP

It has already been noted that Barrow and Tipler's definitions lack linguistic rigor. In need of clarification are the definitions of WAP and SAP. Given the terms defined in the previous section, a clearer version of both is possible.<sup>10</sup> These terms will also be presented in a manner slanted relative to cosmological research as that is the subject of this inquest.

<sup>10</sup> Henceforth, references to WAP or SAP will refer to the following definitions and no longer to those presented above in the citations of Barrow and Tipler.

**WAP:** In cosmological modeling, values for physical and cosmological constants may be freely chosen; however, the laws and models composed of said constants **must yield** a universe such that carbon based life evolves and that the Universe be old enough for such a universe to have already done so.

**SAP:** In cosmological modeling, the physical and cosmological constants **are required** to be such that life evolves therein at some point in history.

The answers to the “why *are* the constants required to be so” is what prompts the sub-versions of the SAP. The answers being either rooted in a teleological / theological / design argument answer or a quantum mechanics argument. The sub-definitions’ original phrasing is acceptable and needs no further clarification. Issues presented by the SAP will be discussed after the WAP is handled.

Evaluation of the Weak Anthropic Principle

The WAP offers the following contentions. First, it contends that all constants’ and variables’ values may be freely chosen from the domain of numbers. Further, it says that if cosmologists wish to render a universe like ours, the values of these constants and variables must be such that the laws allow two things to happen. First, carbon-based life forms must be able to come into being. Secondly, the universe must be old enough (either by assignment or as mathematical result) so that carbon-based life can already have evolved.

The WAP simply expresses something that everyone who is “alive” is aware of. Namely, the universe must have been such (carbon making and old enough, etc.) that humans could exist or else there would be no humans here to discuss whether it was so or not. This statement essentially resolves to a well accepted fact regarding all investigative instruments: “when... measuring anything, [it is] necessary to take into account the particular properties of the measuring instrument.”<sup>11</sup>

The mandate that the WAP delivers to cosmologists is:

- All the research and all the models you make must be such that human life is allowed to form. It merely limits the choice of values which cosmologists may apply to constants and variables.

<sup>11</sup> Casti, John L., *Paradigms Lost: Images of Man in the Mirror of Science*. Morrow, New York, 1989, p. 479.

- By the time the modeled age of the universe matches the current actual age of the universe life should have already come into existence.

Values are, of course, also limited by information acquired through other areas of empirical research.

This AP is surely true in all cases and most cosmologists will accept this as true despite some bitter foot shuffling and commentary that the AP “smacks of tautology.”<sup>12</sup>

The WAP’s Status in Terms of “Scientificity”

The standard of scientific investigation goes from specifying the initial situation and the laws of nature, to predicting subsequent states of the system.<sup>13</sup> The WAP operates in exactly the opposite manner. It takes the subsequent fact, “We are here” and deduces the initial conditions and the laws that govern the system such that the subsequent fact will be fulfilled.<sup>14</sup> Some sentiment against the AP may be an aesthetic disregard for violation of traditional scientific methodology. Critics also charge that wisdom gained by using the WAP does not explain all universes, just the one all humans find themselves in, and that this hypothesis is untestable as there are no “test universes” wherewith to experiment. The role of the cosmologist already establishes that the cosmologist seeks to describe the formation of the universe wherein he finds himself. As such, this criticism is baseless. Such complaints are not grounds for rejection of the principle. The WAP cannot be ruled out for “unscientificity.”

Explanatory Power of the WAP

The WAP lacks a certain “robust” explanatory power. It is highly unlikely that in mathematical research regarding the universe’s interactions the WAP will play any role. A day when a cosmologist programs the WAP into a universe modeling computer and the computer makes use thereof in a groundbreaking discovery is not imminently foreseeable.<sup>15</sup> The WAP has *dismissive* power and has value as a scientific research tool, but it fails to explain why things are as they are. It does not take the mystification out of problems. It is essentially an instinctive truth that everyone - to some degree or another - accepts as a requirement for their own existence. It will never

<sup>12</sup> Ibid., 481.

<sup>13</sup> Ibid., 480-481.

<sup>14</sup> Ibid., 480.

<sup>15</sup> This assuming a meta-logic idea such as the WAP could be programmed without running into Gödelian self-reference issues.

be used to arrive at particular values - its best function is as a razor, trimming bad explanations away leaving fewer bad explanations and the correct explanation.

#### How may the cosmologist use the WAP?

If the WAP proceeds against the general investigative pattern of science and it seems to lack mathematical appeal, how can it be of use in modern cosmology? It can be beneficial as a rejection-algorithm. It is an algorithm that helps reduce the choices of possibilities to the answer (provided, of course, that the disjunctive series is exhaustive). An answer can be found by empirical research and that answer's value is found to such and such or an answer can be found by removing all incorrect values until only the correct value remains. This "reverse-answering" process is how the WAP is best applied in research.

For example, Robert Dicke's employment of the WAP as a confirmation of Paul Dirac's value for the age of the universe is a shining example.<sup>16</sup> The age of a universe is represented by a value constant (remember: difference between constants and variables is insignificant in universe modeling as the universe created is dependent entirely on the cosmologist's input and the pre-programmed laws) and thus could possess any value from the domain of numbers. The choices from  $-\infty$  to 0 were ruled out, as the time value would necessarily be positive. Dicke took the ethos of the AP and asked, "We are here, what must the universe have been?" Banking on the accepted idea that heavier, life-necessary elements are cooked into existence within the nuclear furnace of stars or supernovae, Dicke was able to establish a bottom boundary value,  $\cong 10^{10}$  years.<sup>17</sup> Having established this, all cosmological models could be built in such a way that the age of the Universe would have to be greater than  $10^{10}$ . In addition, a time could also be found at which all stars would be burnt out. Such value would establish the top boundary for the age of the Universe - were such a value so high; cosmologists would not be here to discuss it. An acceptable and fruitful zone for experimentation has been determined by the WAP. A large number of values have been dismissed by use of this AP and a boundary on the range of values has been established. Through the Anthropic Principle, the boundary of  $10^{10} \leq \text{age of the Universe} \leq \text{time that all stars burn out}$  has been established.

<sup>16</sup> The story of Dicke's plan is taken from Casti, p. 480.

<sup>17</sup> This would be the estimate of how long it would take for organic molecules to be made and free from their stellar incubator.

#### The Verdict on the WAP

It is obvious and rather blasé, but philosophically and logically sound. It is a dismissive algorithm rather than a theory that helps construct new ideas and theories. The WAP is able to dismiss theories. While this is not a particularly preferred route wherewith to pinpoint a value, it does indeed help arrive at the correct value. The WAP is worth keeping as a cosmological explanation technique. Its use is valid. Like it or no, most scientists will allow its presence within the vocabulary of scientific explanation

#### Distinction between the WAP the SAP

What is the difference between the WAP and the SAP? The difference between the highlighted words "**must yield**" and "**are required**" seem minor. An example will illustrate the difference in strength between the two versions.

Suppose the gravitational constant G were a million times larger than it actually is. Then the lifetime of a star in its light-giving phase would be about a million times less... If an observer exists in such a universe he would see a universe whose mass would be a trillion times smaller than ours. Question: would life arise in such a vastly accelerated universe? The WAP is totally silent on this issue; the SAP says no, life can exist only if the fundamental constants have values very close to their observed [current] levels.<sup>18</sup>

The SAP was proposed in 1974 by Brandon Carter and is much more speculative than the non-controversial WAP.

#### Evaluation of the SAP

"An implication of the SAP is that the constants and laws...*must* be such that life can exist."<sup>19</sup> In other words, there is no possible universe in which life does not develop, and all cosmological models *must* always yield universes with carbon-based life because the cosmological constants *must* have values very close to those currently observed. The immediate question is "*why must these values be so?*" The varying answers to that question are the source of the radically different interpretations that spawned the different subclasses of the SAP. Two answers hinge on quantum mechanics while the other hinges on a teleological/theological/design argument basis.

<sup>18</sup> Ibid., 481.

<sup>19</sup> Barrow, John D. and Tipler, Frank J. The Anthropic Cosmological Principle. Clarendon Press, Oxford, 1986, p. 21.

**The Teleological - Theological - Design Argument**

Teleological Answer: The universe must have such values because it is the universe's purpose to create human life and its constants' values must accordingly produce and sustain human life.

Teleology asserts that everything in the universe has a "final cause" or a higher purpose. Everything possesses a *telos* or a goal which is that thing's reason for existence. This style of reasoning found its origin in Aristotle who stated all things have a higher purpose that they seek to fulfill. In accordance with its *telos*, the universe has assigned its composite constants values compatible with the production and sustenance of its *telos*-fulfilling end -- human life. This is a dated attitude that has no place in modern science. It is widely disregarded. This answer to the question "*Why must the values be such to produce life?*" of the SAP is unsatisfactory. While the WAP lacked certain robustness, this version of the SAP lacks credibility and believability. It is outright directly rejected.

Theological Answer: *The Universe's constants must have the values they do because of God's Plan*

Present or implied in virtually every religion is that God deliberately made the universe and the earth specifically for the domicile of the human race. Naturally, the constants and the laws of the universe and the universe itself have been made at the behest of the Creator for his creation. Accordingly all constants would be perfectly aligned because of the perfection of God's vision and the supreme mastery of His design. This obviously has no relevance to modern science in any way. Modern science takes place in an atheistic vacuum. This is a personal matter of faith and religion. This fails to answer the "*why must the values be such to produce life?*" question in a fantastic manner.

**Design Argument**

The design argument rests essentially on three key points: that order and harmony are worth appreciating, the classical "watch analogy", and probability. The watch analogy goes as follows. Were one to come across a beautiful Swiss watch, Tag Hauer for example, and one had never seen such a device before, one would be forced to conclude one of two things:

- Some agent created this masterwork of chronographic engineering.

- Through an exceedingly colossal accident this masterwork spontaneously came into existence.

Proponents of design arguments claim that one would naturally assume that it was designed by an agent because the probability of a chronographic marvel, such as the Tag, spontaneously coming into existence would be incredibly small. Proponents assume that the finder of such a device would have a 18th century - Newtonian appreciation for craftsmanship and order and accordingly appreciate the order, harmony, and precision encompassed in the object. The parallel, proponents state, is obvious. The universe - like said chronograph - operates in order, everything is perfectly related in a flawless manner, such majesty could not exist without the hand of an author.

This group of proponents also has modern disciples like Fred Hoyle. While studying nuclear resonance levels in carbon and oxygen Hoyle remarked:

I do not believe that any scientist who examined the evidence would fail to draw the inference that the laws of nuclear physics have been deliberately designed with regard to the consequences they produce inside the stars. If this is so, then my apparently random quirks have become part of a deep-laid scheme...<sup>20</sup>

At this point, the modern example and the Newtonian run into a small problem. At logical end, the design argument necessitates an omnipotent force and overlaps - although far less dogmatically - into the theological division's problems. It suffers from putting God into atheistic science. The Design Argument has a much more scientific subterfuge over its underlying theological core. As such, like the theological sub-class, it must be dismissed as well. In addition, there is no proof that rules out the possibility that a harmonious and orderly universe is not a natural effect of the big bang or any other happening(s) within the universe.

The teleological, theological, and design argument versions of the SAP have failed to explain why the Universe *must* produce life. New life came into the science world's lungs at the turn of the 20th century - quantum mechanics. Quantum mechanics offers recussitation to the ailing corpse of the SAP. In quantum mechanics, the modern day proponents of the SAP found a route by which to justify *why the Universe's constants must have the values that they do*. Their only problem was that quantum mechanics turned out to be slightly more complicated than originally thought. Quantum mechanics came forth in two forms or *interpretations*: The Copenhagen (or standard proposed

<sup>20</sup> Ibid., p. 22.

by Niels Bohr) Interpretation and the Many Worlds Interpretation (proposed by Hugh Everett).<sup>21</sup> The Copenhagen Interpretation is associated with PAP while the Many Worlds Interpretation is associated with PAP2. The two PAP versions of the SAP will now be handled.

PAP: Observers are necessary to bring the Universe into being.

This version of the SAP relies on the Copenhagen Interpretation. According to the Copenhagen Interpretation a quantum occurrence is expressed in the form of a wave function. This wave function “collapses” when an “observer” performs an act of measurement and the quantum occurrence is given a value. A simple example would be the spin of an electron. The two options are spin up and spin down. The value of the electron exists in a quantum state called *superposition* where the electron technically exists in a state of both spin-up and spin-down. When the observation is made, the wave function collapses and a value is given: spin-up or spin-down (represented by +.5 and -.5, respectively).

The way the Copenhagen Interpretation forms the PAP is by assuming that the Universe would exist in a similar state of superposition until the wave function is collapsed by an observer. In such a universe the “observer brings the Universe into being” - so maintains the PAP.

This belief goes directly against the entire ethos of the Anthropic principle and therefore must be disqualified (see initial definitions section in this work). The ethos maintained that because humans exist, the Universe must have already had such properties and been in such a state that humans would be brought into being. Quantum mechanics asserts “Because humans exist and collapse the Universe’s wave function, the Universe’s constants are given values.” Even if this conflict with the ethos of the AP were not to disqualify it from the set of Anthropic Principles, quantum mechanics itself presents several problems.

What is an observer? In quantum radioactivity experiments it is usually a photographic plate. Is that qualified enough to be an observer? Is a mouse qualified? Is a mentally incapacitated human? Is a blind human? The questions are literally endless.

There are logical problems as well. If all properties and all quantum events exist in a state of superposition until an observer brings them into being, how does the observer come into being? If all quantum states are in superposition are they still able to form observers? This

<sup>21</sup> If the reader needs to more information regarding the interpretations of quantum mechanics he can find information in Casti’s book in

view removes every thing to a world of superpositioned potentiality and fails to cite a genesis-actualizing act.

Further, assuming an observer can be brought into existence, as soon as the observer comes into being, he will collapse all the necessary wave functions necessary to bring a universal image into existence. The world will look to the observer (assuming he is conscious) like it was perfectly designed to bring him into existence -- the world will appear to be congruent with the necessary conditions for his existence. He will discover the WAP and the WAP will illustrate its intractable truth.

There are many other questions brought up within the science of quantum mechanics which, simply and regrettably, do not have answers. The most important conclusion is that even if there were no direct conflict with the AP ethos and even if one could manage to tolerate the great gray areas in quantum mechanics and even if one chose to accept the Copenhagen Interpretation, the observer will ultimately see a world compatible with his own existence. The observer will be aware of the truth of the WAP. Due to its vagueness and ethos contradictions, PAP is hereby rejected as an unsatisfactory answer to the question “*why the Universe’s constants must have the values that they do?*”

PAP2: An ensemble of other different universes is necessary for the existence of our Universe

The PAP2 finds its root in the Multiple Worlds theory. As above, this version of the AP is rejected out of hand due to its non-congruence with the ethos of the AP. Were that violation to be set aside it comes to the same end as the PAP1, as we shall see. Following Hugh Everett’s description of quantum mechanics each quantum act of measurement (each collapse of the wave function) splits the Universe into so many universes as there are options. Considering the example the electron spin measurement, the measurement act creates two new universes: one where the value for the electron is spin-up and another where the value for the electron is spin-down – each equally real. Ultimately the same results surface as in the PAP1. In whichever of the infinitely growing universes an observer emerges, he quite naturally sees a universe designed and directly compatible with the his existence. He validates the truth of the WAP for himself.

Further, the language of this principle seems to be in contention as well. The definition stated that an ensemble of worlds was necessary to bring our habitable world (or, in a larger sense *any* habitable world) into existence. This assertion is certainly not shown to be necessary in the Multiple Worlds theory. If there is one more or one less quantum event in the universe, which induces one more or one less split in the

universes, it is hardly apparent that the habitable universes would be maligned by said event.

### The Final Anthropic Principle<sup>22</sup>

The FAP is the most speculative of all AP's and the strongest version of the SAP. The FAP accepts the SAP as true and then uses a logical conclusion to assume the permanence of human life in the Universe. The FAP consists of this statement: "As it is the Universe's goal (in one way or another) to support our life, it is impossible for the human race to die out, lest the Universe negate its success by killing its goal."

