Philosophy in Practice

VOLUME 3 – SPRING 2009

Philosophy in Practice

CONTENTS

ACKNOWLEDGMENTS

EDITORS

Melissa Foote
Peter Miller

FACULTY ADVISOR

Dr. Michael K. Shim

CALIFORNIA STATE UNIVERSITY, LOS ANGELES PHILOSOPHY FACULTY

Mark Balaguer (1992–), Chair, Ph.D. City University of New York. Philosophy of Mathematics, Metaphysics, Meta-ethics, Philosophy of Language, Logic

Mohammed Abed (2008–), Ph.D. University of Wisconsin, Madison. Ethics, Social and Political Philosophy, Philosophies of Violence, Genocide and Terrorism

Talia Bettcher (2000–), Ph.D. University of California, Los Angeles. History of Modern Philosophy, Philosophy of Self, Gender and Sexuality

Anna Carastathis (2009–), Ph.D. McGill University. Social Political Philosophy, Feminist Philosophy, Philosophy of Race, Critical Race Theory, Postcolonial Theory, Phenomenology

Richard Dean (2009–), Ph.D. University of North Carolina, Chapel Hill. Ethics, Kant's Moral Philosophy, Applied Ethics

Jennifer Faust (1993–), Ph.D. University of Illinois, Chicago. Epistemology, Philosophy of Biology, Metaphysics, Ethics of the Human Genome Project, Philosophy of Race

Ann Garry (1969–), Ph.D. University of Maryland. Feminist Philosophy, Philosophical Methodology, Epistemology, Applied Ethics, Wittgenstein, Philosophy of Law

Ricardo J. Gómez (1983–), Ph.D. Indiana University. Philosophy of Science and Technology, Philosophy of Mathematics, Kant, Latin American Philosophy

Ronald Houts (1983–), Ph.D. University of California, Los Angeles. Metaphysics, Epistemology, Logic

Henry R. Mendell (1983–), Ph.D. Stanford University. Ancient Philosophy, History of Ancient Mathematics and Science, Philosophy of Science, Metaphysics

David Pitt (2003–), Ph.D. City University of New York. Philosophy of Mind, Philosophy of Language, Metaphysics

Joseph B. Prabhu (1978–), Ph.D. Boston University. Philosophy of Religion, 19th and 20th Century German Philosophy, Moral and Social Philosophy, Indian and Comparative Philosophy

Sheila Price (1964–). Recent Philosophy, Comparative Religions, Medical Ethics, Environmental Ethics.

Michael K. Shim (2007–), Ph.D. State University of New York, Stony Brook. 20th Century Continental Philosophy, Phenomenology, Husserl, Modern Philosophy, Philosophy of Mind, Philosophy of Language

Kayley Vernallis (1993–), Ph.D. University of California, Berkeley. Moral Psychology, 19th and 20th Century Continental Philosophy, Feminist Philosophy, Ethics, Aesthetics, Gender and Sexuality

EMERITUS PROFESSORS

Sidney P. Albert (1956–1979). Aesthetics, Ancient Philosophy.

Thomas Annese (1961–1992). Epistemology, Modern Philosophy.

Sharon Bishop (1967–2004). Ethics, Political Philosophy, Philosophical Psychology, Feminist Ethics.

Donald Burrill (1962–1992). Ethics, Philosophy of Law, American Philosophy.

George Vick (1967–1997). Metaphysics, Phenomenology, Existentialism, Philosophy of Religion, Medieval Philosophy.

Professor Spotlight: Dr. Jennifer Faust

Dr. Jennifer Faust began teaching at California State University, Los Angeles in 1993 and became a full Professor of Philosophy in 2004. She received her B.S. in Biology and Philosophy from Spring Hill College in Mobile, Alabama, before earning her Ph.D. in Philosophy from the University of Illinois at Chicago in 1995. Her areas of interest include epistemology, philosophy of religion, philosophy of race, philosophy of science, bioethics, and genetics. Her primary area of specialization is applied epistemology, emphasizing the justification of religious beliefs, and also investigating how our knowledge of history has influenced the philosophy of law. Dr. Faust's biology background has contributed to her groundbreaking work in the philosophy of race, specifically dealing with race and genetics in biomedical research. Her work has been frequently published in philosophy journals and other scholarly publications.

In 2008, Dr. Faust left her position as Professor of Philosophy to become Acting Director of Faculty Affairs for California State University, Los Angeles, a position that she currently maintains. She oversees the hiring, tenure, legal issues, and leaves of absence for the entire faculty and staff of the university, which includes 550 tenure-track faculty, 1,200 adjunct professors, and 20,000 students. She offers guidance and direction for the development of campus personnel. However, Dr. Faust says that she misses working with the philosophy students and thinking about philosophy all day, but nevertheless enjoys the new challenges of her administrative position.

Dr. Faust has made countless contributions to the Department of Philosophy, including her role as the original faculty

adviser of *Philosophy in Practice*, which was published under the name *Affirmations* until its third volume. During the journal's first two years, Dr. Faust handled all of the editing work and supervised all of the student contributors. She designed a class around the production of the journal and, during its third year, involved the students in the editorial process. Dr. Faust has always had a deep commitment to ensuring that philosophy students receive the highest quality of education, and her extraordinary work with the journal has proved to be an invaluable resource for the continued academic success of the department.

In 2004, Dr. Faust was given the Distinguished Woman Award from California State University, Los Angeles, and was also honored with an Innovative Instruction Award for the 2000–2001 academic year. She served as a prominent member of the Academic Senate for nearly ten years, and served on the Executive Committee of the Senate. Dr. Faust also chaired the Faculty Rights Committee for the California Faculty Association, Los Angeles chapter from 2003 to 2008. During her sixteen years as an integral member of the Department of Philosophy at California State University, Los Angeles, Dr. Faust has served as Director of Graduate Studies.

The vacancy that Dr. Faust has left in the Department of Philosophy is noticeable, and the students miss her superior instruction and the dedication she had to the personal and intellectual development of all who had the pleasure of taking her classes and seminars. She made exceptional contributions to the department, and supervised numerous M.A. thesis projects in epistemology and philosophy of science. Dr. Faust hopes to one day to return to teaching, and to finish her book on the philosophy of race and science. While the Department of Philosophy misses her key role as professor and adviser, we wish her the best of luck in her new position, and deeply appreciate the contribution she has made to the lives of all philosophy students and faculty.

x

INTERNALISM AND THE LAWS OF LOGIC[1]

Michael Hatcher

§0. INTRODUCTION

I defend internalism about mental content from a popular argument against it. Before I relate this popular argument and my response to it, let me briefly explain what internalism is and why I find it important.

Philosophers of mind like to wonder about how a mental state—say, a particular belief, desire, or fear—gets its *content*. Suppose that an acquaintance says that she fears the same thing that she believes. Then, upon being questioned, she reveals that she both fears and believes *that the economy will never recover from this depression*. The content of her fear and her belief is the proposition *that the economy will never recover from this depression*. Propositions are mind and language-independent abstract objects.[2]

How does a mental state get its content? To ask the same question differently, on the basis of what does a mental state express the proposition that it expresses? *Internalists* answer that the internal features of a mental state determine its content. Specifically, internalists hold that if a person can't tell two mental states apart—i.e., if the states are *subjectively indistinguishable*—then the states have the same content.[3] For example, consider my belief that I'm typing on a computer now. Then suppose that I have a psychologically identical twin who believes that he is typing on a computer now, but who is actually stuck in the Matrix. Internalists hold that, since my mental state and my twin's state are subjectively indistinguishable, they have the same content. *Externalists* deny this. They think that factors external to a mental state have something to do with what content it has.

I'm an internalist. By my lights, the thesis that the contents

of my beliefs are partially determined by external factors implies that I can't tell what it is that I believe just by examining my beliefs.[4] But it seems that I often can tell what I believe just by examining my beliefs. Further, if I can't—prior to looking at the external world—discern the content of my beliefs, how can I compare that content with the world so as to determine whether a particular belief is true or not? Successful comparison of two things, in normal cases, implies access to each such that one can see whether they match.[5] So, in short, deny internalism and it's unclear how we can know the content of our own minds, in which case it's unclear how we can know much of anything at all. But we *do* know much of the content of our own minds, and we *do* know some things about the external world. Now, I hold that being skeptical on account of bad reasons can be as harmful to our lives as making assumptions without any good reasons. If I'm right about what externalism implies, then externalism results in such a form of skepticism.

Many externalists, of course, deny that externalism implies that we can't know the contents of our own minds.[6] But my current task isn't to engage the externalist on that front. The above simply gives a taste of why I think internalism is both true and important.

The popular argument against internalism that it's my task to defuse, put briefly, runs as follows:

(i) There are cases where subjectively indistinguishable mental states differ in reference.

(ii) A difference in reference entails a difference in content. So,

(iii) There are cases where subjectively indistinguishable mental states—contra internalism—*don't* have the same content.

This argument originated in the work of Putnam (1975) and McGinn (1977), among others. The example of my Matrix-bound twin appears to be an instance of (i): his belief and mine seem to be subjectively indistinguishable, but his refers to *him* while mine refers to *me*. So, if referential difference implies a difference in content (i.e., if (ii) is true), then internalism is false (i.e., (iii)

follows).

I will explain this argument more fully and carefully below. I think it's possible at this point, however, to give a preview of my defense of internalism.