- In the theological version, God would not eradicate His creation.
- In the teleological versions, the universe would not negate the success of achieving its telos.
- Design arguments follow the same logic. Humans are considered part of the harmony and have the role of appreciating the majesty and order in the universe. As such the universe would not eradicate humans. It can also rely on the theological as, ultimately, The Author of the universe is the same as God.
- PAP1 concludes that human life can't die out lest the Universe throw itself back into a superpositioned no man's land.
- For PAP2 it makes no difference because given that all combinations of all possibilities are represented life will necessarily emerge in one of the many universes. Persons who believe in PAP2 can always take shelter in the fact that there's a universe where everyone is happy and the human race will live safely and happily in peace forever. The universe will also find this a pleasant arrangement as it guarantees it an observer.

Not only can life not be destroyed after so long, but it will accumulate all the bits of information knowable in the universe and truly have a "mind of God." This is speculative garbage.

In the end, The Final Anthropic Principle is over-investing in the speculation of the SAP. As the SAP was rejected in all its forms, the FAP must be rejected as well. In the least case, the FAP is already far too speculative in its own right.

<sup>22</sup> Earlier I stated that I would treat the FAP as a part of the SAP. I have given closing remarks over the SAP in the previous paragraph because those comments hold for all the SAP sections, while not the FAP. The FAP should still be seen as a extremely strong version of the SAP. Due to its phenomenally speculative version I will handle it this separate section. It is an organizational convenience, not a philosophical division

### The Verdict on All SAP Subclasses

The theological/teleological/design arguments have been rejected due to their lack of substance. The quantum mechanics versions have been rejected due to the fact that their nature is directly against the ethos of the AP and because the science in which they are rooted leaves far too many questions and discontinuities to be counted as a dependable theory. In addition, both quantum versions -- regardless of interpretation and the science's vagueness -- reduce to the fact that any observer simply sees a universe consistent with this existence -- regardless of interpretation.<sup>23</sup> Fortunately for proponents of the WAP, the SAP upholds the truth and universality of the WAP.

### Conclusion

After considering all the varying varieties of the Anthropic Principles known to modern cosmology, the SAP's mandate "That the constants *must* possess the values they currently possess" was found not to be upheld by any of its composite three sub-classes. The SAP is removed from consideration entirely. The FAP was also quickly removed for its speculative and, quite simply, unrealistic ideas. The only AP to withstand this investigation was the Weak Anthropic Principle. The Weak Anthropic Principle is a principle which cosmologists may apply to their research without coming into philosophical error. The WAP must be used as an excluding principle -- it is a winnowing tool. With the WAP cosmologists can exclude values and come closer to finding the answers they seek. In all, however, the AP is a very weak tool and can only help provide a very preliminary narrowing of possible values. While it offers an interesting inversion of the traditional steps involved in scientific inquiry, empirical research offers a more reliable and generally more accepted route of inquiry.

<sup>23</sup> Were there no ethos violation with the PAPs they would actually answer the question "why must.." The answer would be because in both of those states -- a superpositioned universe and a multiple worlds universe -- an observer must ALWAYS come into existence. This is due to the fact that all outcomes for all quantum events are represented. The observer will always come into being and will of course see that the constants could not be any different lest he not be there. By covering the entire area of possibilities, quantum mechanics guarantees itself an observer.

Christopher Furlong  
University of Texas

(Winner: Matchette Essay Contest)

### I

In "Frege on Knowing the Third Realm," Tyler Burge argues for a realist reading of Frege's philosophy. He characterizes as realist any philosophy that accepts the existence of non-spatio-temporal entities that are independent of any kind of human activity, social or mental. Burge argues that, for Frege, numbers and thoughts, among other things, are non-spatio-temporal objects that exist independently of human beings and that therefore Frege is a realist. Hans Sluga and Joan Weiner have rejected this traditional reading of Frege and argue instead that while Frege does deny that thoughts and numbers are physical or mental entities he does not think that they are completely independent of the mind. Rather they are "constitutive of the mind" in that they, in some sense, structure our cognition and make it possible. In this paper I will consider these two readings of Frege and argue that Frege is a realist. I do not, however, entirely agree with Burge's reading. For Burge, Frege's realism is a naïve pre-philosophical assumption wherein the existence of abstract objects<sup>1</sup> is uncritically taken as a starting point. I will argue instead that Frege is led to his ontology by his theory of meaning.

Before I move on to my analysis of various interpretations of Frege's ontology, I should point out that giving a list of entities whose existence Frege is committed to will not do as an explanation of Frege's ontological commitments, nor would a list of the properties of those entities remedy the problem. We certainly need to know what Frege would admit of as existent and we need to know what properties these entities have, but for an explanation to be complete we need to know how Frege *arrives* at them. We need to know their role in Frege's arguments. Are they the basis or the result of his arguments? Burge makes his position clear in his disclaimer:

Although I think that Frege maintained a metaphysical view about numbers and other such entities, I do not believe that this view

<sup>1</sup> Frege never (to my knowledge) uses the term 'abstract objects', and I'm not entirely sure what it means or what kind of philosophical baggage it brings with it so I'm a bit wary of using it. But for the sake of brevity and ease I will use it strictly as a short hand for 'ontologically independent, non-spatio-temporal objects' I reject any connotations the term may bring with it.

*Christopher is a senior philosophy major at the University of Texas at Austin. He is primarily interested in the philosophy of language and epistemology. He plans to pursue a Ph.D. in philosophy after graduation.*

dominated his thinking. His is, for the most part, the relaxed Platonism of a mathematician who simply assumes that there are numbers, functions, and so on, and who regards these as an abstract subject matter which can be accepted without special philosophical explanation.<sup>2</sup> (Burge 1997, p. 4)

For Burge, Frege starts out with abstract objects and builds from there. It will be seen, I hope, by the end of this paper why Burge's characterization is inadequate.

Burge argues for his interpretation in two ways. First he cites and discusses several passages from Frege's writings that support a realist interpretation. Second, he argues that Frege's argument for the possibility of communication and the objectivity of science require the independent, non-spatio-temporal existence of thoughts.

Burge begins by considering two Fregean analogies between abstract and physical objects. Frege writes about numbers:

Just as the geographer does not create a sea when he draws boundary lines and says: the part of the ocean's surface bounded by these lines I am going to call the Yellow sea, so too the mathematician cannot really create anything by his defining. (Frege 1964, p. xiii)

and about thoughts:

The grasp of a thought presupposes someone who grasps it, who thinks. He is the owner of the thinking, not of the thought. Although the thought does not belong with the contents of the thinker's consciousness, there must be something in his consciousness that is aimed at the thought. But this should not be confused with the thought itself. Similarly Algol itself is different from the idea someone has of Algol. [Frege 1997, p. 342]

These analogies suggest, according to Burge, that numbers and thoughts are independent of thinkers *in the same way* that physical objects are. In addition to comparing abstract objects to physical ones, Frege also makes unqualified claims about such objects and their independence from us. Again, with respect to numbers, he writes:

Numbers do not undergo change, for the theorems of arithmetic embody eternal truths. We can say, therefore, that these objects are outside time;

<sup>2</sup> In the debate over Frege's ontological commitments and concerns, interpreters on both sides of the issue seem divided over whether to use the term 'realist' or 'platonist.' As far as I can see, nothing turns on this disagreement. I will use the former term only because I think (perhaps erroneously) that it carries less unwanted philosophical baggage. I take the two terms to be interchangeable

and from this it follows that they are not subjective precepts or ideas. (Frege 1984, p. 230)

and with respect to thoughts:

What we want to assert in using that proposition [that the number three is prime] is something that always was and always will be objectively true, quite independently of our waking or sleeping, life or death, and irrespective of whether there was or will be other beings who recognize or fail to recognize this truth. (Frege 1984, p. 134)

In these passages Frege is quite clear on how he thinks of numbers and thoughts. Numbers are atemporal and objective as opposed to mental. Atemporality presumably entails independence from humans since humans could cease to exist. A thought (that which is asserted by a sentence) is true eternally and independently of us.

Burge's second argument deals with Frege's notion of objectivity. On Burge's interpretation, Frege's arguments for the objectivity of science and the possibility of communication presuppose that thoughts are abstract objects. In his essay "The Thought," Frege says that "if every thought requires an owner and belongs to the content of his consciousness, then the thought has this owner alone; and there is no science common to many." (Frege 1997, p. 336) Frege argues for this view in three ways. First, since if a true thought follows from other true thoughts then it always follows from them, then it follows that if a thought is true at all it must always be true. To account for the atemporality of entailment, thoughts (the bearers of truth and falsity) must be atemporal.

In the second argument, Frege claims that science and communication are not possible if thoughts are ideas in people's minds. After asserting that thoughts cannot be physical he argues that if thoughts (the bearers of truth and falsity and the aim of science) were ideas, then no contradiction between one person's science and another's would be possible. Any dispute over truth and falsity would be "idle" and "ludicrous." So for science to be objective, thoughts must belong to a non-mental, non-physical realm. "Anything belonging to this realm has it in common with ideas that it cannot be perceived by the senses, but has it in common with things that it does not need an owner so as to belong to the content of his consciousness." (Frege 1997, p. 337)

The third argument for thoughts as abstract objects starts with the claim that facts are required for science to have a "firm foundation." But for Frege, facts just are true thoughts. The work of science, then, is to discover true thoughts which must, for science to be possible, be independent of our "varying states of consciousness." One

thing to note about this argument is that a true thought is not *about* facts, in which case science would not require them to be independent of the mind. Rather "a fact *is* a thought that is true." (Frege 1997, p. 342, my emphasis) This oddity is purely terminological, not philosophical. As has probably already become clear, Frege does not understand by 'thought' what we normally do. A thought, for Frege, is more akin to what we mean by 'proposition'.

## II

I will now consider two challenges to the view of Frege as a realist beginning with Sluga. The most charitable way to present Sluga's arguments for a non-realist interpretation is to consider them as a response to Burge's first argument which isn't really an argument but rather an inventory of isolated passages that seem to suggest a realist interpretation. The problem with Burge's strategy is that an interpreter such as Sluga can easily come up with a list of isolated passages that suggest a different interpretation. Consider for example, the following passage from *The Foundations of Arithmetic*:

I understand objective<sup>3</sup> to mean what is independent of our sensation, intuition and imagination, and of all construction of mental pictures out of memories of earlier sensations, but not what is independent of the reason, --for what are things independent of the reason? To answer that would be as much as to judge without judging, or to wash the fir without wetting it. (Frege 1980, p. 36)

It seems that perhaps thoughts and numbers are not so independent of us after all. Burge repeatedly claims that Frege makes no qualification on the independence of abstract objects. He goes so far as to make the highly informative assertion that for Frege, "independence is independence." Unfortunately it is not informative enough. For all the seemingly clear talk of the independence of abstract objects, this passage calls into question Burge's claim that Frege's realism is unqualified.

Up until now, what Frege seemed to mean by saying that abstract objects are independent was that they are ontologically independent; that even if the mental and physical realm were to cease to exist, the third realm would be unaffected. But what does Frege mean by stating that abstract objects are not independent of reason?<sup>4</sup> And

<sup>3</sup> For Frege, what is objective includes both the physical realm and the "third" realm. Abstract objects, then, are objective.

<sup>4</sup> I should note that Sluga's translation of the passage ends "...to say what things are like independent of Reason would be to judge without judging..." In this translation there is an ambiguity over whether independence from reason applies to

what does he mean by 'reason'? In Sluga's gloss to the passage above, he claims that Frege's point is that "it is inconsistent for us to try to say what things are in themselves, independent of our judgments." (Sluga 1980, p.120) To understand what Sluga means, and why he says it, we should consider his general view of Frege.

Sluga views Frege as a Kantian idealist of sorts and rejects the claim that he was concerned with or committed to any kind of ontology. In favor of his interpretation, Sluga enlists a large amount of non-textual evidence concerning Frege's personal associations, memberships in philosophical organizations and the general philosophical climate of late 19<sup>th</sup> century German academia. Very little textual evidence is given, making Sluga's arguments quite circumstantial and of questionable merit. While there is little question that philosophers are strongly influenced by the attitudes and beliefs of their times, evidence dealing with their historical context should never override textual evidence. Sluga would probably respond by saying that his interpretation of Frege proves its worth by allowing him to make sense of some otherwise cryptic passages. We shall see.

The piece of non-textual evidence that Sluga discusses at greatest length is the similarity between Frege's and Herman Lotze's view of objectivity. Sluga presents Lotze's view and, based on *some* explicit agreements with Frege's, concludes that Frege shares Lotze's ontological concerns or lack thereof. For Lotze, "the doctrine of objectivity is not to be taken as ontological, but rather as epistemological."<sup>5</sup> Lotze is quoted as claiming that

In so far as we have and grasp ideas, they possess reality as events, as things that happen in us ... but their content, considered separate from the mental activity we direct towards it, is not something that happens. It does not exist in the way in which things exist; it is simply *valid*. (his emphasis)

One key point of agreement between Frege and Lotze is, according to Sluga, that "the objective is that which can be grasped by more than one human (rational being). The objective, in other words, is the intersubjective." Frege certainly holds this since he says that a

---

things or to saying what things are like. If Frege meant the latter, then there is no problem for the realist interpretation. I don't know German so I can't comment on this ambiguity, but I will be charitable to Sluga and assume that Frege means the former.

<sup>5</sup> All the quotes in this exposition of Sluga will be from Sluga (1980), p. 117-121. I will mention in the text the cases where Sluga is quoting Lotze.

judgeable content<sup>6</sup> is "something objective, that is, something that is exactly the same for all rational beings that are able to grasp it." This seems to support Sluga's claim, but Frege adds "...as, say, the sun is something objective." [Frege 1979, p. 7] Now certainly, the sun is not "simply valid."

In "Objectivity and Reality: Lotze and Frege," Michael Dummett suggests that Lotze's view of objectivity entails a three-way ontological distinction between what is not objective (such as subjective ideas), what is objective and real (such as the sun) and what is objective and simply valid (such as thoughts, in Frege's sense). Frege, however, "admits no category of the intersubjective intermediate between what is private to some individual subject and what is independent of all subjects." (Dummett 1982, p. 117) That is not to say, of course, that thoughts have the same ontological status as the sun since neither is the sun atemporal nor thoughts physical. Frege's view, like Lotze's, does entail a three-way ontological distinction between ideas, physical objects, and abstract objects. But unlike Lotze, Frege's third-realm entities are not simply valid. Physical and abstract objects are ontologically independent of judging subjects *in the same way*.

Frege's theory of objectivity does hold that what is objective "is that which can be grasped by more than one human (rational being);" that what is objective is intersubjective. But what is objective, be it thoughts or celestial bodies, is intersubjective only because it is entirely independent of us. Frege writes:

Not only do thoughts – e.g. natural laws – not need to be recognized by us in order to be true: they do not need to be thought by us at all. A natural law is not created by us, but discovered. And just as a desolate island in the Arctic Ocean was there long before it was seen by men, so the laws of nature and in the same way those of mathematics hold from eternity and not just from the time of their discovery. We may deduce from this that thoughts are not only true, when they are true, independently of our recognition, but that they are altogether independent of our thinking. (Frege 1979, p. 133)

According to Sluga, Lotze thinks that the claim that ideas exist separately but analogously to physical objects is "outlandish," and "...he believes with Plato that empirical knowledge of temporal, changing things presupposes some knowledge of non-temporal, non-changing things" making him an "epistemological platonist rather than an ontological platonist."

---

<sup>6</sup> A judgeable content is just a thought and its truth value. By the time he wrote "The Thought," Frege had split up the notion of "judgeable content" into a thought and a truth-value. Thoughts and judgeable contents, presumably, have the same ontological status.

As we've seen, Frege would agree that knowledge of temporal, changing things presupposes knowledge of non-temporal, non-changing things. But when I considered Frege's argument for this view from "The Thought," I took it to be an argument that Frege is a realist. After all, if knowledge of temporal, changing things presupposes knowledge of non-temporal, non-changing things, then non-temporal, non-changing things must exist. Else, how could we have knowledge of them? Sluga seems to be confusing the issue of motivation with that of commitment. No one, I think, would say that Frege's motivations are ontological. Frege's project is not intended to give an ontological foundation for arithmetic. His motivation is epistemological. He wants to show that we *know* arithmetic in the same way we *know* logic. But it does not follow from this that Frege does not have an ontology. In fact, based on his views of thoughts and how they ground objectivity, he actually *needs* an ontology.

There is another problem with Sluga's interpretation of Frege. He views Frege as a Kantian idealist. What does this entail? On Kant's view, it is nonsensical to speak of things-in-themselves since most of what we know about objects is contributed by us. All we can know about the noumenal world is that it exists. Everything else we know about it, we actually know about, or as a result of, the structure our mind imposes on the it.

This would suggest a plausible interpretation of the passage from the *Foundations* quoted above since nothing of which we have knowledge is independent of the structure our mind imposes on it. But the interpretation loses its plausibility when we realize that Frege simply does not say enough. If Frege really was a Kantian idealist then we would expect him to discuss how his ideas fit into Kant's system. What role do thoughts and numbers play? Are they mind-imposed structures, are they phenomenal or are they noumenal? Not only does Frege not answer or suggest an answer to these questions, he doesn't even adopt Kant's terminology. To my knowledge, the only uniquely Kantian terminology that Frege does adopt are the terms for the analytic/synthetic and *a priori/posteriori* dichotomies over which he is in disagreement with Kant. If Frege discussed his disagreement with Kant in Kantian terms, does it not seem plausible to assume that if his agreement with Kant were substantial he would discuss that agreement in Kantian terms? There is, of course, a Kantian flavor to Frege's preference for 'thought' and 'concept' over 'proposition' and 'property'. But I do think Frege says enough about thoughts and concepts to preclude the possibility that they are in any way mental or mind-dependent as Kant would have them.

Non-textual evidence proves its worth in interpretation by yielding a coherent interpretation that explains certain textual evidence that could not otherwise be explained. For this purpose Sluga quotes, in

addition to the passage from the *Foundations* quoted above, (which he in fact does not explain) the following passage from "The Thought." I here quote the passage along with Sluga's gloss.

Frege notes that to say we grasp objective thoughts is to speak in a metaphor. 'What I hold in my hand can be considered the content of my hand, but it is the content of my hand in quite a different sense and is more alien to it than the bones, the muscles of which it consists, and their tension.' If we take this analogy seriously, it seems to imply that Frege does not hold that thoughts are in the mind as the bird is in the hand, but rather as the muscles and bones are in the hand. The objective is not something alien or external to the mind, but constitutive of it. It is its most characteristic possession. (Sluga 1980, p. 121)

On Sluga's reading, the bones and the muscles of the hand are analogous to thoughts and the hand is analogous to the mind. Dummett claims that Sluga misread the passage. For Dummett, Frege thinks that thoughts "are not mental contents, like ideas, that go to constitute our consciousness, but objects existing independently of us which we grasp in a sense analogous to that in which the hand may grasp a cricket ball." (Dummett 1982, p. 124fn) The analogy in question appears as a footnote to the following passage:

We do not have a thought as we have, say, a sense impression, but we also do not see a thought as we see, say, a star. So it is advisable to choose a special expression; the word 'grasp' suggests itself for the purpose. (Frege 1997, p. 341)

And is preceded by the following sentence:

The expression 'grasp' is as metaphorical as 'content of consciousness'. The nature of language does not permit anything else. What I hold in my hand... (Frege 1997, p. 341fn)

I think it is quite clear that what Frege meant was that while a thought "can certainly be regarded as a content of my [consciousness]; ...it is a content of my [consciousness] in quite another and more extraneous way than are the [ideas] of which [my consciousness] consists." Besides, if Frege meant that thoughts are contents of the mind as bones are contents of the hand, the word 'grasp' would not "suggest itself for the purpose." The fact is that Sluga simply does not make a convincing case for taking Frege to be a Kantian idealist.

In *Frege in Perspective*, Joan Weiner proposes an interpretation of Frege similar to Sluga's. She claims that "The Thought" is not about the possibility of objective knowledge or about the sense of assertoric sentences. Rather, "The Thought" provides another elucidation of the

view that there is a substantive logical source of knowledge without which ...even to think at all would not be possible."<sup>7</sup> What she means is that Frege does not ground objectivity on thoughts as abstract objects but rather "on what is required for grasping a thought." That is to say, for someone to understand the thought expressed by 'Frege is bald', one must understand the basic logical laws. After all, "someone who claims that Frege is bald and that he is also not bald can be said not to understand the thought expressed by 'Frege is bald'."

Frege would agree that sense impressions alone are not sufficient to give us knowledge of the external world. In his anti-solipsism argument in "The Thought," he says that "having visual impressions is certainly necessary for seeing things, but not sufficient. What must still be added is not anything sensible." (Frege 1997, p. 343) Weiner suggests that "the obvious candidate [for this non-sensible thing that makes knowledge possible] is the logical source of knowledge." But what is the logical source of knowledge? All Weiner says by way of explanation is that it "can be tied to Kantian analyticity." (Weiner 1990, p. 71) Presumably what this means is that the logical source of knowledge is that which allows us to recognize the truth of analytic statements. Since, as we have seen, Frege thinks that the question of truth arises only for thoughts, the logical source of knowledge would have to be that which allows us to grasp a set of thoughts (those expressed by analytic sentences) as true. On Weiner's reading then, the argument of "The Thought" is that all knowledge presupposes or requires the ability to grasp a set of thoughts as true.

Suppose I grant this. After all, Frege does say that "to the grasping of thoughts there must then correspond a special mental capacity, the power of thinking." (Frege 1997, p. 341) All Weiner has established is that Frege thinks that all knowledge presupposes or requires the ability to think. What of it? Weiner seems to make the same mistake as Sluga in confusing questions of motivation with questions of commitment. I think it is quite clear that Frege's theory of objectivity is an epistemological theory. It is intended to answer questions regarding the kinds and possibility of knowledge. But none of Weiner's arguments establish (or even suggest) that Frege's theory does not have any ontological commitments. In fact if all knowledge requires the ability to grasp a set of thoughts as true, must not those thoughts exist? Else, what sense would it make to speak of such an ability? So even on Weiner's interpretation Frege is still committed to the existence of thoughts. That he thinks thoughts are abstract objects is clear from his arguments on the possibility of science discussed above.

<sup>7</sup> All the quotes in this exposition of Weiner will be from Weiner (1990), p. 166-169] unless otherwise noted.

### III

Weiner has another, more challenging argument for questioning the interpretation of Frege as an unqualified realist. In the *Foundations of Arithmetic*, Frege asks at one point "how are the numbers given to us?" (Frege 1980, p. 73). On a realist reading of Frege, this seems like a natural question to ask. After all, if numbers are neither mental nor physical objects but are objects nevertheless, then how is it that we come to know anything about them? But Frege reformulates the question as "since it is only in the context of a proposition that words have any meaning, our problem becomes this: To define the sense of a proposition in which a number word occurs." On a realist reading this does not seem to make any sense. Not only does the question not make any sense, but the fact that Frege substitutes one for the other makes even less sense. Weiner argues that "if the considerations that give rise to Frege's question concern only justification [of arithmetical propositions], his answer does not seem odd at all." (Weiner 1990, p. 186)

How is that? Well, if the question asks which logical laws are required to justify sentences in which number-terms appear, it seems reasonable to ask for the meaning of those sentences. I'm actually not entirely clear on how this can be an argument for not taking Frege to be a realist since we are still left asking for the meaning of a sentence which, for Frege is a thought; an abstract object. Nevertheless, Weiner seems to have a point in claiming that Frege's substitution of a question about epistemic access to numbers with one about the meaning of number-terms in the context of a proposition makes little sense on Burge's reading of Frege as an unqualified realist.

Also, consider Frege's admonition in the *Foundations to*

...never to ask for the meaning of a word in isolation, but only in the context of a proposition.... if [this] principle is not observed, one is almost forced to take as the meanings of words mental pictures or acts of the individual mind. [Frege 1980, p. x]

It seems that what Frege is guarding against is the method of defining numbers wherein one takes an isolated number-term and metaphorically looks around for something to attach it to. So the formalist, empiricist and psychological mathematicians that Frege inveighs against are led into the mistake of offering up the numeral itself, collections of physical objects or mental images as the referents of number-terms by ignoring this admonition (hereafter: context principle). Then what about the realist mathematician? Does he not define number-terms out of the context of a sentence by offering up abstract objects as the referents?

So we are faced with two challenges to taking Frege as a realist. (1) Why would he take a question about epistemic access to numbers to be equivalent to a question about the meaning of sentences in which number-terms occur? And (2) how can Frege be a realist (taking the meanings of number-terms to refer to be abstract objects) and avoid violating the context principle? I think its pretty clear that both these problems concern the meaning relationship between words and sentences. So before I answer them I should discuss Frege's view of that relationship.

I'll begin with the context principle. The first thing to determine is whether the context principle is about sense or reference. Given his subsequent distinction between sense and reference, it is unclear what Frege means by "the meaning of a word." Are we to not ask for the sense or the referent of a word in isolation? Well, since sense determines reference, the principle is clearly about reference. The only question to ask is whether it is also about sense. Of course we need to keep open the possibility that the sense-reference distinction was intended to replace the context principle. After all, they were both at least partly concerned with the meaning relationship between words and sentences. I will tentatively, however, take the principle as one about reference.

The next thing to determine is the motivation behind the context principle. What was Frege concerned with when he included it as one of his fundamental principles? I suggested above that he was concerned with avoiding formalist, empiricist and psychological accounts of the meanings of number-terms. But surely this was only the symptom, not the cause. It seems that the cause was misunderstanding the relationship between the meanings of words with meanings of sentences. The formalists, empiricist and psychological mathematicians took the meaning of the word to be primary. That is, words mean what they do as a result of some non-linguistic fact or event (convention, stipulation, etc.). Sentences, on the other hand, mean what they do because of the words they contain. I think the context principle is clear enough that we can safely rule out any interpretation of Frege that commits him to the view that sentences are entirely dependent on words for their reference. But I also do not think we should go too far the other way and claim that Frege viewed words as entirely dependent on sentences for their reference. Consider the following passage from the *Basic Laws*:

The name of a first-level function of one argument has a *referent* (*refers* to something, succeeds in *referring*) if the name that results from filling its argument place by a referring object name always has a referent. An object name has a referent if the name which results from filling with it

the argument place of the referring name of a first-level function of one argument always has a referent. (Frege 1964, p. 84)

This is a rather dense passage, but its import is that since a sentence is a name of an object, the reference-dependence relation between words and sentences is symmetric. As the context principle suggests, words depend on their sentential context for their reference. But additionally, sentences depend on their component words for their reference.

As a side-note, the sense-dependence relation is also symmetric. Frege's view in "On Sense and Reference" that sense is compositional, and lines such as "It is enough if the proposition as a whole has a sense; it is this that confers on its parts also their content," (Frege 1980, p. 71) suggest that the relation also goes in the other direction.

So what have we determined? We have established that words refer to what they do because of the truth-value of the sentences in which they occur and that sentences are true or false because of the referents of their component parts.<sup>8</sup> The same relationship holds for sense. But if the meaning of a sentence confers meaning on the words in it, and the meanings of words confer meaning to the sentences in which they occur, then where does meaning come from? More specifically, how do number-terms and/or sentences about numbers come to be about numbers? How are the numbers *given* to us?

I am now ready to answer the first challenge to taking Frege to be a realist, but I should first say that a completely satisfactory response would require an adequate account of Frege's epistemology which I do not have. But I think we can establish enough to make it likely that Frege's comments do not jeopardize his realism. Frege is concerned with how we have access to numbers. At the time of the *Foundations*, he held that words have meaning only in the context of a sentence.<sup>9</sup> Given his logicism, it would seem that Frege's answer to the question "how are the numbers given to us" would be: "through reason." Also since Frege's characterization of ontological categories (concepts and objects) in terms of their representation in language suggests a linguistic-representationalist epistemology, I think it is likely that Frege thinks that our knowledge of numbers is somehow mediated by

<sup>8</sup> Recall that the reference of a sentence is a truth-value.

<sup>9</sup> The compositionality of sense and reference is nowhere to be found in the *Foundations of Arithmetic*. It isn't until "On Sense and Reference" that Frege suggests that sense and reference are compositional. In fact, in addition to the context principle, Frege makes several other statements that suggest he saw the meaning relation to be asymmetric.

language.<sup>10</sup> In light of this, I think it is quite clear why Frege substituted a question about epistemic access to numbers by one about meanings of propositions; because for him there was no difference. That is, if we have access to numbers solely through language and if number-terms have meaning only in the context of a sentence, then it seems reasonable, if we are interested in our access to numbers, to ask for the meaning of the sentences in which number-terms occur.

I can now also answer the second problem. How can Frege be a realist and avoid violating the context principle? Well, we first need to realize that Frege's realism would be a threat to his context principle only if the relationship between numbers and number-words is primary. That is, if in asking for the meaning of a number-term Frege offered up abstract objects directly, then he would violate his context principle. But this is not what he does, else why would he ask for the meaning of a proposition to determine the meaning of number-term? If the meaning of a number-term were primary, then the meaning of the proposition would not determine them. But then what does he do? He takes the meanings of sentences to be primary and determines the meanings of number-terms from them. This is probably what Frege means when he says:

In our present case, we have to define the sense of the proposition "the number which belongs to the concepts F is the same as that which belongs to the concept G"... In doing this, we shall be giving a general criterion for the identity of numbers. When we have thus acquired a means of arriving at a determinate number and of recognizing if again as the same, we can assign it a number word as its proper name. (Frege 1980, p. 73)

At this point I would like to recall what I said at the beginning of the paper concerning the order of explanation. I mentioned that an account of Frege's ontology had to explain how Frege arrives at the objects whose existence he is committed to. Burge simply claimed that he starts off with them. I think that it's fairly clear from my discussion that this isn't the case. Frege begins with a theory of meaning. Specifically, he begins with a view of the relationships between the meanings of words and sentences. Taking the meanings of sentences to be primary he arrives at the meanings of number-terms (i.e. the numbers themselves).

Burge could perhaps object by claiming that even if he granted that Frege starts out with the meanings of sentences, if meanings of

<sup>10</sup> What I mean is that, for Frege, since all knowledge is propositional and since we have access to propositions only through language, it is likely that Frege thought that all knowledge is somehow mediated by language.

assertoric sentences are themselves abstract objects then Frege still *starts out* with a realist ontology instead of arriving at it as a result of his theory of meaning. I don't have a definitive answer for this objection. But we should keep in mind why Frege makes the meanings of assertoric sentences into abstract objects. The reason he does it is to account for the possibility of truth and objectivity which are semantic and epistemic notions. Even if by the time Frege gets his numbers he is already committed to a realist ontology, this commitment is not a pre-philosophical assumption. He is led to it by his epistemology and semantics.

### Bibliography

- Burge, Tyler. (1997) "Frege on Knowing the Third Realm," in *Early Analytic Philosophy*, ed. W.W. Tait. Open Court Publishing. [originally published in *Mind*, v. 101, no. 404]
- Dummett, Michael. (1982) "Objectivity and Reality: Lotze and Frege," in *Frege and Other Philosophers*, Michael Dummett. Oxford University Press 1991. P. 97-125
- Frege, Gottlob. (1997) *The Frege Reader*, ed. Michael Beaney. Blackwell Publishers.
- \_\_\_\_\_. (1984) *Collected Papers on Mathematics, Logic, and Philosophy*, ed. Brian McGuinness. Basil Blackwell Publisher.
- \_\_\_\_\_. (1980) *The Foundations of Arithmetic*, trans. J.L. Austin. Basil Blackwell Publisher. [originally published as *Die Grundlagen der Arithmetik* in 1884]
- \_\_\_\_\_. (1964) *The Basic Laws of Arithmetic*, University of California Press. [originally published as *Grundgesetze der Arithmetik v.1* in 1893]
- Sluga, Hans. (1980) *Gottlob Frege*, Routledge and Kegan Paul Ltd.
- Weiner, Joan. (1990) *Frege in Perspective*, Cornell University Press

E. Sonny Elizondo  
Harvard University

The central theoretical innovation of Kant's ethical theory is the introduction of freedom as the basic concept of morality. According to Kant, the normativity of moral principles derives solely from the fact that an agent freely endorses them. While moral obligation is grounded in the free legislation of self-given laws, the moral agent cannot select his laws arbitrarily. The will itself is possessed of internal rules of operation – rules of rationality – which direct the will to accept only principles with universal form. This is just what the moral law, the categorical imperative, prescribes. The law of the free will, then, is the moral law. If freedom and morality are made equivalent, though, immorality appears unintelligible. Kant argues that we must be free, else morality is an illusion, but if we are free in Kant's sense, then we would only will morally. In order to make sense of humanity's engagement in immoral action, it is necessary to examine the embodied human will. Only by understanding the necessary constitution of the imperfectly rational human will in its active and regulative domains can we render actual human behavior compatible with our free wills and salvage Kant's ethics of freedom from misinterpretation and irrelevancy.