In the past, some internalists (e.g., Farkas (2003a) and Crane (1991)) challenged (i). Such a challenge, however, falls flat. These same internalists, and others, have rightly recognized that (ii) must be denied. The internalist must deny that referential difference implies a difference in content (see Farkas (2008), Crane (2001), Pitt (unpublished a), and Katz (2004)).

My first move is to bring to light an implication of denying (ii) that, so far as I know, hasn't been given the attention it deserves. To deny (ii) is to say that the *reference* of some propositions depends on context, which is to say that the *truth-values* of some propositions depend on context. But this is to say that the propositions themselves, insofar as they are floating outside of any context in Plato's heaven,[7] are *neither true nor false*. And this is to deny that for all propositions p, p is either true or false, which is a popular way of formulating the Law of Excluded Middle ('LEM', for short). Further, to deny (ii) is to say that one and the same proposition can be true in one context and false in another. And this is to deny that for all propositions *p*, *p* is not both true and false, which is a popular way of formulating the Law of Non-Contradiction (LNC).

On first blush, this seems to reduce internalism to absurdity. Laws of logic like LEM and LNC aren't negotiable: if internalism violates them, then so much the worse for internalism.

My second move, however, is to argue that LEM and LNC can be adequately formulated in an internalist-friendly manner. Further, I argue that this formulation is not ad hoc (i.e., contrived), because there's an independent reason to think it's correct.

My thesis is that there is no cost in formulating LEM and LNC in an internalist-friendly manner. I think this thesis hones in on the pith of a deep disagreement between internalists and externalists. Externalists start with a belief about *logic*: that LEM and LNC, as popularly formulated, both hold (see especially, in this

connection, Fitch and Nelson (2007)). Internalists, however, start with a belief about *the mind*: that, for various reasons, the only plausible view is that a mental state's internal properties determine its content. I don't think the participants in this debate have been clear that these are two of the essential clashing assumptions between which we must arbitrate. The two assumptions are, at least initially, compelling but incompatible.

This paper has three more sections. In **§1**, I explain how internalism is committed to the denial of the popular formulation of LEM and LNC. In **§2**, I formulate LEM and LNC in an internalist-friendly manner and argue that this formulation is both adequate and independently plausible. In **§3**, I recap my argument and emphasize my thesis.

§1. INTERNALISM'S COMMITMENT

Let me express the externalist argument—above, I expressed it as (i), (ii), and (iii)—a bit more carefully. For heuristic purposes,[8] I'll relate it in terms of Putnam's (1975) Twin Earth thought experiment. Twin Earth is just like Earth, except that the clear, thirst-quenching liquid that falls from the sky isn't H_2O but instead a more complex, superficially indistinguishable chemical: XYZ. Back in 1750 ACE (before chemistry), Oscar and Twin Oscar both believe, of what each calls 'water,' *that it is delicious*. While Oscar's belief is about H_2O, Twin Oscar's is about XYZ. Here, then, is

The Externalist Argument[9]

(1) Oscar's belief that water is delicious refers to *water* (i.e., H_2O), whereas Twin Oscar's belief refers to *twater* (i.e., XYZ).

(2) That Oscar's belief has a different referent than Twin Oscar's entails that Oscar believes a proposition with (at *least*) a different constituent *sense*, i.e., a *different* proposition.

but, as is stipulated in the thought experiment,

(3) Oscar's belief is subjectively indistinguishable from Twin Oscar's.

which, when combined with (1) and (2), means that

(4) There's a case where subjectively indistinguishable mental states don't have the same content. In other words, internalism is false.

The Externalist Argument's premise (2) applies a central tenet in Fregean semantics:

The Fregean Tenet: *Sense determines reference*.[10]

To get a good grasp on The Fregean Tenet, note the distinction between 'sense' and 'reference.' '*Creatures with a heart*' and '*creatures with a kidney*' differ in sense (i.e., in meaning—these are different *concepts*) but they refer to (i.e., pick out) the same set of animals, because all creatures with hearts are creatures with kidneys. 'Creatures with a heart' *expresses* a sense. A sense is an abstract object that can be a constituent of a proposition. The reference of the sense in question, however, is simply a set of animals.

So what does it mean to say that *sense determines reference*? To put The Fregean Tenet differently, it's the view that sameness of sense entails sameness of reference, or, in contraposition, that difference of reference entails difference of sense. The idea is that though a given referent can be reached through various senses, whenever there are two different referents, *ipso facto* there are two different senses reaching them: the relation of sense to reference can be many/one but *never* one/many. So, looking again at (2), since Oscar's belief refers to water while Twin Oscar's refers to twater, there must be two different senses reaching these different referents. And this entails that the two Oscars believe different propositions.

At least, this last entailment holds on the view that propositions are simply composites of senses. This, so far as The Externalist Argument is concerned, is all right, because any sensible

5

internalist must hold that view of propositions (this will become clearer in section **§2**). To explain what I mean here by 'composite of senses,' let me introduce a notation that is common in literature on this topic, and which I too will exploit. On the view that propositions are simply composites of senses, the sentence or thought 'Water is delicious' expresses

<the sense of 'Water,' the sense of 'is delicious'>

Angle brackets, i.e. '< >', enclose *ordered sets*, i.e. sets such that the order of membership matters. Since propositions have parts and are structured, it's conventionally held that some such sets are propositions. So the view that propositions are simply composites of senses holds that propositions are like <the sense of 'Water,' the sense of 'is delicious'>, i.e. they're simply senses grouped up and structured in a certain way.

The idea behind The Fregean Tenet is plausible enough, at least initially. With respect to the creatures that have both a heart and a kidney, one might think that, though one can refer to that set of animals by means of different senses, the sense of, say, 'creatures with a heart' *can't* refer to a different set of animals.

The only way to handle The Externalist Argument, however, is to deny The Fregean Tenet, and so, in this way, deny not only (2), but also all of the many surrogates that would do the same work as (2). (There are innumerable cases like Twin Earth.) With respect to the idea that 'creatures with a heart' *can't* refer to a different set of animals, I've got the following response, inspired by Pitt (unpublished a, p. 6). Let's grant that 'creatures with a heart' expresses the same sense at different times. Still, plausibly, that phrase refers to *different* sets of animals at different times. Tomorrow, the set of creatures with a heart, given the birth and death of animals, will differ from the set of such creatures today. So, plausibly, the relation of sense to reference can be one/many. But this strategy will become clearer as we go.

Internalists have targeted premises of The Externalist Argument other than (2), but I find such attempts ineffective. Crane (1991, p. 290-293) denies (1). He thinks Oscar's belief refers to

both water and twater. Farkas (2003a, p. 167-169) is more nuanced. She has the following dilemma for the argument. If the chemical composition of the local watery liquid *matters* to Oscar and Twin Oscar, then (3) is false because they've got different concepts of water. But if the chemical composition doesn't matter, then (1) is false because their beliefs refer to both water and twater.

I don't doubt that many people have a concept of water in light of which Crane's objection works, i.e., a 'watery stuff' concept. Further, if such a loose watery stuff concept and a strict scientific concept were the only possible, then Farkas' dilemma would hold. But it's clear that those aren't the only concepts of water possible. Consider this concept: *that liquid, whatever it is, i.e., whatever its internal structure is, that is clear, quenching, and falls from our skies and fills our oceans*. Plug this concept into The Externalist Argument's premises, and you'll see that both (1) and (3) come out true. The two beliefs will refer to the watery liquid local to the believer, i.e. (1) comes out true, and the beliefs will be subjectively indistinguishable, i.e. (3) comes out true.

Internalists, then, must deny The Fregean Tenet (and, so, deny (2)). In fact, more recently, Farkas (2008, p. 10-11) and Crane (2001, p. 123) have come to recognize this. Let's put this denial as a positive thesis. Now, though I doubt MacFarlane would accept this positive thesis, I find a distinction he (2009, p. 232) makes between indexical and context-sensitive expressions useful in formulating the thesis I have in mind:

- An expression is *indexical* if and only if its *content* depends on the context.

whereas

- An expression is *context-sensitive* if and only if its *extension* [i.e., reference] depends on the context.

Applying this distinction to not only expressions but also sentences, thoughts, beliefs, and so on, to say that Oscar and Twin Oscar's beliefs differ in sense (and so in content) due to differing in reference—as The Fregean Tenet has it—is to say that their beliefs

are *indexical*: their beliefs' contents depend on their respective contexts.

The positive thesis I have in mind is the thesis that though some sentences, thoughts, beliefs, etc., are context-sensitive, *none* are indexical. This thesis could be given many titles. In light of MacFarlane (2009), we could call it 'Non-Indexical Contextualism (Generalized)'; following Katz (2004), we could call it 'Non-Fregean Intensionalism'; and, drawing from Balaguer (2005), we could call it 'The New General Proposition View.' In section **§2**, I will follow Balaguer's terminology when filling out this view, but I think introducing that terminology now would just muddle things. So, to keep things simple, and to emphasize the view's place in relation to The Externalist Argument, for now let's just call it

> **The Defeater View:** Though some sentences, thoughts, beliefs, etc., are context-sensitive, none are indexical.

So, on The Defeater View, (2) is false—though Oscar and Twin Oscar's beliefs differ in reference (and so are context-sensitive), they don't differ in content (they aren't indexical)—and The Externalist Argument is defeated. And the same fate meets any surrogate of (2), for example, the case of my Matrix-bound twin and I. For any such surrogate will apply The Fregean Tenet, and The Defeater View, recall, is designed as a systematic denial of The Fregean Tenet.