In *Groundwork* III Kant describes two dimensions of freedom: the negative and the positive. He defines the negative conception of freedom as the "property of a causality [the rational will] that can be efficient independently of alien causes determining it" (4:447). Defined in this way, freedom is juxtaposed with natural necessity in which all events can be described in terms of an infinite chain of causes and effects. In the natural, empirical world there can be no first cause, no initiator of causal schemes that was not itself effected by some antecedent cause. A negatively free will is self-determining. It is the beginning of a unique casual chain which it freely selects to effect in the world. This sense of freedom, however, does not indicate that the free will is a lawless, capricious will. The free will is not random, arbitrary or particularistic. It must will according to laws or principles. That is, it must will universally.

The will is a causality, and as such, must not be considered "lawless but must instead be a causality in accordance with immutable laws" (4:447). A free will, then, must act according to laws, but since it is free, these laws must be self-selected. It may then be asked how the

*Sonny is a junior philosophy major at Harvard University with in a primary interest in moral and political philosophy. He plans to attend graduate school in philosophy and hopes to pursue a career in academia.*

free will determines what principles it shall accept as its own. We know from *Groundwork* I that the will is just practical reason; it will act only if presented with a reason to do so. Since reasons are derived from principles, the free will must have a fundamental principle which it uses to determine what counts as a reason. But it appears a mystery as to how the free will selects its first principle since there can be no principle before the first by which the first can be determined.

The answer to this quandary is that since the will is a causality, the will is charged by its nature to select a law for itself. The content of that law is insignificant—the will just must select a law. This, though, is just what the moral law demands: "act only in accordance with that maxim through which you can at that same time will that it become a universal law" (4:421). The law of the free will, then, is just the moral law. This is the positive dimension of freedom. A free will, a completely self-determining will, behaves according to the moral law; a free will is a moral will, and morality is the active expression of the will's true freedom.

It should be obvious, though, that human beings engage in immorality with disturbing frequency and regularity. If humans are to be considered free – a necessary supposition if morality as defined by Kant is to be real for us – then we must understand how it is that a human will, a supposedly free will, can behave immorally. Kant argues in *Religion within the Limits of Reason Alone* that evil can be understood only in terms of freely chosen maxims. He says, "We call a man evil not because he performs actions that are evil (contrary to the law) but because these actions are of such a nature that we may infer from them that the presence of evil maxims" (16). In Kant's view, moral accountability derives only from the fact that the agent freely selects his own maxims. It is clear, then, that morality requires that humans have free will, but it is not clear how a free will can have an evil maxim if, by the previously given argument, the free will is by its very constitution the moral will. A satisfactory resolution of this problem must be offered if Kant's ethics are to be taken to comport with the actual moral condition of humanity.

While Kant does not appear to answer this problem explicitly in the *Groundwork*, he does indicate the general form an answer might take. After completing the analytic of *Groundwork* I and II, Kant takes up the question of whether or not his explication of the pure, rational will has implications for the moral lives of humans. He says that "if someone asked why the universal validity of our maxim as a law must be the limiting condition of our actions...we could not give him a satisfactory answer" (4:450). Of course, though, Kant does have an answer which it is part of the project of *Groundwork* III to give. The answer, broadly, is that the human presupposition of freedom provides

rational justification for a moral incentive that makes submission to the moral law rationally desirable for our imperfect human wills.

Before this incentive is discussed properly, though, it is necessary to understand what the mere existence of a moral incentive indicates about the human will. Morality requires an incentive only because of the imperfect rationality of human beings. The will of a perfectly rational agent could only act under the moral law. This is why Kant argues in the *Critique of Practical Reason* that an incentive should be “understood [as] the subjective determining ground of the will of a being whose reason does not by its nature necessarily conform with the objective law” (5:72). The issue of vital import here, though, is the constitution of the human will in all its imperfection. To what are moral incentives addressed? What actually is the human being we observe acting both in ourselves and in others?

In *Groundwork* III, Kant claims that the aspect of us motivated by incentives is a part of the phenomenal world, and, as such, is subject to the natural causality of desires and inclinations. From the perspective of pure practical reason, this aspect of ourselves cannot be understood as free. Freedom, if not a illusion, must exist behind the phenomena in some supposed noumenal reality which directs the dynamics of appearances. This proposal, though, seems incompletely articulated, for it cannot be doubted that we take ourselves to be. We feel ourselves acting freely in the world whether of not we act according to the law of the perfectly rational and free will, the moral law. This has led some thinkers, such as Sidgwick, to claim that Kant does not employ a consistent usage of the term ‘freedom’.<sup>1</sup> A further explanation of the our imperfectly irrational natures is required if these claims are to be refuted and our actual moral experience is to be rendered intelligible in a Kantian system.

Human beings are self-conscious creatures. We are aware of the operations of our own minds and conceive of a self to whom those operations belong. This self is, fundamentally, an agent. Self-consciousness is an act, the act of turning consciousness in on itself. The most primitive self-consciousness, then, is simply an awareness of one’s agency, a view of oneself acting upon a world independent to that self. This distinction between subject and object is not purely external. The subject or self is also taken to be self-directed in its actions. We have many natural desires such as hunger, thirst, and sex. We do not, however, feel as if these desires take immediate control over us, overriding our power to suppress the desire. Self-consciousness, then,

<sup>1</sup> Sidgwick’s Appendix to *The Methods of Ethics*, “The Kantian Conception of Free Will” p.511 Hackett Ed.

allows us to achieve a certain distance from our incentives. This apparent capacity to either endorse or reject an incentive gives rise to our first ideas of freedom.

Freedom understood in this way is a sort of pre-philosophical comprehension of Kant’s negative freedom. The self that determines action is seen as independent from the body that provides it with incentives to act. These incentives appear alien to the willing self, incapable of directly compelling the agent without his assent. In Kant’s language when we act upon an incentive, we incorporate that incentive into our maxims. This incorporation occurs only according to the free choice of the agent. The process of active deliberation is just the determination of what incentives are to be incorporated into maxims. We take ourselves to be agents – causes which effect our will on the world – only when self-consciousness provides us with the idea of freedom by distinguishing the acting self from a world of incentives external to it. The self that we take to be freely acting in ourselves and others will be called the active will.<sup>2</sup>

At this point Kant’s claim that we must act under the idea of freedom seems naturally evident. He says in *Groundwork* III that “the will of a [rational] being cannot be a will of its own except under the idea of freedom” (4:448). It has been shown how the concept of an active self, the active will, arose through the process of self-consciousness, and with it, the idea of freedom. Kant, though, does not want to attribute freedom to this active will. Although the active will takes itself to be free in that it feels no compulsion to act except according to its own volition, Kant believes that freedom does not consist merely in liberty from obvious coercion. For Kant, the free will is a causality governed by its own laws. The law of the free will, ultimately the moral law, is just the law of rationality. A free will is a moral will is a rational will. Freedom is something that properly exists only for perfectly rational wills. Clearly, though, humans are not perfectly rational. The proof for this is just the fact that we do act immorally, something of which a perfectly rational will, a divine will, is incapable. This freedom that we experience as self-conscious beings seems then to be faulty or in some way incomplete.

Again, we return to the problem of how we can be both free and immoral. Kant clearly does not want to call the active will, which obviously is capable of immoral action, free. If the active will is to be

<sup>2</sup> The active will described here is very similar to Kant’s *Willkür* characterized in *Religion within the Limits of Reason Alone*. I am greatly indebted to John Silber’s “The Ethical Significance of Kant’s *Religion*,” included in the HarperRow edition of that book.

free and accountable at all, then it must somehow be possible to connect the incomplete freedom of the active will with the complete freedom of the perfectly rational will. The imperfect rationality of human beings must have some necessary relation to perfect rationality. As with the incomplete freedom of the active will, it is self-consciousness that furnishes the imperfect human will with its connection to complete freedom.

Self-consciousness is self-comprehension. It is self-conscious reflection that allows us to conceive of ourselves as free at all. Further reflection, though, indicates that our freedom must have a foundation deeper than simply the distance we can achieve from our incentives. True freedom from the determination of alien influence is not, as we may care to believe, a radical liberty to follow our whims and transgress any and all laws. Such particularistic willing stands in contradiction to the very idea of a self distinct from inclinations. If the active will simply obeyed its whims, then it would just be its whims. The distance between the active will and its incentives would be reduced to nothing. The will, then, must be directed by some rules, but if it is to be free, these rules must be internal and not given from outside the will. The will does possess such rules: the rules of rationality.

The active will must, then, conceive of an aspect of itself that acts solely according to its self-given laws of rationality. Without this idea of a perfectly free self, the active will cannot properly take its own actions to be free at all. The active will's freedom is actual to the extent that it stands in relation to the posited free self. We will call this self the regulative will.<sup>3</sup> This regulative will is identical to the pure rational will which Kant claims exemplifies both the negative and positive aspects of freedom. The regulative will is negatively free in that it is completely self-determining, and it is positively free in that its law is the moral law. The regulative will, then, gives the moral law to the active will. Morality simply prompts the active will to fulfill the conditions of its own freedom.

Returning to the discussion of moral incentives, we can now understand how the moral law influences the active will by way of moral incentives. The moral incentive is the form the regulative will's claim on the active will takes. In the *Critique of Practical Reason* Kant claims that "respect for the moral law is ... the sole and also the undoubted moral incentive" (5:79). In respecting the moral law we recognize it as the same law of our true, free self, the regulative will. This also holds for Kant's talk of incentives in *Groundwork* III. The idea of a noumenal existence we realize through morality gives us an

<sup>3</sup> The regulative will described here is very similar to Kant's *Wille*, also characterized in *Religion*.

incentive to act morally. But this only makes sense if membership in the intelligible world is already part of our nature. This, though, is exactly what the regulative will demonstrates. The moral incentive represents morality as an ennobling enterprise that actualizes our own true free and efficacious natures.

It may be thought that the active and regulative wills actually correspond to two different types of freedom, such as Sidgwick's moral freedom and rational freedom. This is not, however, the case. First of all, both of the active and regulative wills are just dimensions of the same unitary human will. The division of the will into these dimensions is just the form that consciousness of freedom takes for imperfect human wills. To imagine a human will without either dimension is simply incoherent. It makes no sense to speak of an active will without a regulative will, since the active will must conceive of the regulative will as a condition of its own existence. An active will without a regulative will could only be unreflective; and although it may take itself to be free, it is incapable of actualizing true freedom. Without the idea of the regulative will, the active will would not know how to be free. Conversely, the regulative will is perfectly rational, but if we were to be so we would no longer be human. Both conceptions of the human will are necessary, then, for the reflective human will to conceive of itself as free.

Furthermore, the freedom of the active will is to be considered merely a corrupt imitation of the regulative will. Freedom refers to a singular concept of which the regulative will is the perfect representation. The active will falls short of this freedom because of its imperfect rationality, but both are being measured, so to speak, by the same standard. One exemplifies true freedom while the other does not. It does not make sense to divide the two in a way that would mask the essential relation that Kant wants to emphasize. Morality as freedom makes sense only if imperfectly moral can be understood as imperfectly free. Dividing freedom into two distinct parts would, as Sidgwick claims, vitiate the entire system. There is no reason to do this, however, for the system is consistent and coherent in itself<sup>4</sup>.

Another way to consider the issue is in terms of rationality in general. In either theoretical or practical applications of reason, we

<sup>4</sup> The point can also be made using Kant's idea of autonomy, which he uses in two different but not altogether distinct senses. If an agent is autonomous in the weak sense, then he conforms to a principle simply because he conceives of it as a law (aristocratic codes of honor, for example); he need not know why he regards it as such. If an agent is autonomous in the stronger sense, then he conforms because he recognizes what makes a law, i.e. its universal form. These are not two distinct ideas of autonomy; the former is fully realized and expressed by the latter.

never attain perfect rationality—that is not our nature. But just because we are not perfectly rational does not mean that we do not possess rationality at all. Our rationality is limited by our natures, but the rationality expressed by our actual imperfect existence need not be considered of a different kind. We express our reason when behaving rationally and an instance of irrationality does not undermine our entire identity as rational creatures. Since the relation between freedom and rationality is clear from *Groundwork* III, this holds for freedom as well. In acting as if we were free, we express the freedom of our rational nature and make ourselves practically free.

It may not be clear at this point how the problem of immoral willing is solved by understanding the human will in its active and regulative dimensions. Obviously, if immoral willing is to occur it must be the product of the active will, for the regulative will is just the standard of perfect rationality and freedom against which the active will must measure itself. The active will takes itself as free on the condition of its being at the core of its nature no different from the regulative will. The agent's choice in any moral situation is to achieve freedom through morality or to deny that possibility and let himself be overcome by his sensuous nature, subordinating the moral maxim to the maxim of self-love.

Although the will that chooses evil must be the active will, there is an aspect of immoral action that is decidedly passive. In *Religion*, Kant argues that it is impossible for a human being to violate the moral law for the sake of violation alone; such an act would be diabolical and wickedness of that sort is not part of human nature. The evil done by man is termed 'perversity of the heart.' Kant claims that "such a heart ... arises from the frailty of human nature, the lack of sufficient strength to follow out the principles it has chosen for itself" (32). Understood in this way, immorality is a weakness of will. Because we must take our fundamental nature to be free, the active will receives the moral law from the regulative will. In this way, the moral law is the de facto law of the active will. When the active will acts immorally, though, it allows prudential concerns to override moral ones. As Kant says, "Genuine evil consists in this, that a man does not will to withstand ... inclinations when they tempt him to transgress" (51).

This passive character of immoral action is not meant to indicate that the active will does not adopt immoral maxims. When I behave immorally and make self-love my fundamental maxim, I determine myself to act upon my strongest inclination. I have chosen my maxim freely, but if the mere presence of an inclination can move me to act regardless of my assessment of its rationality (morality), then my freedom is clearly restricted. I express my freedom in the first act of self-determination, the choice of a fundamental maxim, but do not

follow through on it. Instead of freely effecting my ends in the world, I allow myself to become a tool of inclinations over which I have no control. Immorality, then, is a sort of failed self-determination, purely unintelligible from the perspective of the regulative will, though all too real from our own active perspectives.

In this way we can understand immorality as a free acquiescence to the incentives of self-love. Kant's statement "only freedom in relation to the internal legislation of reason is properly a capacity; the possibility of deviating from it is an incapacity" (*Metaphysics of Morals* 226-27) makes perfect sense in this context. Deviation from the moral law is an incapacity to actualize one's true identity and obey one's self-given laws. Immorality is a denial of the fundamental freedom of our nature and an abdication of our place in the intelligible world.

It may be asked, though, whether the version of immorality described can truly be considered willing at all. Kant is clear that evil actions must be attributed to evil wills if morality is real. He says that "man himself must make or have made himself into whatever, in a moral sense, he is or is to become. Either condition must be an effect of his will; for otherwise he could not be held responsible for it and could therefore be neither morally good nor evil" (*Religion* 40). Immorality must, then, be a product of immoral willing, the active will's self-conscious denial of its regulative dimension.

On the other hand, while we must be able to infer an immoral maxim behind an immoral act, immorality is not the product of the pure will. True willing, the active expression of freedom and efficacy, can only be realized by the perfectly rational will. Our imperfect wills do not will in the same way that the pure will does. Immoral willing is possible in that the active will freely subordinates morality to self-love, but the manner of assent the active will offers is perfunctory. The endorsement our will gives to immoral maxims is real, but weak, a sort of rubber-stamp on incentives that the will is not powerful enough to resist. Willing, as the effective activity of rational causality, is not realized in immorality. Immoral willing is the will's passively granting its desires dominion over itself, rather than its actively expressing its own nature and laws. Immoral willing is willing but only incompletely and imperfectly.

It is clear, then, that immoral willing is not incompatible with the absolute spontaneity of the embodied human will. Self-conscious reflection gives rise to two understandings we have of our own will, the active and regulative dimensions. The imperfect freedom of the active will, expressed in immorality, arises from our own imperfect rationality. Though the regulative will informs the active will of its own true nature in freedom and morality, the active will is free to accept or deny its nature, to accept or deny its freedom. Such a denial, the essence of

immorality, stems from the active will's inability to resist the incentives of its sensuous nature. To be sure, though, the incentives of self-love do not compel the active will against its consent. The active will can always resist the temptation; resistance may be difficult, but no one ever claimed that morality was easy.