Internalism, then, is committed to The Defeater View. In fact, as soon as one grants that some beliefs are context-sensitive, The Defeater View could be seen as a way of formulating internalism. Indexicality is *the property a mental state has when its content-determination depends on context* and context is plausibly thought of as *mind-external*. So to say that mental states are non-indexical is just to affirm internalism: content-determination depends solely on what's *mind-internal*.

There's a worry with The Defeater View, however. LEM and LNC, popularly, are formulated as follows:

LEM: For all propositions *p*, *p* is either true or false.

LNC: For all propositions *p*, *p* is not both true and false.

Now, it might not be immediately obvious that The Defeater View entails that LEM and LNC, as popularly formulated, are false. But it does.

On The Defeater View, context-sensitive thoughts or sentences like 'He is in Paris,' 'It is raining now,' 'That object is interesting,' and 'I am starving' express full-fledged propositions which, considered independently of context (i.e., place, time, thinker/speaker, possible world, etc.), *do not have truth values at all*. In other words, <the sense of 'It,' the sense of 'is raining now'>, for example, is neither true nor false *in itself*. That proposition needs a context. In a given context, what's expressed by 'It is raining now', i.e. <the sense of 'It', the sense of 'is raining now'>, will refer to the time present and the place local to the context, and so will be either true or false. If it's raining at that time and place, then the proposition is true in that context; in another context what's expressed by 'It is raining now' will be false. Note well: it's the same proposition, i.e., <the sense of 'It', the sense of 'is raining now'>, expressed in either case.

So The Defeater View entails that LEM, as popularly formulated, is false: the proposition *that it is raining now*, outside of a context, has no truth-value—it's not either true or false. And propositions do exist outside of contexts: they're abstract objects, and abstract objects exist necessarily. As Balaguer (unpublished, p. 17) puts it, on this view "sense doesn't determine extension [i.e., reference] all by itself, and so in general… propositions don't have fixed truth values."

Likewise, The Defeater View entails that LNC, as popularly formulated, is false: the proposition *that it is raining now* will be true in one context and false in another. Hence, on The Defeater View, that proposition is both true and false.

What commitments must the internalist make to deal with The Externalist Argument? The internalist is committed to denying The Fregean Tenet by means of The Defeater View, and

is thereby committed to denying both LEM and LNC, as popularly formulated.

But is this a cost of internalism?

§2. INTERNALIST-FRIENDLY LOGICAL LAWS

Taken by itself, the argument of section **§1** could be seen as a *reductio ad absurdum* against internalism. LEM and LNC, as popularly formulated, sure seem like laws of logic, and surely any view that denies such a law is *ipso facto* false! Now, some (e.g., Pitt, Balaguer, Priest, and others)[11] feel little anxiety about disposing with LEM *in general*. I, however, feel differently about both laws. These logical laws, by my lights, are meant to capture the intuition that there can be no *claim* that is neither true nor false or both true and false. For now, think of claims simply as *assertions that something is the case*.

If denying LEM and LNC, as popularly formulated, amounts to holding that a claim can lack a truth-value, or have two of them, then internalism is doomed. Here I suggest that internalism isn't doomed because, in short, a proposition like <the sense of 'It', the sense of 'is raining now'>, *outside of a context*, is not a claim. And there's a formulation of LEM and LNC available that entirely protects the relevant intuition about claims.

My formulation of LEM and LNC will help clarify my suggestion that <the sense of 'It', the sense of 'is raining now'> is not a claim. Now, in service of clarifying this formulation, I will take a look at two views on the nature of propositions. On the one hand, there's the specific view that falls out of internalism's commitment as indicated in section **§1**; on the other, there's a view that many externalists find attractive, which view I will construe broadly.

Suppose that Barack Obama, reflecting on his recent election, utters quietly to himself the following: "I am both anxious and excited." On The Defeater View—recall, the view that though some sentences, thoughts, beliefs, etc., are context-sensitive (i.e., variant in reference due to context), *none* are indexical (i.e., variant

in content due to context)—this utterance expresses a proposition composed solely of senses, i.e.

<the sense of 'I', the sense of 'am both anxious and excited'>,

which determine reference *only relative to context*.

After Balaguer (2005), let's rename The Defeater View the following: 'The New General Proposition View' (or, for short, 'The New GP View'). To get a feel for this terminology, note that propositions that contain *referents* are often called *singular* propositions. The New GP View is a "general" proposition view because its propositions, as distinct from singular propositions, are composed solely of senses and contain no referents. It's "new" because, in denying that sense determines reference, i.e. The Fregean Tenet, it departs from traditional Fregeanism.

Many externalists, however, countenance singular propositions. On this view, Obama's utterance expresses the singular proposition

<Barack Obama, the property of being both anxious and excited at time t>

The above ordered set's first constituent is *the actual, flesh-and-blood man, Barack Obama*. Let's call this broadly construed view, of which there are many variants, 'The SP View'. It should be clear enough why the view is a singular proposition view (its propositions contain referents).

How does distinguishing these views help me formulate LEM and LNC in an internalist-friendly manner? Well, the first thing to see is a point Balaguer (unpublished) makes. On The New GP View, the referents of the constituents of

<the sense of 'I', the sense of 'am both anxious and excited'>,

given the context of its utterance (i.e., uttered by Barack Obama at time t), compose, when grouped up in an ordered set,

<Barack Obama, the property of being both anxious and excited at time t>

11

But this *just* is a singular proposition. So, The New GP View's general propositions ('GPs', for short) and The SP View's singular propositions (SPs) bear a certain relation: an SP is going to be *the ordered set of the referents, in a given context, of the constituents of a GP*.^ Balaguer (unpublished, p. 17) puts this as follows: SPs are identical with the truth conditions for GPs. The GP we've been considering is true if and only if Barack Obama has the property of being both anxious and excited at time t. (Of course, the above relation doesn't hold for GPs expressed by utterances not containing singular terms, i.e., terms that purport to refer to an individual object, but it isn't about such utterances that the two views importantly differ, so I ignore them for now.)

SPs/truth conditions are *abstracta* (abstract objects), and I (following Balaguer (unpublished, p. 17-18) and all adherents of The SP View) see nothing wrong with thinking that they have truth-values. Perhaps here we have what we need: a kind of abstracta that approximates what we might think of as *claims* (i.e., assertions that something is the case), and therefore a kind of abstracta in terms of which LEM and LNC can be formulated in an internalist-friendly way.

The formulation—let's combine LEM and LNC for brevity here—might go like this: For all singular "propositions"/truth conditions p, p is either true or false and p is not both true and false. (The scare-quotes indicate that the internalist doesn't consider SPs to be *propositions*.) Then the internalist could say that the following *contextualized* law holds of *propositions* (i.e., GPs). Where a context includes time, place, subject, and possible world, for all propositions p and for every context C, p is either true in C or false in C and p is not both true and false in C—*by virtue of referring, in C, to a singular "proposition"/truth condition that (outside of any context) is true or false and not both*.

The above position is somewhat plausible, in my view, and, for present purposes, anyone who really likes it is free to hold it. It's clear that there's a sort of abstracta in terms of which the popular formulation of LEM and LNC holds even if internalism (and thus, The New GP View) is true. And this sort of abstracta

approximates what we might think of as *claims*, i.e., assertions that something is the case.

I, however, advance a more nuanced view, for two reasons. First, I think there are some claims that are not SPs. Second, I think it's worthwhile to formulate LEM and LNC in terms of *propositions*, i.e., since I'm an internalist, GPs.

Here's an utterance that expresses a claim, but no (full-fledged) SP: "Santa Claus loves his reindeer." The SP would be <Santa Claus, *being-one-who-loves-his-reindeer*>, but, because Santa Claus doesn't exist, there is no such SP. In the mouth of a five-year-old (i.e., one who actually believes that Santa Claus is real), "Santa Claus loves his reindeer," by my lights, expresses a claim that is false. Here, then, is what I mean, precisely, by 'claim':

- *X* is a *claim* if and only if (i) *X* purports to refer to something and (ii) *X* purports to predicate something of that something.

In the case of claims involving names and indexicals, the 'something' in (i) will be an object, and the 'something' in (ii) will be a property (e.g., *that man over there is tall*, and *I feel better now*). In the case of mass nouns, however, the 'something' in (i) will be a set (e.g., *all bachelors are lonely*). And, in some cases, the 'something' in (i) will itself be a property (e.g., *triangularity entails trilaterality*). That *X* need only purport to refer to something to count as a claim covers the *Santa Claus loves his reindeer* case. That *X* need only purport to predicate something covers cases of (false) claims like *that figure has the property of being both square and circular* (the predicated property doesn't exist).

With this in mind, it's easy to see that every GP in a given *context* (i.e., specification of subject, time, place, and possible world) is a claim. A GP in a given context, by that very fact, will purport to refer to something, and will purport to predicate something of that something. Of course, some GPs are also claims outside of any context, e.g. *triangularity entails trilaterality*. (One might think that *all bachelors are lonely* is a claim outside of any context, but I'm not so sure. A world—a big context, so to speak—needs to be specified in order to get the set of bachelors of

which the predication is made (see Pitt, unpublished a, p. 6).) But, since GPs that are claims outside of any context are also claims (the same ones) inside every context, they are claims inside any given context, and hence there's no real loss in stipulating (which I do now) that

A claim is a GP in a context.