## Lusting After the Absolute

Justin C. Hall  
University of Texas

*The fact is that every writer creates his own precursors. His work modifies our conception of the past, as it will modify the future.*  
Jorge Luis Borges

To some extent the quest of modernity has been the search for the Absolute – modernity's path to certainty. A metaphysical certainty is the necessary foundation for truth and certainty in the natural sciences. This was the conviction of modern philosophers before Kant. However, three of the most influential modern philosophers, including Kant, criticize philosophy's attempt to establish empirical certainty on knowledge of the Absolute. Instead, they argue that absolute, unmediated knowledge is impossible to establish. Among these three thinkers, I find compelling yet distinct arguments that challenge not only modernity, but ultimately, the traditional Western notions of truth and value. Kant is the culmination of human reason's attempt to explain the world in terms of itself; Hegel is the inheritor of Kant's enlightenment ideals, while Kierkegaard cultivates the seeds of their deconstruction. After passing through Hegel, Kierkegaard radically restates a strengthened version of Kant's critique: objective knowledge is always only approximate. This is to say that thought can never fully represent the actual object to which it is supposed to correspond. Derrida takes this strain of arguments to a radical conclusion: philosophy's search for the absolute is an impossible and even nonsensical search. Language and philosophy deconstruct themselves, and thereby make meaning and truth always already indeterminable.

If there ever was a philosopher who lusted after the absolute it was Hegel, and yet Hegel's philosophy is in large part a reply to Kant's *Critique of Pure Reason*. Hegel is responding to Kant's claim that the Absolute is unknowable. According to Kant there are two worlds: the world as it actually is and the world as it is represented to the human mind. With this dualism as his foundation, Kant redefines the traditional notion of truth as correspondence between thought and object. Pre-Kantian modern philosophers held that the mind's representations, or thoughts, are grounded in the presence of the world as it actually is. In other words, mediated knowledge is somehow ultimately founded on immediate knowledge. Thus, the traditional, albeit not uncontested, view of truth is that a statement or thought is

*Justin Hall is an Honors Philosophy senior at the University of Texas at Austin. His current interests are metaphysics and the relationship between philosophy and literature. At all cost, he hopes to avoid an academic career in philosophy.*

true if it correctly corresponds to the actual, real object or circumstance that it purports to represent. Kant calls this actual thing the object in itself or the object independent of experience, and Hegel later names it the Absolute. It is worth noting that the Absolute has different connotations in Kant and Hegel. However, the Absolute holds roughly the same structural or functional role in both Hegel's and Kant's systems. According to Kant and in contrast to Hegel, this Absolute, this thing in itself, is unknowable to the human mind.

In many ways Kant's most enduring and important contribution to metaphysics is his Copernican revolution. Before Kant, the assumption was that the mind conformed to the objects of the world. Kant's new paradigm reverses the directionality of the relationship between the mind and the object. Objects, to be experienced by the mind, must conform to the mind. For Kant, transcendental idealism means that the scope of human knowledge is limited to objects of experience while objects independent of experience are by definition unknowable to the human mind. Transcendental idealism becomes the necessary condition for empirical realism.

Since the formal ground of the objects of experience is located a priori in the mind itself (Kant, p. 126), it can only apply to objects of experience, not to objects themselves. Raw sensory data, which Kant calls the manifold of intuition, are interpreted by the formal intuitions of space and time, transcendental unity of apperception and the transcendental synthesis of the imagination in the context of the "a priori conditions for the possibility of experience" (ibid., p. 126); Kant names these *a priori* conditions the categories. For Kant the categories, the transcendental unity of apperception, and the transcendental synthesis of the imagination are the structures of consciousness that govern the mind's interpretation of the world. The world of human experience is a world of representations of objects, and consequently, human experience necessarily does not deal with the actual, unmediated object. Certainty is not gained from knowledge of the Absolute (which is unattainable); rather, we can know that certainty exists in our representations because we ourselves introduce this certainty with the structures of our consciousness. In Kant's words, "the order and regularity in the appearances, which we entitle nature, we ourselves introduce" (ibid., p. 147).

Kant recognizes the creative role of the subject in the creation and maintenance of reality (the world of experience). Kant says "that experience itself should only be possible by means of this transcendental function of the imagination is indeed strange, but is nonetheless an obvious consequence of the preceding argument" (ibid., p. 146). Here Kant is admitting that if reality is confined to interpretations, the mind must and does have a creative role in at least partially determining

these interpretations. In this way Kant places humanity at the center of the universe more radically than any previous modern philosopher. Reality's creation and maintenance is dependent upon the human subject, and yet there are definite, if not determinable, rules and parameters on how and to what extent the mind creatively affects objects of experience. A representation necessarily represents something; individuals do not create their experience *ex nihilo*. Instead, the world of experience is an integration of the thinking subject's interpretation and of some unknowable given. The given is unknowable only insofar as by knowing it one refers to an absolute knowledge independent of experience. Kant's only evidence that there even exists an object independent of experience is this accepted definition of representation, and the intuitive idea that the mind does not create experience from nothing.

Experience of one kind or another is the only way humans interact and understand the world, and yet anything mediated by experience can only be understood within certain constraints. The human mind can never make contact with the Absolute because experience as we know it is mediation. Some of the most obvious *a priori* limitations on knowledge mediated by experience are time and space. Nothing can be known in any absolute sense if it is known in time and space. It can only be known as it was at time X, at location Z. Kant's claim is that to be human is to experience the impossibility of knowing things absolutely; this is also to say that human experience is guaranteed the possibility of knowing objects of experience with certainty. Therefore, the desire for the Absolute is a misdirected but inevitable desire. This is Kant's almost irresistible temptation. The representation indicates the existence of a represented, but the represented is necessarily outside the scope of human knowledge.

A key component of Kant's argument is the recognition of the limits of reason; one must realize that the scope of reason is confined to appearances. Hegel has a problem with this, claiming that one transcends limits by recognizing them. The idea is that to realize the limitations of a system, a faculty outside the system must be used. In other words, to posit that one does not know anything about an object is to claim to know something about that object (Hegel, p. 133-136). Hegel does not believe, as he contends traditional Western metaphysics believes, that knowledge is a means to an end. Knowledge is not "the instrument by which to get possession of absolute Reality" or anything else (ibid., p. 131). This notion of knowledge ignores the fact that knowledge of the object is always altered by the process of knowing. Consciousness comes to know something other than itself by explaining it in terms of itself. Consequently, both the object and the subject, the perceived and the perceiver are altered in the process of knowing. As a

result, human consciousness' understanding of the subject and object is always developing. Hegel holds that this unfolding and development of consciousness yields ever more clear and adequate understanding of reality, and thus, consciousness comes ever closer to the Absolute.

"Everything depends on grasping and expressing the ultimate truth not as Substance but as Subject as well" (ibid., p. Hegel, 80). Knowledge is the expression of the unity of subject and object. This is how Hegel attempts to transcend Kant's transcendental idealism – what Kant calls appearances, Hegel exalts as the Absolute. If, as Kant contends, human experience is composed solely of representations, then the sum of all representations is the Absolute. According to Hegel, Kant's Absolute, the world of objects independent of experience, does not exist. There is only the Absolute constantly becoming more aware of itself. Like Kant, Hegel is replacing the old notion of truth as correspondence with a new interpretation of truth as correspondence. The further consciousness develops, the better it corresponds with reality; Absolute knowledge is attained when consciousness corresponds perfectly with reality.

Kierkegaard's critique of Hegel shares similarities with Kant's critique of traditional rationalism: although it is an interesting and brilliant thought-experiment, it simply does not accurately describe the world. Existence cannot be demonstrated. Existence is a point of departure, and the only way a proof can demonstrate existence is by presupposing it (*Philosophical Fragments*, p. 51). Suppose an attempt were made to prove Napoleon's existence from Napoleon's deeds. While the existence of Napoleon does explain his deeds, "the deeds do not prove *his* existence, unless I have already understood the word 'his' so as thereby to have assumed his existence" (ibid., p. 50). The proof for Napoleon's existence is a delusion because it presupposes a relationship between him and his deeds (i.e., it assumes his existence). Thus, the proof is merely an elaboration of a presupposition.

This faulty reasoning is one of Hegel's crimes: Hegel attempts to demonstrate the existence of the Absolute (God or the Unknown) from its deeds. Hegel proposes to derive the Absolute from manifestations of the Absolute; which is to say he is developing an ideal interpretation that he had already presupposed. In *Concluding Unscientific Postscript* Kierkegaard asserts that "there can in all eternity be no direct transition from the historical to the eternal" (p. 89). Even if the Absolute exists, it cannot be reached through arguments based on contingency, yet Hegel must ground his absolute and universal knowledge on particular, historical facts. In Kantian terminology, Hegel tries to attain unmediated objects in themselves by way of mediated knowledge of representations. Therefore, Kierkegaard's opposition to "an attempt to create a quantitative transition to a

qualitative decision" (ibid., p. 88) reformulates Kant's warning against lusting after the Absolute.

For Kierkegaard, Hegel at best asserts that there is no difference between abstract thought and the abstract object. Hegel's formulation of truth "becomes a tautology. Thought and being mean one and the same thing, and the correspondence spoken of is merely an abstract self-identity" (ibid., p. 170). Kierkegaard replies that if this is what is meant by truth, then truth is merely a reduplication, a tautology. At worst, Hegel blurs the distinction between absolute and existential. He wants to argue the impossible; namely, that eternal truth can be derived from historical fact. Hegel forgets that his Absolute system is necessarily founded on existential subjectivity; the Hegelian system originates in an existing human author, and everything existential is uncertain. To choose any particular existential point of origin over any other is arbitrary. For Kierkegaard, existential uncertainty is irrefutably exemplified by the continual possibility of death for any existential subject; "all positive security is thus rendered suspect" (ibid., p. 76). An Absolute that is uncertain is no longer absolute. To put it another way, that which is essentially non-temporal and non-spatial cannot be made to exist in space and time.

From here Kierkegaard argues that the world is essentially fragmented for the existential subject. Hegel lusts after a merging of the Absolute and the existential that is for Kierkegaard impossible. Meaningful truth must originate in subjectivity because the "poor existing individual is confined to the strait-jacket of existence" (ibid., p. 172). Knowledge, on the other hand, is essentially inaccurate. Due to the uncertainty necessarily entailed in knowledge, only an approximation of correspondence is possible between the absolute object and the existential subject. The two are in states of constant change; to pin one down is to lose the other. Or, that which is essentially temporal and spatial cannot be understood in terms of the non-spatial and the non-temporal. Thus, truth is paradoxical.

"Thus the truth becomes an approximation whose beginning cannot be posited absolutely, precisely because the conclusion is lacking, the effect of which is retroactive" (ibid., p. 169). However, there is never any existential conclusion, and the existential striving after the truth is infinite. Truth is never attained, but only continually in the process of becoming. Which is to say, truth does not exist in a traditional sense. It is not a finite and explicable concept. In Kierkegaard's words "that the existing subjective thinker is constantly occupied in striving, does not mean that he has, in the finite sense, a goal toward which he strives," and "his thought must conform to the structure of existence" (ibid., p. 84, 74). If existential thought must always be striving, then so must the truth of the existential.

If Kierkegaard claims that, at best, object and thought can only approximate correspondence, Derrida contends that this entire project of truth is essentially self-deconstructing. The true, or correct, correspondence for which Western thought lusts is always already indeterminable. Derrida does not merely critique the Western metaphysical tradition, he also attempts to deconstruct the vessel carrying it: language. He changes the focus from the relationship between thought and object to the more fundamental idea of the sign.

One of Derrida's technical terms is '*differance*'. '*Differance*' is derived from the French verb that means both to differ and to defer; Derrida employs it with various meanings. First, it refers to the difference between the signifier and the signified, the representation and the represented. "The sign represents the present in its absence" (Derrida, "*Differance*," p. 9). The word 'tree' is not the same thing as the object tree any more than the visual image of the tree experienced by the human mind is the tree in itself. If the representation were the same as the represented, then it would not be a representation of but identical with the referent. Therefore, even in the presence of the signifier "the signified concept is never present in and of itself, in a sufficient presence that would refer only to itself" (ibid., p. 11). According to Derrida, there does not exist any transcendental signified – an original, non-sign signified concept.

A system of signs like language is nothing other than a system of differences. Nothing innately tree-like inhabits the sign 'tree'; rather, the only thing that defines the word 'tree' is its difference from other signs. The only way to define a sign is in terms of other signs. Every sign signifies another sign that is a sign for another sign, *ad infinitum*. Furthermore, every utterance of a particular sign is slightly different from the last; when the word 'tree' is uttered or written it is not signifying the actual tree. It is only signifying some other instance of the sign 'tree', and is consequently at least slightly different from this previous utterance. In this way, the effect of language is not a movement toward determinable meaning, but the continual and necessary distancing itself from it. The written sign breaks with the context under which it is written, or as Derrida says writing must "dissimulate its own history as it is produced" (*Of Grammatology*, p. 3). This is, among other things, Derrida's formulation of the idea that unmediated knowledge cannot be attained through mediation. On the contrary, further mediation can only push the immediate farther away. One can never say what one meant, but only what one thought one meant. Philosophical reflection does the opposite of what it intends. Western metaphysics is self-defeating precisely because mediation precludes immediacy. Derrida claims that *differance* "is the non-full,

non-simple, structured and differentiating origin of differences. Thus, the name 'origin' no longer suits it" ("*Differance*," p. 11).

These are some of the senses in which Derrida employs *differance*, and it is by virtue of these continual differences that meaning is always already deferred. In the absence of any non-sign origin, there is no authoritative reference for meaning. Authentic meaning can only be deferred infinitely. Unfortunately, Western philosophy has misunderstood the nature of language, and has exalted certain meanings at the expense of others. "Western thought...has always been structured in terms of dichotomies or polarities: good vs. evil, being vs. nothingness..., speech vs. writing" (Johnson, p. viii). This last pair of supposed opposites is a main theme in *Of Grammatology*. Speech is prized over writing; the idea being that there is less room for misunderstanding in speech because both the speaker and the listener are present. Writing, on the other hand, is a derivation of speech. Speech is the sign of some referent outside the system of signs, namely, thought. Writing is merely the sign of a sign, and therefore, of secondary significance. This is Western thought's marginalization of writing.

In light of Derrida's development of *differance* this exaltation of speech over writing is deconstructed. According to Derrida, every sign is second-rate. In several places Derrida has infamously made remarks such as "there has never been anything but writing" (*Of Grammatology* 159). He is not calling for the marginalization of speech. Derrida's idea is that there has never been anything in existence except textuality – the endless play of signs on signs, the play of the *differance* between sign (signified and signifier) and some other sign acting as a referent. Derrida calls this textuality *arche-writing*, and it is *arche-writing* that he contends is prior to speech. Thus, when referring to the opposition between literal and figurative meaning, Derrida says that it is not "a matter of inverting the literal meaning and the figurative meaning but of determining the 'literal' meaning of writing as metaphoricity itself" (*Of Grammatology*, p. 15). Literalness is essentially contained in the idea of metaphor and vice versa. Neither speech nor writing has any authentic right to a place of privilege over the other. Instead, it is *arche-writing*, the system of differences, that is prior to both speech and writing. Notice that Derrida's deconstruction, in opposition to Hegel's system, is fragmented, undecidable, and unverifiable.

Derrida reverses the relationship between the presence of the actual object and representation. Representation is not grounded in presence, but rather, representation is the necessary condition of presence. "Thus one comes to posit presence – and specifically consciousness, the being beside itself of consciousness – no longer as the

absolutely central form of Being but as a determination and as an effect" ("Differance", p. 16). Derrida's critique goes beyond either Kant's Copernican Revolution or Kierkegaard's strengthened reformulation of it. Kant was mistaken in assuming both that representation indicates a transcendental signified and that the conditions for the possibility of experience are necessary. Meaning is not grounded in the *a priori* conditions for the possibility of experience but in a system of differences. By grounding presence in *differance*, Derrida makes the concept of representation self referential insofar as it lacks any non-sign referent outside the system of representation. Therefore, representation deconstructs itself. This presence, this Absolute after which Western metaphysics has lusted, is nothing more than a myth.

Some of Derrida's detractors maintain that what results from his deconstruction is that everything is meaningless; anything can mean anything. This is to (mis)interpret the subtlety of his arguments. The irony is that if Derrida either believed or succeeded in destroying all meaning, none of his readers would know it. Of all Derrida might be guilty of, he has not proved, nor has he claimed to have proved that there is no meaning; it is simply not true that any meaning may be derived from any text. Instead, he has demonstrated the possibility of multiple meanings, and explored why none of these meanings may authoritatively marginalize any of the others. To restate this in the context of Kierkegaard and Kant, human experience is necessarily mediated, and a result of mediated, existential experience is a melee of possible multiple meanings, interpretations, or representations, none of which may legitimately claim a metaphysically privileged position over any of the others.

Another criticism of Derrida is that deconstruction deconstructs itself. Derrida states it himself: "the enterprise of deconstruction always in a certain way falls prey to its own work" (*Of Grammatology*, p. 24). This does not necessarily undermine the project of deconstruction; indeed it could *almost* definitively support it. Western metaphysics, not Derrida, desires to escape being situated within the context of language. Derrida calls this the double effacement: "Metaphysics – the white mythology which reassembles and reflects the culture of the West: the white man takes his own mythology, Indo-European mythology, his own *logos*, that is, the *mythos* of his idiom, for the universal form that he must still wish to call Reason" ("White Mythology," p. 213). Deconstruction does not merely raise the possibility of situating itself within the philosophical traditions of the West; it demands it. The fact that deconstruction deconstructs itself supports Derrida's claim that there is nothing outside the system of signs. The Western metaphysics of presence yearns after

the transcendental signified; the signified outside the system of signs. In this way, philosophy lusts after the impossible.

Instead of a transcendental signified, there is only the *differance*. Presence is always the trace of the past and the trace of the future; the passing of one into the other. A sign breaking with the context under which it is presented. Thereby, representing a previous presence, a previous here and now. Derrida is a re-presentation of his predecessors, just as Hegel is a sign of Kant and Kierkegaard a sign of Hegel, and on *ad infinitum*. The trace of one shifting into the trace of another. Most importantly, what Derrida is offering us is a new strategy for reading and understanding. The history of philosophy is the history of philosophers informing their readers on what previous philosophers actually meant, and where they were mis-taken. This is how the future is modified: by understanding the past in a new and different way, and not by attaining the Absolute. Derrida has simply shaken the foundations more dramatically, challenged the preceding paradigm more radically. He has replaced the traditional metaphysics of presence with a metaphysics of *differance*. Quite simply, his claim is that there is nothing other than the shifting of paradigms and the *differance*, their incessant and arbitrary differing and deferring amidst a melee of meaning. The result is that reality is not grounded in the presence of the actual object, but that presence is an effect of representation. However, this is paradoxical, and consequently, the Absolute, or metaphysical certainty, is unattainable.

#### Works Cited

- Bass, Alan. "Translator's Introduction to *Writing and Differance*." by Jacques Derrida. Chicago: U of Chicago P, 1978.
- Derrida, Jacques. "*Differance*." in *Margins of Philosophy*. Trans. Alan Bass. Chicago: U of Chicago P, 1982.
- \_\_\_\_\_. *Of Grammatology*. Trans. Gayatri Spivak. Baltimore: The Johns Hopkins UP, 1976.
- \_\_\_\_\_. "White Mythology: Metaphor in the Text of Philosophy" in *Margins of Philosophy*. Trans. Alan Bass. Chicago: U of Chicago P, 1982.
- Hegel, G. W. F. *The Phenomenology of Mind*. London: Allen & Unwin, 1931.
- Johnson, Barbara. "Translator's Introduction to *Dissemination*" by Jacques Derrida. Chicago: U of Chicago P, 1981.

Kant, Immanuel. *Critique of Pure Reason*. Trans. Norman Kemp Smith. New York: St. Martin's Press, 1965.

Kierkegaard, Soren. *Philosophical Fragments*. Princeton: Princeton UP, 1962.

\_\_\_\_\_. *Concluding Unscientific Postscript*. Princeton, Princeton UP, 1941.

## Falsehood in the *Republic*

Michael Keith Plocek

University of Texas

*to be false to one's soul about the things that are, to be ignorant and to have and hold falsehood there, is what everyone would least of all accept, for everyone hates a falsehood in that place most of all.*<sup>1</sup>

These words of Socrates, the mouthpiece of Plato, do not strike us initially as problematic. Plato's theory of knowledge, throughout his entire literary corpus, was an attempt to defend his conviction that realism is true, in order to refute any claims of Protagorean relativism. It does not strike us as odd that such a defender of truth would believe that a falsehood is hated not only by the gods, but also by human beings, since any believer in truth would want to grasp it (382c).<sup>2</sup> However, in outlining his *kallipolis* of the *Republic*, Plato advocates the use of falsehood, through the means of noble lies, to preserve the good of the city (389b-c). In light of the symbiotic dependence of Plato's interpretations of justice in the soul and justice in the city, it appears that Plato's *kallipolis* could never be just, since its denizens would have falsehood in their souls. It seems that this falsehood in the souls of the inhabitants, including the future rulers, of the *kallipolis* would disrupt the psychic harmony of division of labor of the *logistic*, *thymoeidic*, and *epithymetic* parts of their souls.<sup>3</sup> If reason is always straining to know where the truth lies (581b), then reason adulterated by falsehood would not be performing its appropriate task. Hence the souls of the denizens would be unjust.

The words that follow will elaborate upon and examine these prefatory remarks, in an attempt to test the internal consistency of Plato's overall argument for justice in the *Republic*. Understanding of the issues raised will be illuminated by several pertinent factors: the context of Plato's remarks on falsehood in the soul, how false Plato

---

<sup>1</sup> *Republic*, 382a, translated by G.M.A. Grube, revised by C.D.C. Reeve, Hackett Publishing Company, 1992.

<sup>2</sup> See also: *Meno*, 86b-c. ...we will be better men if we believe that one must search for the things one does not know, rather than if we believe that it is not possible to find out what we do not know and that we must not look for it. Translated by Grube, Hackett Publishing Company, 1981.

<sup>3</sup> Roughly translated, *kallipolis* is good-city, *logistic* is reason driven, *thymoeidic* is spirit driven, and *epithymetic* is desire driven.

*Michael Keith Plocek is a student of Philosophy and Ancient Greek. His extracurricular interests include playing guitar, learning foreign languages, and cavorting with his cat.*

actually believed his noble lies to be, the relationship between justice in the soul and justice in the city, his motivations for falsehood and Plato's awareness of the limitations of education in reality. This investigation shall proffer the notion that Plato's *kallipolis* is not as just as he contended, but this fact will not endanger his contention that the most just man is happier than the most unjust one.

### Falsehood in the Soul

In order to begin his defense of justice against the remarks of Thrascymachus, Plato examines justice in a city, rather than in the soul, since it is easier to learn about justice from the larger entity. Upon completion of the basic units of his *kallipolis*, Plato outlines the education of the citizens of it. His pedagogical remarks begin with the censorship of poetic passages that stated disparaging remarks about the gods. Since the gods are foremost good, they would never tell falsehoods; for falsehood in the soul is hated by all gods and humans (382a). In Plato's understanding of causality, only good can come from good, therefore the good gods would never taint the souls of men with falsehood.

Despite his conviction that falsehood in the soul is rightly hated by the gods, Plato states, a few words later, that lies can be useful for humans. Since we do not know the truth about those ancient events involving the gods, we can make a falsehood as much like the truth as possible, thus making it useful (382d). He later states that it is appropriate for the rulers of the city to use falsehoods for the good of the city; Plato develops his city upon a foundation of lies (389b-c).

At this point it is pertinent to examine exactly how much Plato actually thought that a falsehood could be like the truth; for if a falsehood cannot be like the truth, then the education of Plato's philosopher-kings would be an indoctrination of ignorance. In Book V of the *Republic*, Plato stated that knowledge is set over what is, while ignorance is set over what is not (477a). Opinion is the intermediate between the two. In the case of Plato's noble lies, since Plato himself stated that they are falsehoods, they must be at most opinions, because they are only like the truth, and are not set over what is. The natural phenomena described by each myth, e.g. innate abilities, learning, and death, cannot be adequately described by these myths, since they are falsehoods. So the myths are not adequate representations of each phenomenon as it is and as it is not. Strictly speaking, the myths are false. One could wonder whether or not Plato even considers them to be cases of true belief. Plato's myths would place falsehood in the souls of those they educated.

A Platonic apologist might reply to these preceding conclusions on several grounds. First, it could be argued that the whole purpose of

the education of the rulers of the city is to equip them with the proper tools for grasping the Good. Obviously this equipment will not be knowledge, since all knowledge of the other forms is grounded upon the Good. However, if the Good is ever to be grasped, then epistemological skepticism is not the way to go about searching for it. Rather, the dialectical skills taught to the rulers in their thirties would enable them reach the Good, and thereby proceed to deduce true knowledge from it.

These words may evoke sympathy for Plato's position, yet they do not disarm the contention that Plato's education of lies places falsehood in the souls of the rulers; rather, the aforesaid defense concedes this contention as a presupposition of the dialectic. Preliminarily, for we have yet to examine fully the relationship between the souls of the denizens and the city they inhabit, we must conclude that Plato's *kallipolis* would not be just, since the reasoning parts of the souls of its citizens would be directed towards what is not, thereby not performing reason's own task.

At this point in the argument, our hypothetical Platonist could reply that Plato's *kallipolis* is the most just that is possible in reality, and Plato was concerned with creating a possible city, rather than a merely theoretical one. The *kallipolis* would not be completely just, but it is the most just, and thus does serve as an adequate model for the purpose of Plato's arguments against Thrascymachus. However, the validity of this claim, that the *kallipolis* is the most just possible city, can be brought into question. Would not a city created by the philosopher-kings of the *kallipolis*, after they had grasped the Good, be more just than the *kallipolis*? If the *kallipolis* is possible, then the city created by the fruits of its labor is also conceivably possible. However, arguments of this kind are somewhat dogmatic, in the sense that they attack Plato's position in an unwarranted fashion, because increased support of them blinds one to the fact that the *kallipolis* may be just enough for Plato's purposes of contrasting the happiness of just people with the happiness of the unjust.

### The Soul and the City

In his quest for justice, Plato made frequent use of the connection between city and soul, at times merely drawing an analogy between the two, while at other times establishing a causal link. For the ends of this analysis, we will explore the latter connection, in an attempt to determine if the falsehood in the souls of the denizens, particularly the rulers, of the *kallipolis* would detriment the justice of the city. If the *logistic* parts of the philosopher-kings' souls are to be directed towards falsehood, rather than where the truth lies, then reason would not be able govern the spirit and appetites, and the soul would then be unjust. Consequently, if a necessary connection surfaces

between the justice of the souls in the city and the justice of the city itself, then Plato's belief that his *kallipolis* is just will be refuted, since the injustice in the souls of the denizens that results from falsehood in their souls would entail injustice in the *kallipolis*.

Once Plato had established justice in his *kallipolis*, through the money-making, auxiliary, and guardian classes each doing its own work, he returned from the larger example of the city to the soul (434c-e). A just man would not differ at all from a just city in respect to the form of justice (435b). In order to strengthen his analogy between the attributes of the city and the attributes of the soul, Plato explicitly stated his belief in a causal link:

Must we not acknowledge that in each of us there are the same principles and habits which there are in the State; and that from the individual they pass into the State?—how else can they come there? Take the quality of passion or spirit;—it would be ridiculous to imagine that this quality, when found in States, is not derived from the individuals who are supposed to possess it.<sup>4</sup>

These words unequivocally support the thesis that the *kallipolis* will not be just, since its attributes are to be dictated by the souls of the citizens.

However, if a city is just only if all its citizens are just, then each class of Plato's *kallipolis* will be *logistic*, *thymoeidic*, and *epithymetic* only if its respective members are. This presents a problem for Plato's *kallipolis*, since the city will be in a constant state of strife, for the lower classes will not be ruled by the *logistic* parts of their souls and hence will not understand the edicts of the ruling class. Totalitarian measures would have to be introduced.

Yet perhaps we have constructed a straw man out of Plato's argument. At times he seems to have stressed that it is actually the characters of the leaders of the city that are responsible for the constitution. The passage at 544d-e, where Plato begins his argument for the tripartite soul, states:

there must be as many types of human character as there are forms of government. Constitutions cannot come out of stocks and stones; they must result from the preponderance of certain characters which draw the rest of the community in their wake.<sup>5</sup>

<sup>4</sup> *Republic*, 435e, Jowett translation. Grube has translated this part ambiguously.

<sup>5</sup> *Republic*, 544d-e, translated by Francis Cornford, Oxford University Press, 1945.

This seems to be hinting at the fact that it is the characters of the rulers that are responsible for the constitution of the state. At other points in the work, Plato noted the importance of the rulers in determining the attributes of the *kallipolis*. The divergent forms of constitution are engendered by the ruling classes, and are defined according to the characters of such rulers. Hence if the rulers are just, then the city will be just.<sup>6</sup>

These aforesaid remarks help rebuild Plato's argument for justice in the *kallipolis*. Although the rulers will have been told falsehoods in their youths, they would eventually learn that these lies are false. This knowledge would be revealed to them at some point in their education, and if it is not, they should be able to find these myths to be false themselves, once they have mastered dialectic. Since the rulers will have a hand in all public policy, it follows that the city's constitution would most resemble their characters. Because they eventually would not possess falsehood in their souls, the city appears to be just. This was one of the reasons why Plato outlined the education of the rulers in such detail; all falsehood must be removed from the souls of the potential rulers. However, these remarks do not alleviate the problem that the rulers will not be just until they have been fully educated, so the city will not be just. Plato might have been willing to concede this problem, for he was well aware of the difficulties surrounding education, in general as well as in his *kallipolis*.

### Education and the *Meno* Problem

Plato was aware of the difficulties involved in educating the guardians towards the Good, since leading the rulers to the epistemological grounding of the Good presupposes an understanding of it. Since his pre-dialectic understanding of the Good caused him to posit its existence, and to speculate that it is what makes all other forms intelligible, Plato must have believed that some degree of belief is necessary to pave the way to knowledge. Once the Good has been grasped, then the philosopher-kings would be able to descend from it and deduce true knowledge. However, it has already been argued that these falsehoods would make the rulers, and thus the city, unjust, since the *logistic* parts of the souls of the rulers would be directed at falsehood. Yet Plato was aware of the similar difficulties involving any inquiry; this awareness was exemplified in one of his earliest dialogues, the *Meno*.

<sup>6</sup> See Bernard Williams' *The Analogy of the City and Soul in Plato's Republic* for a more detailed analysis of such difficulties. (present in *Exegesis and Argument*, Mourelatos and Rorty, ed. Van Gorcum, 1973, p. 196-206.)

In the search for an adequate formulation of the definition of virtue, Plato referred to a debater's argument, that

a man cannot search either for what he knows or for what he does not know. He cannot search for what he knows—since he knows it, there is no need to search—nor for what he does not know, for he does not know what to look for.<sup>7</sup>

Plato reconciled this hermeneutic circle by stating that all knowledge is recollection of the forms. Learning is the clearing up of the understanding of the forms present in us from birth. As far fetched as this reconciliation may appear to our modern metaphysical conceptions, it is internally consistent; it also defends Plato against certain critiques of circularity, since his philosopher-kings could be instructed with the pre-dialectical "knowledge" human beings already possess of the Good and the other forms. Plato stated that the role of education is to turn the mind's eye from particulars to the forms, not to put sight into blind eyes (518b-c).

We all can have access to the forms, if we are properly educated. Therefore, the education of the rulers is paramount; they, above all, must grasp the forms. In his allegory of the Cave, Plato stated that if someone was compelled to look at the light itself, her eyes would hurt, and she would turn and flee towards the things she is able to see, believing them to be clearer than the light (515d-e). Perhaps Plato felt falsehoods were necessary, as a means of preliminary education; the future rulers would be blinded if they looked directly from the shadows on the wall to the sun. Plato must have believed that falsehoods, if carefully formulated, could provide the impetus for further understanding. Our current pedagogical practitioners proffer falsehoods to children, in order to help them grasp fundamental concepts. Any student that enrolls in a decent college course is usually dumbfounded at the lack of certainty that applies to practically everything she had learned in her preliminary education. Children are lied to, being given simplified versions of science, in order to facilitate their understanding. Only after a huge system has been inculcated into their minds are they permitted to question the foundations of anything. The revelation that falsehood may be necessary for pedagogical purposes weakens the earlier claim that a more just city could be developed after the rulers have obtained access to the Good, for these knowledgeable rulers may still have to lie to their pupils, in order not to blind them with too much at once. Plato planned that the education of the rulers would require many decades for a reason: the path to the Good is long and this path

cannot be traversed overnight; the eyes must adjust themselves to the light slowly.