So, where 'propositions' are *GPs* and not SPs, an internalist-friendly LEM and LNC can be formulated quite well, as follows, to cover all *claims*:

$LEM_{Propositions\ in\ Contexts}$ ('LEM_{PIC}', for short): For all propositions p and for every context C, p is either true in C or false in C.

$LNC_{Propositions\ in\ Contexts}$ (LNC_{PIC}): For all propositions p and for every context C, p is not both true and false in C.

To bring us back to our focal worry: in light of LEM_{PIC} and LNC_{PIC}, is the denial of LEM and LNC, as popularly formulated, problematic? Recall: the basic reason to think such laws true in the first place is the intuition that there can't be some *claim* that's neither true nor false, or both true and false. LEM_{PIC} and LNC_{PIC}, however, fully account for this intuition.

LEM_{PIC} and LNC_{PIC} make clear that on internalism's The New GP View there aren't any abstracta out there that are *claims* that are neither true nor false or both true and false. Some GPs, *unless put in a context, aren't claims at all*. For example, consider <the sense of 'That flower', the sense of 'is yellow'>. This proposition, outside of a context, doesn't even *purport* to refer to a flower, and it doesn't even *purport* to predicate yellowness of some flower. GPs like this, outside of a context, aren't claims. They make no assertion that something is the case: they neither purport to refer to something nor purport to predicate something of that something. Hence, their lack of truth-value should trouble us no more than the truth value*less*ness of abstracta like *numbers*, *sets*, *properties*, etc. Likewise, when in different contexts, and so

when predicating a property of *different* objects, it's not troubling that one and the same GP can be true in one context and false in another. So, it seems that substituting LEM$_{PIC}$ and LNC$_{PIC}$ for the popular formulation of LEM and LNC is *not* a cost of internalism.

Indeed, it isn't a cost. In fact, there's reason to think LEM$_{PIC}$ and LNC$_{PIC}$ are the right way to formulate LEM and LNC *independently* of the internalist/externalist issue. But, before I relate that independent reason, I think a brief recap of what I've argued so far is in order.

Internalists, in light of The Externalist Argument, must systematically abjure The Fregean Tenet. In other words, internalists must endorse The Defeater View, which *just is* The New GP View. That view, however, entails that LEM and LNC, as popularly formulated, are false, which *prima facie* reduces internalism to absurdity. But there's a way out. LEM and LNC needn't be formulated merely in terms of propositions as

LEM$_{Propositions}$ ('LEM$_P$'): For all propositions *p*, *p* is either true or false.

and

LNC$_{Propositions}$ ('LNC$_P$'): For all propositions *p*, *p* is not both true and false.

but can instead, by the internalist's lights, be more accurately and adequately formulated in terms also of *context* as LEM$_{PIC}$ and LNC$_{PIC}$. And LEM$_{PIC}$ and LNC$_{PIC}$ capture exactly what these laws are supposed to: the intuition that there can be no *claim* that's neither true nor false or both true and false. So countenancing LEM$_{PIC}$ and LNC$_{PIC}$ instead of LEM$_P$ and LNC$_P$ is not a cost of internalism.

Here's the independent reason to prefer LEM$_{PIC}$ and LNC$_{PIC}$ to LEM$_P$ and LNC$_P$. It's a reason that has to do with *modality* (i.e., possibility, contingency, and necessity). It's commonly recognized that some propositions have their truth-values *necessarily* whereas others have them *contingently*. The standard way to explain this is to say that some propositions are true in every

15

possible world (i.e., roughly, every *total way things could have been*) whereas others are true in some possible worlds and false in others. The former have their truth-values necessarily, the latter contingently. So, for example, the proposition that *Barack Obama is the 44th President of the United States* is true in the actual world but false in the possible world in which John McCain won the 2008 presidential election.

LEM$_P$ and LNC$_P$, however, have a hard time accounting for this. Here I develop and apply an argument to which Pitt alludes (unpublished a, p. 7). Let 'W^1' be the actual world, and 'W^2' be the possible world in which John McCain won. Consider the following uncontroversial modal statement:

(A) *That Barack Obama is the 44th President of the United States* is true in W^1 and false in W^2.

LNC$_P$ entails that (A) is false. According to LNC$_P$, for all propositions *p*, *p* is not both true and false. But *that Barack Obama is the 44th President of the United States* indeed is both true and false. It's true in W^1 and false in W^2. It's *one and the same* proposition that's true in W^1 and false in W^2, *contra* LNC$_P$. This is similar to how *that it is raining now* is true at one time and place and false at another according to internalism's The New GP View. As I've defined 'context' (i.e., as a specification of subject, time, place, and possible world), these two propositions both differ in truth-value in different contexts. And this just means that LEM$_{PIC}$ and LNC$_{PIC}$, as opposed to LEM$_P$ and LNC$_P$, are appropriate for them. LEM$_{PIC}$ and LNC$_{PIC}$ allow some propositions to be *contingently* true. LEM$_P$ and LNC$_P$ yield all propositions as *necessary* truths if truths at all (let's call this view '*necessitism*'). Necessitism, however, is implausible.

To see that this isn't some quick trick, note that externalism's tendency towards The SP View, i.e., the view that propositions sometimes contain a *referent* as a constituent, also tends towards necessitism. Kaplan, for example, thinks that Obama's utterance of "I am both anxious and excited" expresses <Barack Obama, the property of being both anxious and excited> (1989, p. 727). I didn't

include 'anxious and excited *at time t*' because Kaplan seems to be a *temporalist*, i.e., he seems to think that one and the same proposition can be true at one time and false at another (Ibid., p. 735). One of his reasons, however, for thinking that Obama's utterance expresses <Barack Obama, the property of being both anxious and excited> as opposed to <the sense of 'I', the sense of 'am both anxious and excited'>, is that the former, but not the latter, obeys LEM_P and LNC_P (Ibid., p. 743). But if obedience to LEM_P and LNC_P is what's desired, <Barack Obama, the property of being both anxious and excited> won't do either: that proposition can be true at one time and false at another. But then, even <Barack Obama, the property of being both anxious and excited *at time t*> won't satisfy LEM_P and LNC_P: that proposition, while true in W^1, might be false in W^2. What's really needed is <Barack Obama, the property of being both anxious and excited at time *t in W^1*>. The problem here, however, is that this makes what's expressed by Obama's utterance a *necessary* truth, because <Barack Obama, the property of being both anxious and excited at time t in W^1> is true in (or, 'at')[12] *every possible world*. This necessitism, though, is downright implausible. What Obama's utterance expressed is, if true, clearly a *contingent* truth. What Obama said could have been false.

Let me address a possible worry with this argument. To say that *p* is true in the actual world and false in some other possible world is just a regimented way of saying that *p* is true and that *p* could have been false. This isn't to say that *p* literally is both true and false, so contingent propositions aren't counterexamples to LEM_P and LNC_P.

Now, some philosophers might respond that possible worlds talk is more than simply a way of talking. But, for now, let's grant that it is just a way of talking. The modal argument for LEM_{PIC} and LNC_{PIC} can be recast as follows. A contingent proposition is true in virtue of *a relation to something else*.[13] In the case of *that Barack Obama is the 44th President of the United States*, that proposition is true *in virtue* of being in the correspondence-relation to *the way things actually are*. If things had been different,

then that very same proposition *would have* been false. Suppose that we can't exploit possible worlds talk in order to establish directly that LEM_P and LNC_P fail. Still, it's granted that the very proposition that is actually true *would have been* false if things had been different. And this implies that we'd need a good reason to not grant that the very proposition that is true in one context in the actual world can be false in another. In other words, in order to protect LEM_P and LNC_P, one would have to explain why the admitted difference between possible worlds and contexts within the same world generates a *principled* difference between how propositions get their truth-values in worlds and how they get them in contexts within a world. But it's unclear what such an explanation might be.[14]

Suffice it to say, then, that a consideration from modality offers support for LEM_{PIC} and LNC_{PIC} over against LEM_P and LNC_P. Now, admittedly, there are philosophers who have done work in modality that think the exact opposite that I do (e.g., Kaplan).[15] But I haven't the space to engage them in this paper. I offer the above consideration simply to stave off the counterargument that, while my formulation of LEM and LNC as LEM_{PIC} and LNC_{PIC} protects internalism from absurdity, it's an *ad hoc* (i.e., contrived) maneuver. The charge is that if the only reason to think LEM_{PIC} and LNC_{PIC} are the correct laws, as opposed to the more popular LEM_P and LNC_P, is internalism *itself*, then LEM_{PIC} and LNC_{PIC} are contrivances developed merely to defend internalism.

Now, I think that even if internalism were the only reason to think LEM_{PIC} and LNC_{PIC} true, it'd still be a good reason. (My view here would be more plausible in light of substantive argument for internalism, something that I can't provide in this paper.) But, in any case, as the above consideration from modality indicates, internalism isn't the only reason to think LEM_{PIC} and LNC_{PIC} true. The view that LEM_{PIC} and LNC_{PIC} are preferable to LEM_P and LNC_P is part-and-parcel of a general philosophical position on the nature of propositions and how they get their truth-values.

Thus, it's not merely the case that internalism's commitment to LEM_{PIC} and LNC_{PIC}, instead of LEM_P and LNC_P, is not absurd.