Plato's rulers would hold some degree of falsehood in their souls, but this is because they do not yet have a grasp upon the Good. However, Plato did not feel that this fact should impede the development of the *kallipolis*. Plato did believe that there was some truth to his myths (382d). Even if the myths are not knowledge, in the strict sense of being grounded by the Good, they still useful when made like the truth (382d). Therefore the myths do serve a positive purpose in the *kallipolis*; they help govern the rulers, and thereby the city, towards the Good and the Just.

### Relevance of These Conclusions to the Overall Argument of the *Republic*

It has been demonstrated that the use of falsehoods, within the purview of Plato's thought, would be deleterious to the overall justice of the *kallipolis*. However, this demonstration does very little damage to Plato's main thesis that the truly just man would be happier than the truly unjust man would be. Plato's portrayal of the tyrannical soul, ruled by the most base of *epithymetic* desires, is one of a man that is far less happy than any of the other four types of souls described. Regardless of the fact that the aristocratic *kallipolis* has been demonstrated to be not completely just, it, and its denizens, can still be regarded as the most just of any of the five types. It cannot be cogently argued that the tyrannical man is happier than the aristocratic man; the soul ruled by reason has a much better chance of being satisfied, and thus happy, than the soul ruled by appetite, since the *epithymetic* desires are insatiable. Moreover, if the consequences of justice and injustice are to be a means of evaluating which of the two is to be preferred, then justice is the felicitous choice, since the unjust tyrant will be in constant fear of those around him, as he is at war with the city in a manner analogous to his internal strife. Although falsehoods in education would lead to falsehoods in the soul, and subsequently to falsehoods in Plato's city, the citizens of Plato's ideal state would be far happier than the subjugated wretches of a tyrant. Thus, the debasement of the justice in Plato's *kallipolis* is not detrimental to his overall thesis: the just man is happier than the unjust man.

<sup>7</sup> *Meno*, 80d-e, translated by G.M.A. Grube, Hackett Publishing Company, 1981.

# Politics, Texts, Residues: Foucault's Philosophy and the Possibility of Social Activism

Rafeeq Hasan  
University of Chicago

What I call text is also that which 'practically' inscribes and overflows the limits of such a discourse.

--Jacques Derrida<sup>1</sup>

Between techniques of knowledge and strategies of power, there is no exteriority...

--Michel Foucault<sup>2</sup>

In an interview conducted for the 1984 Channel 4 Voices Series, *Knowledge in Crisis*, philosopher Ron Aronson made the following comment on post-structuralist philosophy:

Post-structuralism...seems to make the whole world into a text...What troubles me is the seeming self-exclusion of the world, the absorption of everything in the world into a world of discourse, the world of words, the world of texts, and the lack of concern for those problems of human suffering...that occupy most of my waking energy.<sup>3</sup>

Aronson's view condenses and typifies the general critique of post-structuralist philosophy made by its detractors within the academy. More problematic than this somewhat crude definition and subsequent condemnation of a sophisticated and multi-faceted corpus of philosophical work, however, is the response to Aronson provided by Gayatri Spivak, a post-colonial literary critic who avowedly professes her allegiance to the post-structuralist philosophies of Jacques Derrida and Michel Foucault. Her reply to Aronson hinges on an implicit assumption regarding the post-structuralist view of the text that I wish to argue is not all pervasive within the discourse of post-structuralism itself. She claims that the Derridean understanding of 'text' is not limited to the verbal text. Rather it deals with

<sup>1</sup>Jacques Derrida, *Positions* (Chicago: University of Chicago Press, 1981) p. 59.

<sup>2</sup>Michel Foucault, *The History of Sexuality, Volume 1: An Introduction*, trans. Robert Hurley (New York: Pantheon, 1978) p. 98.

<sup>3</sup>These discussion are collected in Gayatri Spivak, *The Post-Colonial Critic* (New York and London: Routledge, 1990) p. 24.

Rafeeq Hasan is a sophomore philosophy major at the University of Chicago. He is currently exploring the relation between post-structuralism and second generation immigrant identity.

'a network, a weave...politico-psycho-sexual-socio, you name it.'<sup>4</sup> This rather cryptic explanation of the 'text' is used to conclude that a real event of 'human suffering,' such as the possibility of nuclear holocaust, could be understood as a textual event in that it would be the product of certain historically determined ideas and consciousnesses that would function as legitimators for this action. Spivak says that such a catastrophic and intuitively 'real' event could not happen 'without the history that we are speaking of at great length here.'<sup>5</sup>

Thus, Spivak claims that the role of the text in the post-structuralist view of social change is one of legitimization. The 'text,' whether it be discursive or spoken, legitimizes political action which creates 'real' change. This has the effect of subsuming the political into the textual. If political actions are always already determined by certain kinds of discourse, a discourse which creates avenues for certain kinds of thoughts or ideas, the human subject becomes only a minor player in overturning the conditions which oppress him. For many modern critics who declare their affinities with Foucault, thinkers like Spivak or Homi Bhabha, this notion of the textual and its relationship to the political leads to a methodology in which it is the sites of logical breakdown inside any particular discourse *itself* that eventually brings an end to the discourse and its peripheral systems of power. In the work of Homi Bhabha, for example, it is the tension between the desire to make the colonized subject in the image of the colonizer and the need to create a debased subject to be metaphorically enslaved that ends the colonial moment.<sup>6</sup> Thus, colonialism falls because of the fissures in colonial rhetoric, and human agency or struggle on the part of the colonized themselves seems to be of minimal importance.

Removing the specifically Derridean bent of Spivak's methodologies, I wish to suggest that this notion of political change as that which occurs from within the formal logic of the 'text' itself is not a view shared by all of post-structuralism, and that consequently, a totalizing critique of post-structuralism on the grounds of its lack of pragmatism or de-emphasis of human agency must be formulated differently.

<sup>4</sup>Spivak, p. 25.

<sup>5</sup>*Ibid.*, p. 25.

<sup>6</sup>See, for example, Homi Bhabha, 'Of Mimicry and Man: The Ambivalence of Colonial Discourse,' in *The Location of Culture* (New York and London: Routledge, 1994). It should be noted that my characterization of the writings of both Bhabha and Spivak is reductionist at best. Indeed, like Foucault's own distinctions between 'ars erotica' and 'scientia sexualis' in *The History of Sexuality, Volume 1*, the opposition that I have drawn between these thinkers and Foucault serves more as a convenient theoretical starting point than as a completely accurate methodological exposition.

Before I progress any further, I must adamantly assert that the label 'post-structuralism' is profoundly limiting in describing the diverse and often opposing works of those who fall under this moniker. Indeed, it seems that 'post-structuralism' refers more to the kind of Parisian modishness and inaccessibility of language in the writings of several French philosophers rather than to any one underlying or unifying epistemological principle.<sup>7</sup> Keeping this in mind, I wish to use the work of Michel Foucault, arguably the best known and least modish of the post-structuralists, as well as the writings of some of his philosophical contemporaries, to suggest that a more complicated view on the relation between the textual and the political exists in at least one branch of post-structuralism.

Specifically, I want to show that Foucault and his contemporaries do not consider a text 'political' because it legitimates a political action, or because the political is subsumed into the textual. Rather, I claim that in the Foucaultian methodology, a text is the site where the political and the discursive become closely imbricated--the text shows where political ideas have sunk the most solidly into the status of knowledge. By 'overflow[ing] the limits' of practical discourse, the text illuminates where political power is concentrated and where it is residual. Foucault attends to texts because it is in the text that a philosopher can point to the breakdown between action and the systems of truth which legitimates action, i.e. the gaps in the edifice of power. But, unlike Spivak, the impetus behind illuminating these 'textual' gaps is to illustrate situations into which a human subject can insert himself to create new relations of power--to create change. To return to the Bhabha example, I am suggesting that Foucault would locate the tension between mimicry and mockery, the need for sameness and difference, not in order to show how this tension internally creates colonialism's downfall, but to show one specific way (by mimicking yet mocking) that a colonial subject would be able to effect a particular oppressive aspect of colonialism. To return to the Spivak example of nuclear threat, I quote a minor passage from an essay by Paul Veyne: 'Foucault....hardly believed that the emergence of nuclear terror could be traced back to an unfortunate proposition by Descartes.'<sup>8</sup>

Most importantly, I want to explore how this illumination of the fissures of power, and its relation to the human subject, coincides with

<sup>7</sup>See, for example, Edward Said, 'Michel Foucault 1926-1984,' in *After Foucault: Humanistic Knowledge, Postmodern Challenges*, ed. Jonathan Arac (New Brunswick and London: Rutgers University Press, 1988). p. 2.

<sup>8</sup>Paul Veyne, 'The Final Foucault and His Ethics,' collected in *Foucault and his Interlocutors*, ed. Arnold Davidson (Chicago and London: University of Chicago Press, 1997). p. 227.

Foucault's own view on what political activism should entail. Using Arnold Davidson's Foucaultian approach to reading Freud, along with passages in Foucault's *History of Sexuality Volume 1: An Introduction*, I wish to suggest that this imbrication of the textual and the political actually allows for an analysis of power that, in turn, creates the conditions of possibility for pragmatic social change.

Running through the thread of Arnold Davidson's argument in 'How to Do the History of Psychoanalysis: A Reading of Freud's *Three Essays on the Theory of Sexuality*,' is a view on the relationship between the text and the 'outside' world that provides a clear introduction to my claims about the real nature of the Foucaultian view of the text. In Davidson's effort to assess Freud not from the standpoint of Freud himself but in terms of 'anonymous rules for the formation and production of statements,'<sup>9</sup> what in this case he calls a 'psychiatric style of reasoning,'<sup>10</sup> Foucault's philosophical affirmation of the death of the author can be everywhere felt. Yet while Davidson claims to align himself more with the early Foucault's project of formulating an archaeology of knowledge internal to disciplines that claim that status of objective truth, Davidson's argument ultimately has political implications which slide into the framework of the later Foucault as a genealogist of power--a Foucault found most prominently in *The History of Sexuality Volume 1*.

In this article Davidson analyzes Freud's thought on perversion and the sexual instinct in light of the earlier psychiatric theories of sexuality provided by authors such as Kretz-Ebbing. These psychiatric (rather than psychoanalytic) theorists moved away from the anatomical explanations of sexual perversion and created the concept of sexual instinct--a mental site for sexual behavior. In order to sustain themselves as members within a diagnostic tradition, in order to constitute sexuality as a disease, this creation of the sexual instinct led to a need to create distinctions of 'normal' and 'pathological' within the instinct itself. Perversion could only be a disease in need of a cure if its root, the sexual instinct, could be classified in certain manifestations as deviant; a deviance which could then be effaced to make the instinct 'normal.' This explanation for perversion as lying at the heart of a historically specific and seemingly minor struggle for legitimating power certainly evokes Foucault's project to 'identify the accidents, the minute deviations...that gave birth to those things that continue to exist and have value for us,'

<sup>9</sup>Arnold Davidson, 'How to Do the History of Psychoanalysis,' *Critical Inquiry* Winter 1987, p. 254.

<sup>10</sup>Davidson, p. 255.

what he beautifully characterized as the primacy of 'the exteriority of [the] accident.'<sup>11</sup>

Yet if all Davidson's article did was explain perversion, with all of its very real political and legal realities, in terms of the discursive shifts and textual strategies of psychoanalysis and psychiatry then perhaps Aronson's claim of a post-structuralist elision of the political would be appropriate. But as Davidson moves on to his discussion of Freud the article goes beyond this textual discursive realm.

Davidson attributes to Freud the notions of a sexual drive existing prior to a sexual object, as well as a 'normal' (in this case, normative) sexuality arising from component sexual instincts that would be separately regarded as perverse. Davidson concludes that these novel notions radically reassessed the theories of nineteenth century sexuality by reformulating the logical rules of psychiatric discourse to the point where 'cases of inversion [could] no longer be considered pathologically abnormal.'<sup>12</sup> Yet, while the Freud of his own texts (the discursive Freud) could then no longer logically categorize certain sexualities as normal or pathological, the very real Freud continued to do so. Davidson illustrates this tension between the discursive Freud and the real Freud by examining the passage where Freud writes, 'For if the genital zone is weak, this combination, which is required to take place at puberty, is bound to fail, and the strongest of the other components...will continue its activity as *perversion*' (emphasis mine). Davidson reads this passage and concludes that from the logic of the text that has preceded this statement, the last clause should simply read 'the strongest of the other components...will continue its activity.'<sup>13</sup>

To explain this discrepancy in what the logical Freud should have meant and what Freud actually said we must necessarily turn away from the textual, or rather, look at the textual in a political light. Davidson locates this textual/'real' contradiction in what French historians have called a *mentalite*, or, as he defines it, 'a set of mental habits or automatisms that characterize the collective understanding and representations of a population.'<sup>14</sup>

This *mentalite* takes into account those elements of understanding and action that are largely unwritten, as they have submerged themselves into the psychic realm of 'collective understanding.' Just as Freud's texts

<sup>11</sup>Michel Foucault, 'Nietzsche, Genealogy, History' collected in *Language, Counter-Memory, Practice: Selected Essays and Interviews* reprinted in Paul Rabinow, *The Foucault Reader*, (New York: Pantheon Books, 1984), p. 81.

<sup>12</sup>Davidson, p. 265.

<sup>13</sup>This encapsulation contained in Teresa de Lauretis, 'Freud, Sexuality, Perversion' p. 229-229.

<sup>14</sup>Davidson, p. 275.

show how the political and the textual relate, how residues of the political do not allow for a complete objective pursuit of new knowledge, Davidson's archaeology of psychiatric knowledge leads one to question the claims of an elitist, textual based post-structuralist philosophy of which he is, at least in some vague sense, a follower. What Davidson's reading of Freud has shown is not that the subjugation of homosexuals or others labeled perverts is a result of the writings of Freud that legitimized these actions, nor that Freud's rhetoric has the necessary consequence of subjugating homosexuals or perverts, methods that Aronson or Spivak (respectively) would attribute to post-structuralism. Davidson has instead shown that a seemingly non-political text can be read to illustrate what aspects of the political sphere had saturated themselves into Freud's psyche, implanting themselves as experiential or epistemological a-prioris for Freud's subsequent thought. We can use both Freud's and Davidson's texts to see just what aspects of the oppressive power against deviants were so pervasive as to take root in 'objective' science.

To generalize away from the specific context of Davidson's article, this approach locates for the activist where power is so invidious as to produce perpetuating effects by pointing to the places where power incorporates itself into the very thought that attempts to rid itself of its effects. The ambivalences that the article points to within the rhetoric does not provide a kind of predictive schema for change (*pace* Bhabha), but instead locates possible points of resistance for those rendered inferior by the rhetoric. To again speak as concretely as possible, an article such as this does not provide an account of the beginnings of homophobic rhetoric in order to predict what outcomes the lack of logic within this rhetoric will eventually create, but instead points to lacks of logic as places where an oppressed homosexual could attempt struggle.

Yet Davidson is obviously not writing for a politically activist audience, and to claim that his text has overt political implications is to perhaps correct for Aronson's claim to the other extreme. Indeed, if we heed Foucault's advice and look to where '[a text] has been used, how can it circulate, and who can appropriate it for himself,'<sup>15</sup> the publication of Davidson's article in a scholarly journal necessarily makes its political implications subtle rather than overt. I began with the academic Davidson rather than with the more popularly known Foucault to show more generally how Foucault's methods for inspecting power, while seemingly entirely discursively based, are, in the final analysis, directed toward extra-textual considerations. Given that we have seen how Foucault's methods

<sup>15</sup>Michel Foucault, 'What Is an Author,' collected in *Textual Strategies: Perspectives in Post-Structuralist Criticism* reprinted in *The Foucault Reader*, p. 120.

have been used to read other authors and their historico-political positionings, in an at least implicitly political light, I now wish to show how Foucault's methods can be used to read Foucault himself. Specifically, I want to do a Davidson-type reading on *The History of Sexuality Volume 1*, albeit with a reverse impetus. Using Foucault's archaeology of knowledge to locate continuities rather than ruptures, I wish to see how the seeming ambiguity in Foucault's shifting distinctions between sex and sexuality both correspond to his genealogical views of power as well as to show how he may have implicitly encoded a view of the text's relationship to politics that functions in the imbricating rather than the legitimizing (or self-generating) way that I have sketched.

For the reader who wishes to see Foucault's *History of Sexuality Volume 1* as more than a simple polemic against the repressive hypothesis as authored and subsequently reformulated from the early psychiatry of Kretz-Ebbing up to Lacan, one of the more perplexing aspects of the text is the distinction, or lack thereof, that Foucault draws between 'sex' and 'sexuality.' Indeed Foucault appears to fluctuate between these words, using them simultaneously to refer to seemingly different concepts.

On one level the main task of the work is to show how sexual desires and preferences have been constituted, through a historical process beginning with religious confession and ending with psychoanalysis, as the ultimate truth of a person's very being rather than as simple acts of pleasure that a person may pursue. Yet Foucault certainly problematizes the notion that this is the work's purpose by often speaking of a construction of the truth of 'sex,' a term which seems to imply the sexual act rather than sexualized consciousness, instead of a truth of 'sexuality.' For instance, near the book's beginning Foucault writes, 'the essential thing...[is]...the existence in our era of a discourse in which sex, the revelation of truth, the overturning of global laws...are linked together.'<sup>16</sup> Yet elsewhere when Foucault examines how the creation of confession as a procedure for exacting the 'truth' of the subject relates to the primacy given to sexuality, he comments, seemingly more logically, that 'sexuality' rather than sex was made something to be interpreted.<sup>17</sup> Indeed the most obvious ambiguity seems to appear in the latter half of the work, where Foucault both claims that sex is constituted within the deployment of sexuality<sup>18</sup> and that 'We...are in a society of 'sex,' or rather a society 'with a sexuality.'<sup>19</sup>

These purely textual contradictions seem further problematized given the general thrust of Foucault's argument in the work.<sup>20</sup> If sexuality is constituted as all determinative, the main problem for Foucault seems to rest in the fact that sexual acts are no longer valued, as he felt they were in societies possessing an *ars erotica*, from the standpoint of their possibilities for the creation and intensification of pleasure. Foucault seems to object to the fact that we have created homosexuals (a group whose very identity lies in specific sexual practices) rather than sodomites (those who simply engaged in acts of non-standard sex). If, as he claims, sexual acts were reclaimed and given greater currency within the apparatus of sexuality, why would he preface his claim that we are a society 'with a sexuality,' with 'We...are in a society of sex'? And furthermore, why would he write 'It is through sex...that each individual has to pass in order to have access to his own intelligibility,'<sup>21</sup> given that his project of showing how sexuality is constituted as the secret has led to a need to know yourself not through the specifics of your sexual acts, but in the larger implications for your very being that these acts entail?

I wish to suggest that this is actually not an ambiguity, or in the context of the Davidson article, that this is not a kind of ambivalence on Foucault's part stemming from some French homosexual *mentalite*. Rather I wish to point to the differing contexts in which Foucault uses 'sex' rather than 'sexuality.' Specifically, it seems that when Foucault is speaking of political change he uses 'sex,' and when speaking of discursive conditions for the possibilities of the self's conception of its own unity he speaks of 'sexuality.' To explain why Foucault makes this distinction, and, most importantly, what this distinction may entail for the post-structuralist view of the text (in this case, their own texts), I wish now to turn, at some length, to Foucault's specific conceptions of power.

In his extremely lucid essay on the use of Foucault in certain gay movements such as Act Up, David Halperin comments on one of the most often misunderstood and yet important aspects of Foucault's philosophy--the meaning and prevalence of power. He writes, '[In] Foucault's view, power is not a substance but a relation...[it] is therefore not *possessed* but *exercised*...power is what characterizes the complex relations among the parts of a particular society' (emphasis in original).<sup>22</sup> Thus, Foucault does not conceive of power in the same way that Marxists may conceive of it. Power does not imply the property or ability to exercise domination over another group as this is merely one form of power, and furthermore, the

<sup>16</sup>Michel Foucault, *The History of Sexuality, Volume 1* 7.

<sup>17</sup>*History*, p. 67.

<sup>18</sup>*Ibid.*, p. 156.

<sup>19</sup>*Ibid.*, p. 147.

<sup>20</sup>No pun intended.

<sup>21</sup>*Ibid.*, p. 155.

<sup>22</sup>David M. Halperin, *Saint Foucault* (New York: Oxford University Press, 1995), p. 16-17.

dominated group's resistance to those who dominate it is itself 'never in a position of exteriority in relation to power.'<sup>23</sup> Instead, power is present in every relation between individuals or communities. Foucault once candidly expressed this by saying, 'We all fight each other.'<sup>24</sup> Rather than take the form of possession, which would lead to a language of economics, power as a relational field means that it must be spoken of in the discourse of tactics or games; in what Foucault calls an 'analytics of power.'

This omnipresence of power means that a general theory of power is not only impossible, but that to propose one actually obscures the real way in which power functions. Indeed, Foucault continually argues against the most common of the general theories of power which holds that power is held solely in the hands of a sovereign ruler. In an interview entitled 'The History of Sexuality,' Foucault says:

It seems to me that the problem [of power] is too often reduced...to the problem of sovereignty...between every point of a social body...there exist relations of power which are not purely and simply a projection of the sovereign's great power over the individual; they are rather the concrete, changing soil in which the sovereign's power is grounded, the conditions which make it possible for it to function.<sup>25</sup>

In other words, even if we use the terms formulated by theories of power, words like 'sovereign' or the 'social body', these theories themselves do not take into account that any seemingly totalizing, possessional form of power is the result of numerous micro-cosmic power relations that traverse all of society.

The specific form of this relation(s) of power that Foucault felt characterized our modern times was that of bio-power, treated at some length near the end of *History of Sexuality*. Foucault writes that the relations of micro-power, and subsequently, the relations of the sovereign power formed in this 'soil' (it is no mistake that he refers to this agent of biological reproduction), have moved from the ability to condemn certain people to death into the 'tendency to align itself with the exigencies of a life-administering power.'<sup>26</sup> To explain further, Foucault writes that at the advent of this bio-power the body came to be viewed as a machine which needed to operate at its maximum efficiency. The techniques for maximizing this efficiency came from various codes and sanctions which

<sup>23</sup>*History*, p. 95.

<sup>24</sup>Michel Foucault, 'The Confession of the Flesh,' in *Power/Knowledge: Selected Interviews and Other Writings 1972-1977*, ed. Colin Gordon (New York: Pantheon, 1980), p. 208.

<sup>25</sup>Michel Foucault, 'The History of Sexuality,' *Power/Knowledge*, p. 87.

<sup>26</sup>*History*, p. 136.

later came to be known as the disciplines or discourses of medicine, i.e. anatomical physiology, psychiatry, etc. From this focus on extending the life of the individual came its counterpart, extending the 'life' of the species. At this time, the sovereign became interested in conditioning species life to function at its maximum potential. This emphasis on creating and prolonging, on maximizing, the potential for the life of the community (in addition to the individual) led to the creation of normalizing institutions such as hospitals, asylums, and prisons, whereby those that currently held little value for engendering further species life could be corrected, pushed aside, or, in extreme cases, killed. To use Foucault's own terms, from the individual body as a machine, the "anatomy-politics of the human body," arose the concept of the social body, the "bio-politics of the population."<sup>27</sup> For Foucault, the most important thing about these two axes of bio-power is that they represent the two poles of his philosophical inquiry--the *anatomy-politics* can be viewed as the discursive systems for producing true individuals, and the *bio-politics* as the actual political actions created as a result of these systems. Sexuality was given its ever-present importance in the West because it mediated between these two levels of bio-power that can be characterized in these two different ways: power (control) and knowledge (rules), or the body as a body and the body as a species.<sup>28</sup> I now wish to turn back to my premise of examining sex/sexuality in light of Foucault's ideas on power/bio-power by looking first at what these conceptions of power held for Foucault's conception of political liberation.

Given that power is an ever-present field in the functioning of society, we may infer from Foucault's logic alone that he would be distrustful of any liberation movement that claimed an all-encompassing status for the overthrow of power as we know it (particularly the Marxist conception of an overthrow of existing bourgeois society). If power does not emanate from one source, any movement that attempts to overthrow this fictive source necessarily falls prey to a juridical conception of power that Foucault felt hid the more sophisticated workings of modern bio-power. Foucault's statement that '...there are many different kinds of revolution...as many kinds as there are possible recodifications of power relations,' as well as his view that the his role as an archaeologist and genealogist of power was only to expose the possible places where certain manifestations of power could be overthrown, seem to provide the reader with a much needed continuity between the logic of an author's writings and the author's actual feelings. Unlike Freud's ambivalences, as sketched

<sup>27</sup>*Ibid.*, p 139.

<sup>28</sup>*Ibid.*, p. 146.

by Davidson, these candidly personal statements made by Foucault seem to follow from his theoretical conceptions of power.

This idea of multiple points of resistance, of seizing power where it may be most fragile, leads to Foucault's distinction between 'sex' and 'sexuality' in *The History of Sexuality Volume 1*. Let us look again at the seemingly confusing claim of Foucault's claim that it is 'sex' (rather than 'sexuality') through which each individual has to pass in order to have access to his own intelligibility. In light of his views on power, bio-power, and activism, it seems that by speaking of 'sex' instead of 'sexuality,' he is consciously dealing with the political rather than discursive realm, the realm of controls rather than the realm of the disciplines. He is discussing the possible objection made by an imaginary Aronson like interlocutor who argues that by investing so much in sexuality, which is the product of discursive systems, Foucault has elided the 'sex' which is biological, functional, anatomical, in short, the real stuff of politics. To meet that objection, Foucault uses the idea that bio-power hinges on sexuality to mediate between the population and the individual, between rules and codes, and presents an individual's sex, meaning his sexual acts, as a possible place where knowledge and power become disjoined, separable to the eye of the genealogist/archaeologist. He points to one possible shortcoming in the rhetoric of sexuality as a possible point of *human* resistance. If the discourse on sexuality mediates between the human body and the species body, then sexual acts that do not directly lead to procreation are in a precarious position. By both living the life of a citizen under the species body, and engaging in other acts that do not prolong the biological body, the individual can bring to the forefront an ambivalence within the rhetoric of sexuality or bio-power. He can, in short, show that it is possible to fulfill the functions of a citizen under the dictates of species life without following the exact mandates of biologically correct reproduction, or, as it later came to be known, psychologically correct desires. Like Bhabha or Spivak, Foucault is pointing to sites of rupture, but, unlike these thinkers, he is providing a way for *human subjects* to use this ambivalence as a way to wield power.

Furthermore, by using 'sex' rather than 'sexuality,' Foucault is concerning himself with the logistics of discursive systems only in so far as they apply to the pragmatics of social action. An individual cannot effectuate change by attempting to recodify the systems of knowledge that have created the sexual self (sexuality); but he can alter the sexual acts that, when combined by these systems of truth-seeking knowledge, sketch the meaning of the self. By exploring different bodies and forms of pleasures through the accessible avenue of personal sexual choices, people will obscure the logical rules of a kind of sexuality discourse, they will no

longer understand the 'ruses of sexuality.'<sup>29</sup> This form of emancipatory power is thus exercised not from some authoritative point of emanation which undoes the oppressive by-products of bio-power (the sovereign/judicial conception); this rebellion is instead waged at the local, micro-cosmic level of sexual choice.

Now that I have hopefully used the concept of a 'style of reasoning' on Foucault himself in order to show that the ambiguity between 'sex' and 'sexuality' is a reflection of the human agent's role in manipulating the possible discrepancies between the textual and the political (a discrepancy that, *pace* Aronson, post-structuralism does in fact recognize), I want to turn to the possible implications that the imbrications of 'sex' and 'sexuality' may have for the post-structuralist view of the text itself.

In his original preface to the *History of Sexuality Volume Two*, Foucault writes:

There is no experience that is not a way of thinking and cannot be analyzed from the viewpoint of the history of *thought*; this is what might be called the principle of irreducibility of thought. (my emphasis)<sup>30</sup>

While at first glance this statement seems to coincide with Spivak's claim that political events do not happen without systems of thought, or that action can be deduced from thought, a more careful reading reveals an important difference between these theorists' views on knowledge and experience. Foucault is not claiming that experience is only a way of thinking or writing, but that experience (which leads to 'real' events) can in part be explained by examining the ways in which the subject's recognition of himself has been formed. Thus, we are not to avoid the political, but to see how the political has been informed by the textual, or where it rests in a necessary limit of the already existing textual systems, where, as I have suggested, sites of possible manipulation may occur. And as Davidson's article shows, Foucault's use of the word 'thought' does not necessarily need to refer to textual or written thought, as thought is both a product of textual strategies and the larger impulses of the *mentalite*.

Rather than draw superficial and disjointed oppositions between Foucault and Spivak, however, my more general project has been to illuminate how the embedded distinction between sex/sexuality in *The History of Sexuality Volume 1* is itself representative of an embedded

<sup>29</sup>Ibid., p. 159.

<sup>30</sup>Michel Foucault, 'Preface to The History of Sexuality Volume Two,' collected in *Ethics Subjectivity and Truth*, ed. Paul Rabinow (New York: The New Press, 1997), p. 201.

distinction between the political and the discursive. Although the discursive may subsume the political in order to hide places for overturning the political, to deduce what ultimate political implications a discourse may carry from an internal analysis of the discourse itself is not what Foucault has in mind. To return to Veyne, nuclear war cannot be traced from Descartes.

I have argued that these two poles of experience, knowledge and power, or discourse and action, play off one another, that they have similarities and differences. What is more important, however, is that by existing together, without 'exteriority,' systems of truth and effects of power create a monolithic structure that perpetuates itself by presenting and validating what are in fact false avenues for change. Politically oppressive climates obscure the distinction between oppressive actions and the systems of knowledge which allow such action--they prevent the subject from actually seeing where epistemological breakdowns between thought and action actually occur. As Foucault writes, '[power's] success is proportional to its ability to hide its own mechanisms.'<sup>31</sup> Foucault elaborates on this insight by showing how sex/sexuality seem to be used interchangeably in the discourse of sexuality, but in fact serve different functions.

By including this ambiguity to be discovered, however, he is also questioning his own status as a complete harbinger of truth, as someone who can work outside of the blurred distinctions between relations of power and systems of knowledge. Just as the ways in which the interrelations of knowledge and power in an oppressive instances must be worked out by the careful genealogist/archaeologist, so must the relations between emancipatory powers and 'liberating' techniques. By forcing the reader to undergo a kind of careful detective work to find 'the answer,' thereby allowing for multiple interpretations, multiple answers, Foucault is supporting his own advocacy of the need to think of many points of resistance, or to avoid a totalizing theory of emancipation which would preclude the subject from finding new ways to free himself from the specificities of the ever changing oppressive modalities of power. To force the reader to draw his own conclusions about what 'sex,' or 'sexuality' actually means in the workings of bio-power, Foucault is supporting his dictum that 'If everything is dangerous, then we always have something to do.'<sup>32</sup> Indeed, he is making even his own text a bit 'dangerous.'

The embeddedness of this distinction, however, is of only minor importance. More important than stylistic or personal concerns on

Foucault's part, however, is the inherent political power of this method, a power that calls into question both Aronson's charges of elitism and Spivak's equally incorrect, yet rhetorically different claims of the overarching importance of 'discourse.' Foucault's careful and difficult separation of the discursive and the political in order to analyze the ways in which they are joined and disjoined, this method of reading historical and literary texts in order to see how they cooperate and collaborate in the workings of power, blends the two criticisms, rendering them both, like the repressive hypothesis itself, falsely oppositional. In other words, both post-structuralism's critics and some its defenders indulge in a debate on post-structuralism that does not allow it to be shorn of its elitist implications. Neither only discursive or discursively political, Foucault's more sophisticated method of reading is what Halperin has characterized as 'Foucault's political approach to discourse.'<sup>33</sup>

To conclude, Foucault seems neither concerned with only the textual or the world of overarching ideology. His methodology merely seeks to use the text as its text, as it were, to locate the interrelations between political power and knowledge, to locate places where the two are disjoined and where a manipulation of this disjunction can occur. And furthermore, it is a methodology that seems to attempt these complex philosophical questions in a playful, non-egotistical, non-Aronson like way.

<sup>31</sup>History, p. 86.

<sup>32</sup>Michel Foucault, 'On the Genealogy of Ethics,' collected in *Ethics Subjectivity and Truth*

<sup>33</sup>Saint Foucault, p. 31.