LEM$_{PIC}$ and LNC$_{PIC}$, in addition, are not ad hoc. There's an independent reason to think that they're correct, which lends reason to think that internalism's commitment comes at no cost at all.

§3. CONCLUSION

Internalists and externalists start from different places. Externalists start with LEM$_P$ and LNC$_P$, from which The Fregean Tenet follows. Accordingly, in the face of a case like Twin Earth, where Oscar and Twin Oscar's subjectively indistinguishable mental states differ in reference, the externalist infers that their states have different contents. This is The Externalist Argument. Internalists, having their reasons for internalism, start with the conviction—once they are clear on the issue—that something must be wrong with The Fregean Tenet. But, so far as I know, internalists have not yet been adequately clear that to deny The Fregean Tenet is to deny LEM$_P$ and LNC$_P$. And they haven't been clear that this is, at least initially, worrisome.

It is this gap in the internalist's strategy that I've tried to fill. First, I've made explicit that internalism is incompatible with LEM$_P$ and LNC$_P$. The Fregean Tenet's denial, i.e. The Defeater View—otherwise known as The New GP View—countenances GPs that lack truth-value absent context and that have one truth-value in one context and another in another.

Second, I've offered formulations of LEM and LNC that do all the work we could ever ask of such laws. LEM$_{PIC}$ and LNC$_{PIC}$ protect the intuition that there can be no claim that's neither true nor false or both true and false. For example, the proposition <the sense of 'That object', the sense of 'is spherical'>, unless put in a context, is not a claim, i.e., an assertion that something is the case. That proposition, outside of a context, neither purports to refer to an object nor purports to predicate sphericality of some object, so its lack of truth-value is no more troubling than the truth-valuelessness of numbers, sets, properties, etc. Likewise, when in different contexts, and so when predicating sphericality of different objects, it's not troubling that it can be true in one

context and false in another.

Finally, I argued that LEM_{PIC} and LNC_{PIC} are not ad hoc formulations of LEM and LNC. There is a consideration from modality that's an independent reason to prefer LEM_{PIC} and LNC_{PIC} to LEM_P and LNC_P. Namely, it's plausible to think that LEM_P and LNC_P are not friendly to the modal fact that some truths are contingent, whereas LEM_{PIC} and LNC_{PIC} are. Or, at least, given contingent truth, the advocate of LEM_P and LNC_P has some explaining to do. She must explain why, though a self-same proposition *would have* been false if things had been different, it's not plausible to think that a self-same proposition can be *actually* true in one context and false in another.

The upshot of all this is that there's no cost in formulating LEM and LNC in an internalist-friendly manner, which is my thesis. This thesis, if correct, de-fangs The Externalist Argument.

It's good to be clear about the limits of my thesis. My thesis does not count as a full-fledged defense of internalism. I deal with only one argument against internalism; there are others.[16] And, besides a teaser in section **§0**, I offer no positive argument on behalf of internalism.[17] Now, I think that a full-fledged defense of internalism is indeed possible, but that's an argument for another day.

Notes

1. Many thanks are due to Dr. David Pitt and Dr. Mark Balaguer for invaluable interaction with prior drafts of this paper.

2. A note for those unfriendly to abstract objects: So far as I've given the matter thought, I think I am as well. Feel free to think of my abstract-object-talk as loose talk. I think the whole internalist/externalist debate could be run in nominalistic terms, but that would be more time-consuming.

3. There are other construals of internalism. For example, some see internalism as the view that mental states that correlate with (or, are) physically identical brain states express the same proposition. I, following Farkas (2003b), don't find that to be the really interesting internalist/externalist debate about content. And, for the record (though I'm agnostic about this in the main text), I think externalism wins that debate. See Fisher (2007) for an interesting argument to that effect.

4. It also seems so to Pitt (2004). Pitt argues that, since (a) we can identify the thoughts we're thinking, (b) there must be a unique *cognitive* sort of phenomenology. Pitt's paper is a good place to start for those intrigued by my admittedly inadequate pro-internalist paragraph.

5. For interesting discussion relevant to this point, see Ten Elshof (2005), pp. 7-8.

6. See, in this connection, Ludlow and Martin's (1998), a collection of papers on the topic of whether externalism is compatible with self-knowledge.

7. Plato, Socrates' best student, is famous for the idea that there are Forms, i.e., abstract objects, which account for the fact that objects and actions have the properties they do. For example, a given flower is beautiful because it imitates or participates in the Form of Beauty, and protecting the innocent is just because it imitates or participates in the Form of Justice. 'Plato's heaven', straightforwardly, refers to the non-spatiotemporal and non-causal realm in which abstract objects exist, but it could just as easily (and perhaps more accurately) be taken as a metaphor for the non-spatiotemporal and non-causal way in which abstract objects exist.

8. I say 'for heuristic purposes' because I'm not concerned with Putnam-exegesis in this paper. Putnam's Twin Earth case is just a good way to set the stage.

9. This is an argument, to be specific, for *natural kind* externalism, so-called due to the fact that the internal-structure of the natural kinds involved (see *water*) play the role they do. I don't address social externalism in this paper. For the record, I think *social* externalism is a total nonstarter, see Pitt (unpublished b) for why.

10. Balaguer, in conversation, claims that Frege's view is that it's *sense-plus-the-actual-circumstance* that determines reference. Balaguer's probably right. I don't express The Fregean Tenet like this because (i) 'sense determines reference' can be interpreted in Balaguer's way if one likes, (ii) expressing Frege's view in that way would muddle things in the main text, and (iii) it would make no difference at all to The Externalist Argument or my response.

11. Pitt and Balaguer have expressed as much in conversation, for different reasons. Priest (1998) expresses as much in print, and he's the co-editor of a collection (2006) devoted to examining similar divergences from classical logic.

12. There's a worry on The SP View that an SP can't be true in a world in which one of the SP's constituents doesn't exist, for then the SP doesn't exist *in* that world. For example, <Obama, *being-one-who-was-never-born*> doesn't exist in those possible worlds in which Obama was never born. Some think this can be remedied by saying that such an SP is nonetheless true *at* that world. See Matthew Davidson (2007) for interesting discussion of this view.

13. Pitt argued convincingly, in conversation, that necessary propositions are also true in virtue of a relation to something else. For example, the necessarily true proposition that *triangles are trilaterals* is true in virtue of something else, namely the properties of *triangles* and *trilaterals*.

14. I owe the recast modal argument to Pitt, in conversation.

15. One wishing to construct a reply on Kaplan's behalf could start with the 28th footnote of his 1989, pp. 735-6.

16. On my view, the most important distinct argument against internalism is that no *empirically adequate* semantic theory for natural languages (i.e., a theory that is adequate to native speakers' intuitions about the *contents* of the sentences uttered in their language) can be internalist-friendly, i.e., such that sentence utterances express GPs in the manner of The New GP View. For example, suppose that Michelle Obama says that she is hungry (by saying "I am hungry") and Barack Obama says that he is hungry (ditto). There's a strong intuition, contra The New GP View (on which both utterances express <the sense of 'I', the sense of 'am hungry'>), that these two sentences say different things, i.e. express different propositions. Many externalists think that internalism's The New GP View simply can't handle such empirical constraints, and others like them. For example, Williamson holds that these considerations "put the burden of proof on the internalist" (2000, p. 54). And McGinn said as much before Williamson did: "We are dealing here with a semantic *datum*—something any reasonable theory should try to respect, not flout" (1989, p. 38).

    But, by my lights, Balaguer (unpublished) convincingly argues that The SP View and The New GP View can each be developed in such a way that both turn out to be empirically adequate. Obviously, I can't relate his entire argument, but to get a flavor for it, note how Balaguer deals with the case of Michelle and Barack: there's really just an intuition that they say different things in one sense and the same thing in another sense (unpublished, p. 18). And that intuition can be satisfied on both views: The New GP View has it that they say different things in the sense of the *truth conditions* (recall, these will be SPs) of their utterances, the same thing in the sense of the GP expressed by their utterances. And The SP View can simply introduce *character* as that which the two utterances have in common, such that Michelle and Barack say different things in the sense of the SPs their utterances express, and the same thing in the sense of the character of their utterances (Ibid). As Balaguer (unpublished, p. 25) puts it: "We don't have any intuition to the effect that they *really*—or *ultimately*, or *primarily*—say the same thing or different things."

    The above is Balaguer's argument that both The SP View and The New GP View satisfy the *what-is-said* constraint on the empirical adequacy for semantic theories of natural languages. There are four other constraints with respect to which Balaguer makes the same argument (Ibid., p. 16), and I'm am convinced that he is successful. I should note, though, that my

explicitness about LEM$_\text{PIC}$ and LNC$_\text{PIC}$ provides some clarity I think critical with respect to Balaguer's argument about what we could call the *truth-bearer* constraint: native speakers' intuitions about what sort of entities are the bearers of truth-value in connection with a sentence utterance. I strongly encourage those interested to read Balaguer's (unpublished), as well as his (2005).

17. I think another good argument for internalism, though, is as follows:

(1) Mental states have their contents essentially, i.e., a state with a different content is a different state.

(2) Mental states don't have any *relational properties* essentially, i.e., my twin in the Matrix is in the same mental state as I am, though our states share no relational properties. Therefore,

(3) The *non*-relational (i.e., *internal*) properties of a mental state determine its content: internalism is true.

I admit, though, that developing this argument would take another paper. See David (2002) for interesting discussion of (1).

Bibliography

Balaguer, Mark. (2005) "Indexical Propositions and *De Re* Belief Ascriptions", *Synthese* 146, pp. 325-55

_____ (unpublished) "Is there a Fact of the Matter Between Direct Reference Theory and (Neo-)Fregeanism?"

Crane, Tim. (1991) "All the Difference in the World", reprinted in: Pessin, Goldberg, and Putnam (Eds) (1996) *The Twin Earth Chronicles*, pp. 284-304 (New York: M.E. Sharpe)

_____ (2001) *Elements of Mind: An Introduction to the Philosophy of Mind* (New York: Oxford University Press)

David, Marian. (2002) "Content Essentialism", *Acta Analytica* 17, pp. 103-114

Davidson, Matthew (Ed) (2007) *On Sense and Direct Reference: Readings in the Philosophy of Language* (Boston: McGraw-Hill)

_____ (2007) "Transworld Identity, Singular Propositions, and Picture-Thinking", in: Davidson, Matthew (Ed) (2007) *On Sense and Direct Reference: Readings in the Philosophy of Language*, pp. 720-782 (Boston: McGraw-Hill)

Farkas, Katalin. (2003a) "Does Twin Earth Rest on a Mistake?", *Croatian Journal of Philosophy* 3, pp. 155-169

_____ (2003b) "What is Externalism?", *Philosophical Studies* 112, pp. 187-208

_____ (2008) "Phenomenal Intentionality Without Compromise" forthcoming April 2008, in a special issue of *The Monist*, on "Intentionality and Phenomenal Consciousness", edited by Uriah Kriegel and Terence Horgan.

Fisher, Justin. (2007) "Why Nothing Mental is Just in the Head", *Noûs* 41, pp. 318-334

Fitch, Greg and Nelson, Michael. (2007) "Singular Propositions", in *Stanford Encyclopedia of Philosophy*, accessible online at http://plato.stanford.edu/entries/propositions-singular/

Kaplan, David. (1989) "Demonstratives", reprinted in: Davidson, Matthew (Ed) (2007) *On Sense and Direct Reference: Readings in the Philosophy of Language*, pp. 720-782 (Boston: McGraw-Hill)

Katz, Jerrold. (2004) *Sense, Reference, and Philosophy* (New York: Oxford University Press)

Ludlow and Martin (Eds) (1998) *Externalism and Self-Knowledge* (Stanford: CSLI Publications)

MacFarlane, John. (2009) "Nonindexical Contextualism", *Synthese* 166, pp. 231-250

McGinn, Colin. (1977) "Charity, interpretation, and belief," *Journal of Philosophy* 74, pp. 521-535

_____ (1989) *Mental Content* (Oxford: Basil Blackwell)

Pessin, Goldberg, and Putnam (Eds) (1996) *The Twin Earth Chronicles* (New York: M.E. Sharpe)

Pitt, David. (2004) "The Phenomenology of Cognition Or *What Is It Like to Think That P?*", *Philosophy and Phenomenological Research* 69, pp. 1-36

_____ (unpublished a) "Demonstrative Thoughts"

_____ (unpublished b) "The Burgean Intuitions"

Priest, Graham. (1998) "What's So Bad About Contradictions?" reprinted in: Priest, Beall, and Armour-Garb (Eds) (2006) *The Law of Non-Contradiction: New Philosophical Essays*, pp. 23-38 (Oxford: Clarendon Press)

Priest, Beall, and Armour-Garb (Eds) (2006) *The Law of Non-Contradiction: New Philosophical Essays* (Oxford: Clarendon Press)

Putnam, Hilary. (1975) "The Meaning of 'Meaning'," reprinted in: Pessin, Goldberg, and Putnam (Eds) (1996) *The Twin Earth Chronicles*, pp. 3-52 (New York: M.E. Sharpe)

Ten Elshof, Gregg. (2005) *Introspection Vindicated: An Essay in Defense of the Perceptual Model of Self-Knowledge* (Burlington: Ashgate Publishing Company)

Williamson, Timothy. (2000) *Knowledge and Its Limits* (Oxford: Oxford University Press)

REIDENTIFICATION AND FIRST-PERSON EXPERIENCE

Peter Miller

Philosophy of the self is a branch of philosophy that is chiefly concerned with providing an account of personal identity. To grasp the importance of personal identity, we might ask questions like: What makes me the person that I am? What differentiates me from others? If someone else existed who was exactly like me, would we share the same identity? Contemporary philosophers of self have spent considerable time on another question. The reidentification question, as posed by Derek Parfit and others, seeks criteria that could explain what makes a person at two separate times the same person. Parfit's solution is the psychological continuity theory. Marya Schechtman argues—in part through the use of puzzle cases like transitivity and fission, but more forcefully in what she calls "the extreme claim"—that this approach does not do justice to the problem of personal identity, specifically by not adequately accounting for the relationship between personal identity and intuitively basic features of personal existence, or what Schechtman calls "the four features" (Schechtman 1996, p. 2). As an alternative, Schechtman offers the characterization question, which asks what psychological features make someone the person she is. Schechtman's answer to this question is what she calls "the narrative self-constitution view" (Schechtman 1996, pp. 1-2).

In this paper, I argue that Schechtman's four features hint at, but do not capture, what matters most in identity: namely, the first-person "mineness" of experience. Further, this mineness is not addressed by the characterization question, and so not properly accounted for in the narrative self-constitution view. I also argue that certain metaphysical commitments outside of the problem of personal identity invite the question of whether or not it is possible

for someone to be identical at t_1 and t_2, and so make the reidentification question necessary. Once mineness has been introduced to the reidentification question and the psychological continuation theory, I will re-evaluate Schechtman's arguments for the transitivity problem, fission cases, and the extreme claim. I conclude that transitivity and fission cases may or may not pose a problem for the psychological continuity theory, depending on those same external metaphysical commitments: specifically, it depends on whether or not one holds a four-dimensionalist view of time, or whether or not one holds a reductionist view of persons. As for the extreme claim, it has no bearing on a psychological continuation theory that is based on a concept of mineness.

However, it will first be necessary to spell out Schechtman's argument in greater detail. Schechtman claims that most analytic philosophers do not appreciate the complexity of the problem of personal identity; there is not one, monolithic identity question, but instead a multitude of questions whose distinct answers illuminate different aspects of personal identity. The identity question that most analytic philosophers treat as *the* question of personal identity is the reidentification question, or what makes a person at t_1 identical to a person at t_2. The answer to this question takes the form of a reidentification criterion. The problem arises when these philosophers recognize the great practical importance of personal identity, as captured in "… four basic features of personal existence: survival, moral responsibility, self-interested concern, and compensation," or what I will simply call the four features (Schechtman 1996, p. 2). These philosophers assume that their reidentification criterion must capture the link between the reidentification question and the four features, and it is here that they "run into trouble" (Schechtman 1996, p. 2).

There are two basic answers to the reidentification question: the bodily continuity theory and the psychological continuity theory (Schechtman 1996, p. 13). In the former, a person is identical at t_1 and t_2 if we can trace the ancestral relation of their body over time, or if there is qualitative similarity of the body at both t_1 and t_2. This is the standard we use in everyday encounters.

Schechtman believes it is the right answer to the reidentification question, and worth keeping solely for its evidentiary application. We assume the bodily continuation theory when, for example, we want to make sure that the body of the man in the police line-up is the same body that was seen robbing a liquor store. In the psychological continuity theory, a person is identical at t_1 and t_2 if we can trace the continuation of a "single psychological life," or there is a qualitative similarity of the mind at both t_1 and t_2 (Schechtman 1996, p. 13). The psychological continuity theory has gained acceptance, Schechtman suggests, because of its appeal to our intuitions about the four features (Schechtman 1996, p. 17). For example, it would seem strange to say that someone had survived an accident that altered or erased her entire psychological life, regardless of whether or not the body remains alive and intact.

However, for the psychological continuation theory to serve as an answer to the reidentification question, it must "have the logical form of an identity relation," and Schechtman argues that it does not (Schechtman 1996, p. 26). Schechtman shows this in two ways: the transitivity problem and with puzzle cases like fission or duplication. The transitivity problem is demonstrated in Reid's "Brave Officer" thought experiment, which was meant to show a flaw in what he took to be Locke's memory theory of personal identity. Reid asks us to picture a young boy who was flogged for stealing from an orchard. The boy grows into an officer who remembers the incident, but the officer eventually becomes a general who does not. We are asked to consider whether or not the young boy is identical to the general. On the view ascribed to Locke, they are not the same person (Schechtman 1996, pp. 27-28). Put another way, any identity relation should be transitive, so that if A is identical to B, and B is identical to C, then A will be identical to C. In this example, A is identical to B, and B to C, but A is not identical to C. In the fission case, we are asked to imagine someone who has, whether by brain-splitting surgery or some other science fiction scenario, been split or duplicated. Where once there was a single person, now there are two. This violates our understanding of identity relations, because in such a

situation, A would be identical to both B and C, but B and C would not be identical to one another (Schechtman 1996, pp. 27-28).

There is another problem of logical form, which Schechtman calls "the problem of determinacy and degree". This is related to the extreme claim. Personal identity must be all or none, so that A either is, or is not, identical to B. The psychological continuation theory, however, can admit various degrees (Schechtman 1996, pp. 42-43). This is because the reidentification question demands that its solution offer a criterion for determining the degree of similarity between two states, t_1 and t_2. There are two basic solutions to the problem of logical form. One solution is to view persons as four-dimensional objects, and the other is to adopt a reductionist view of person (Schechtman 1996, pp. 49-50). The four-dimensionalist view takes people not to be objects that endure over time, but instead people exist only over their entire duration; a person is not entirely present at any single point in her life (Schechtman 1996, pp. 11-12). In this way the states at t_1 and t_2 do not have to be identical, but instead they can be related in such a way as to show that they are parts of the same person (Schechtman 1996, pp. 38-39). The reductionist view also states that a person is divisible into parts, which could be conscious episodes or the infinitesimally-small time-slices of the four-dimensionalist view. What distinguishes reductionism is that these parts are ontologically more basic than the whole person, which is itself a construct. This reductionism is said to be "eliminative" because it effectively destroys the concept of the person as we understand it (Parfit 1995, p. 656). By way of example, Parfit suggests that we might have an eliminative reductionist view of nations, where nations do not have any kind of independent existence, but instead are constructions made up of people living in certain ways on some territory (Parfit 1995, p. 656). One need not affirm both four-dimensionalism and reductionism, though Schechtman does suspect that the former entails the latter (Schechtman 1996, p. 12).

The extreme claim follows from these two solutions. The first premise of the extreme claim is, like the problem of determinacy and degree, that personal identity requires numerical

identity, that A must be identical to B, not merely qualitatively similar (Schechtman 1996, p. 52). Schechtman makes this demand because of the four features. For example, if I am told that tomorrow I will face execution, I am concerned with my own survival, not whether someone *like* me will survive. The second premise is that the psychological continuation theory recognizes no distinction between identity and qualitative similarity (Schechtman 1996, p. 53). Given the reductionist view of persons, even if the time-slices at t_1 and t_2 belong to the same construct of a person, they are merely qualitatively similar and not identical. As Schechtman puts it, "…all it is for some future person to be me is for that person to have a psychological life qualitatively like mine…." (Schechtman 1996, p. 54). The extreme claim is meant to show that the psychological continuation theory cannot account for the importance of personal identity, as it cannot adequately make room for our intuitions about the four features. This is especially problematic because Schechtman claims it was our intuitions about the four features that motivated us to choose the psychological continuation theory over the bodily continuation theory. This, Schechtman argues, is the inevitable conclusion of our reasoning about the reidentification question.

As a means to better accommodate the four features, Schechtman offers the characterization question, which asks what "…beliefs, values, desires, and other psychological features make someone the person she is" (Schechtman 1996, p. 2). Another way that Schechtman puts it is that "characterization theorists ask what it means to say that a particular characteristic is that of a given person" (Schechtman 1996, p. 73). Schechtman is not concerned with "identity" in the sense of an identity relation, as it is treated in the reidentification question, but instead in the sense of an "identity crisis," where a person "…is unsure about what those defining features are, and so is unsure of his identity" (Schechtman 1996, p. 74). To answer the characterization question, one would need to provide a criterion for explaining which characteristics are merely a part of a person's history, and which ones are constitutive of someone's identity (Schechtman 1996, p. 77). This criterion will

need "…to define the relation that holds between a *person* and *particular actions*, *experiences*, *or characteristics* that are hers" (Schechtman 1996, p. 77). Since this relationship holds between a person and characteristics, it is not a comparison between two like terms, and so, unlike the reidentification question, there is no demand that the answer be an all-or-none relationship or non-transitive (Schechtman 1996, p. 79).

Schechtman's answer to the characterization question is the narrative self-constitution view, in which a person's identity is grounded in the formation of a coherent narrative that can be articulated and is consistent with reality. This narrative is self-constituting in that a person has the agency to create who she is by choosing which facts about her history are relevant to the narrative, and Schechtman suggests that organizing information in this way is what makes a human being a person at all (Schechtman 1996, pp. 94-95). This view acknowledges temporality, in that a person may choose which characterizations from the past, or even anticipated characterizations in the future, can be made central to a person's identity. But it is a different kind of temporality than that of the reidentification question: Schechtman emphasizes that individual time-slices cannot be addressed outside of the context of the narrative, or the person as a whole (Schechtman 1996, p. 97).

While Schechtman offers the characterization question and the narrative self-constitution view as a means to better explain our intuitions about the four features, it is my contention that intuitions about the four features will lead us away from the characterization question before the narrative self-constitution view ought to be considered an answer at all. This is because the four features are not in themselves what matter in identity, but instead point us to what matters: the first-person mineness of experience. Thomas Nagel gets at this quality in *The Objective Self*. Nagel asks the reader to imagine the world from an objective point of view; we can imagine a world composed of all people and all the facts about those people, and in spite of the apparent completeness of this description, it would seem to leave something out: namely, which one of those people I am (Nagel 1989, pp. 54-55). Further,

the truth and force of this insight can only be experienced from within the first-person perspective. If we attempt to formulate an objective picture of the world that leaves room for the relationship of conscious experience to individual bodies, we will still omit the fact that I—the one thinking about the world—am a particular person in that world (Nagel 1989, p. 56). Nagel calls this the "ineliminability of indexicals from a complete conception of the world" (Nagel 1989, p. 57). "I am Peter" is a valid identity statement because the first-person experiences of Peter are *experienced by me*. They have that quality of mineness.

It is not difficult to see how the four features each point us to this quality of mineness. Once again, the four features are moral responsibility, compensation, survival, and self-interested concern. In the case of moral responsibility, my guilt or satisfaction that results from an act are directly and subjectively enjoyed by me in a way not experienced by others. This is because there can only be moral responsibility with moral agency, and my agency has that quality of mineness perhaps more than anything, because it relates to actions that originate in my intentions. To illustrate, it would be paradoxical to say that I feel both guilty about, and not responsible for, some act. Similarly, in the case of compensation, I want to make sure that the labor with which I had a direct and possibly unpleasant experience will result in rewards that will be *experienced by me*. Mineness is very clear in the case of survival: When I survive, it is the continuation of my first-person conscious experience—my future of pains or pleasures. If I die, it is the extinction of this. Last, the connection may be most obvious for self-interested concern. I care for myself in a way that is different than how I care for other people, because my experiences have a different quality for me than the experiences of others. It is helpful to consider Nagel's objective view in this case; when I consider the world of persons objectively, I might say that when someone is injured they will generally experience a pain. But it would be an important omission if I were not to mention that my way of knowing that Peter is in pain is entirely different from my way of drawing this conclusion for anyone else. Peter's pains are mine, so

I will naturally care about them in a different way.

Once the four features are discussed in this way, it is clear that the characterization question and the narrative self-constitution view are not adequate to explain what matters in identity. Even if Schechtman were to argue that an experience could only be made part of a narrative if it had the quality of mineness, the narrative itself would not be the basic feature of identity; the mineness would. Additionally, this quality of mineness does not seem to be at all related to the characterization question, which directs our attention toward beliefs, values, or desires. Mineness is none of those things, and it is only arguably a psychological feature. My identity consists in the capacity to feel desires or form beliefs, but those beliefs and desires are not themselves me. I shape them; they do not shape, or constitute, me.

Another way to illustrate this is a thought experiment: imagine there exists another person who has the exact same beliefs, values, desires, and psychological features as me. We could say that orbiting a distant star there is a planet exactly like Earth, and on it a person exactly like me. I will call him Peter-2. To put this in terms of Schechtman's narrative self-constitution view, Peter-2 has chosen the exact same events and characterizations from his history to constitute who he is as a person. We could distinguish Peter-2 from me with facts like his location in the universe as opposed to my exact location in the universe. But, like me, Peter-2 does not really know or care about where in the universe he is relative to other galaxies, so this fact has not made its way into his narrative, or mine. If we were to consider me and this Doppelgänger in Nagel's objective view, there would be no way to tell which one of these people I am. The objective view would tell us everything we needed to know about my self-constituting narrative, but my self-constituting narrative would not help anyone in determining my identity. Mineness is lacking.

I do not mean to suggest that the characterization question has no value. I would draw a distinction between identity and character, and I would suggest that the characterization question tells us about character and not identity. To illustrate this

distinction, it is possible that beliefs can change, memories can be forgotten, and values can be betrayed, while the first-person mineness of experience persists unaltered. If that is possible, then none of these features are essential to identity. It is also possible for two people to have the same beliefs and values—the same character—and not be identical. In this way, the characterization question could require an evidentiary standard comparable to the bodily continuation theory, and ask what beliefs, values, desires, and psychological features are essential to our character, which might be used to distinguish our psychological lives from those of others, while leaving issues of identity to another question. For example, what beliefs or values do I possess that are essential enough that, if changed, I would become unrecognizable as Peter to those who feel like they know me? Or, how would I have to change, or what actions would I have to perform, to become unrecognizable to myself? This question might bleed into questions about what it means to know someone or to know one's self.

It is also worth exploring the relationship between identity and character. I have shown that belief and values do not dictate identity, but it is accurate to say that identity, as defined in terms of first-person mineness, has some power to affect one's beliefs and values, and so to form character. Richard Moran, in *Authority and Estrangement*, draws a distinction between the theoretical stance, in which we take ourselves to be one observable object among many, similar to Nagel's objective view, and the deliberative stance, in which we are first person agents (Moran 2001, p. 58). In the deliberative stance, we do not discover our beliefs by looking at the contents of our minds, as we might in what Moran calls "the perceptual model," but instead we decide our beliefs—or values, judgments, etc.—by considering the objects of the world themselves (Moran 2001, pp. 58-59). This illustrates how two people could have an identical character but not the same identity. As a consideration of my own beliefs will lead to a consideration of the features of the world that led to those convictions, Moran writes "…it is quite a different matter to take one's own belief about something to be true and to take someone else's belief to be

true, even when these beliefs concern the very same proposition" (Moran 2001, p. 76).

While we are not given reasons why beliefs, values, desires, and psychological features should be a part of a question about personal identity, the quality of mineness previously discussed does give us reason to believe that the reidentification question is a valid one. The first-person mineness of experience has a definite temporal aspect, and the reidentification question explores the relationship between identity and temporality in a way that the characterization question does not. Nagel mentions that the nature of first-person indexicals is parallel to a problem about the identity of time; we could describe the world objectively using words like "previous" and "subsequent and it would appear to be complete, except that it lacks "identification of a particular time as the present" (Nagel 1989, p. 57). It is worth noting that the mineness of experience exists in the present in a way that it does not in the past or the future. We can find an element of this in Moran's discussion of the deliberative stance and Moore's paradox. Moore's paradox is the apparent contradiction in statements like, "P is true, but I don't believe it" (Moran 2001, p. 69). Moran notes that it is not a formal contradiction; it would not be logically improper to say of anyone else that something is the case but that they believed the contrary (Moran 2001, p. 69). If follows that this is a problem of the deliberative, or first-person stance, and not a problem of the theoretical. But it also has an undeniable temporal aspect. The paradox dissolves if we consider the past-tense, "P was true, but I did not believe it." Moore's paradox is a contradiction because in the deliberative position, which can only exist in the first-person present tense, our beliefs are up to us and our interpretation of the world; my perception and rational consideration of features of the world have that element of mineness in the present tense, and in no other.

That mineness exists only in the present tense takes some defending, because it may not be obvious. After all, memories seem to have some quality of mineness, even though the events to which they are tied may have passed long ago. But it is impor-

tant to consider that having a pain, right now, in the present, is very different from considering the memory of some pain. Even in the most unpleasant memory, the pain itself is no longer *felt* in the present. A commonsense interpretation is that the mental state that experienced the pain belongs to the same identity as the mental state that is remembering it, but it is also possible that these two mental states belong to separate identities. Memories themselves are only experienced—that is, accessed or reconstructed, depending on one's view—in the present tense. While a memory may point to some previous experience, that experience lacks an immediate kind of mineness. The act of remembering itself, however, is something that I experience now, and so it preserves the importance of mineness. We can reintroduce the reidentification question to this discussion by considering whether or not one who is remembering a past pain is identical to the one who experienced that pain. The answer to the question "was it me who experienced that pain?" depends partly on metaphysical commitments that rest outside of the problem of personal identity, namely: the answer depends on one's view of the ontological status of persons, as either reductive or not. But I also take the answer to depend on one's view of time; like Schechtman, I suspect that a four-dimensionalist view of time entails a reductionist view of persons.

In his influential essay, "The Unreality of Time," McTaggart refers to the view of time that runs from the past, up to the present, and into the future as the "A series" (McTaggart 1908, pp. 456-473). The A series of time is the view of time that includes the subjective present tense, and it is the commonsense view that we take to be real. We can also call this the "tensed" view of time. The "B series" view of time takes all temporal events to exist relative to one another, describable in terms like before, after, earlier, or later; it omits that there is a subjective present (McTaggart 1908, pp. 456-473). This would be the description of time that Nagel says exists "from no point of view within the world" (Nagel 1989, p. 57). The B series of time is also comparable to a four-dimensionalist view of time, where time is relative and there is no absolute present. Both the B series and a four-dimensionalist view

of time are "tenseless." However, this does not mean that the B series and four-dimensionalist view of time are interchangeable. The four-dimensionalist view is incompatible with the tensed view or A series, but McTaggart suggests that the A series and B series are not only compatible, but that we take them both for granted (McTaggart 1908, pp. 456-473). This is because the four-dimensionalist view denies the subjective present, while the B series merely omits it. For my purposes, I will contrast the tensed view with the four-dimensionalist view, since they are mutually exclusive.

For a person to be a four-dimensional object the person would have extension not only through space but through time; the limits of a person's temporal extension would be the beginning and the end of his or her existence. If we are committed to a tensed view of time, then my self-conscious awareness, the part of me that has experiences of mineness, will endure through time. My awareness will "move" from the past to the present. The answer to the reidentification question, on this view, would be that it is possible for someone to have the same identity at t_1 and t_2. However, if I adopt a four-dimensionalist view of time, then it does not make sense for this awareness to endure, because that would require some kind of avowal of a "flow" of time. Instead, we can say that this awareness perdures; a person has temporal parts that exist in a sequence and are causally-linked, thus producing a whole person. Since each individual moment has a quality of mineness, but each individual moment is itself not a complete person, it is clear how reductionism follows from four-dimensionalism. Part of our concept of a person includes the notion that a person will endure through time. If we replace endurance with perdurance, then we have eliminated the construct of an ontologically basic person; in other words, we have accepted reductionism.

On this view, our answer to the reidentification question would be that it is impossible for someone to have the same identity at t_1 and t_2. When we consider the nature of identity and the nature of time, the reidentification question is unavoidable. Additionally, the psychological continuation theory, as an answer to

the reidentification question, is unavoidable; the consideration of first-person mineness is more relevant to questions about the mind than to mere evidentiary questions about the body. It is also worth noting that Schechtman framed reductionism and four-dimensionalism as solutions to the problems of logical form that the psychological continuity theory faces. In Schechtman's argument, these solutions only create more problems by making the psychological continuity theory susceptible to the extreme claim. I would like to emphasize that reductionism and four-dimensionalism could be true *regardless* of one's views on identity. In fact, one's interpretation of the special theory of relativity, for example, might entail a belief in four-dimensionalism, and arguably reductionism. From these metaphysical commitments, the psychological continuity theory as an answer to the reidentification question follows.

To illustrate this, I will now reconsider Schechtman's objections to the psychological continuation theory: transitivity, fission, the admittance of degrees, and the extreme claim. In the case of the transitivity problem, on the tensed view, a single conscious life with the quality of mineness could move forward through time and enjoy a continued existence, preserving its identity, regardless of whether the features of its body or its memory have changed. It is likely that this is more what Locke meant in his discussion of consciousness than the memory interpretation given by Reid (Schechtman 1996, p. 108). On the four-dimensionalist view, to use the "Brave Officer" example, A would not be identical to B, and B not identical to C, so there is no transitivity claimed, and hence no violation of transitivity laws. This is similar to the four-dimensionalist solution to the fission case; since the mineness of experience is episodic, A would not be identical to either B or C, in the same way that A would not be identical to B if the situation were non-branching. On the tensed view, the fission case becomes more bothersome. It is difficult to describe what would happen. For fission cases, Parfit claims that there simply would be no solution: the claim that B or C is identical to A would be neither true nor false (Parfit 1995, p. 659). Parfit also claims that the question itself would be "empty," since we could know and understand

exactly what happened in the case even without being able to give an answer (Parfit 1995, p. 659).

Regardless of the answer provided in defense of the psychological continuation theory, Mark Reid shows that Schechtman's narrative self-constitution view is also susceptible to fission cases. Reid asks us to imagine a man, Joe, who has lived with his girlfriend Sarah for six years. Joe loves the outdoors, but he also loves Sarah. Joe then undergoes a procedure that splits him into Joe-A and Joe-B. Realizing that they both cannot stay with Sarah, Joe-A moves into the mountains and becomes a forest ranger. Joe-B stays with Sarah and is common-law married to her on their seventh year, a practice that is retroactive so that they are considered married even before Joe underwent the fission procedure. Later, when Joe-A and Joe-B are reunited, they have formed entirely different narratives which include different interpretations about the six years that Joe lived with Sarah before the split. Joe-B believes that in that time, he was married to Sarah and living in accordance with his dream of a life spent devoted to her, while Joe-A concludes that he was wasting time that could have been spent on what was always his true passion: living in nature. While Joe-A and Joe-B, in the course of their discussion, might come to some agreement, Reid's interpretation of Schechtman shows that the problem persists, as some of the information exchanged will be of the "mutually exclusive" variety (Reid 1997, p. 211). A move is open to both Schechtman and Parfit on this issue: in such science fiction examples, the thought experiment shifts the discussion into a world that is so different from the one in our everyday experiences that it is not possible to comment on it.

Setting aside the fission cases, it is clear that the problems of degree and the extreme claim do not apply to a psychological continuation theory that takes mineness to be the defining feature of identity. Identity on this view remains an all-or-none relationship. Either I am having an experience that contains the quality of mineness, or I am not. This version of the psychological continuation theory would not say that I am identical to someone who is mostly like me, unless this person also had the same subjective,

first-person experience at the same time. This unity of identity does not depend on one's commitment to a particular view of time. On the tensed view, while many of my psychological features may have changed, if it is the same "I" that is having experiences, then it is the same identity. For the four-dimensionalist view, no future Peter is identical to the Peter that is having experiences at this exact moment.

While the reidentification question has shown itself more useful than the characterization question for getting at the heart of personal identity, the characterization question is still valuable. It asks us what defines character and prompts questions about the relationship between character and identity. Just as Schechtman did not toss out the bodily continuation theory after it was shown to produce more trivial conclusions than the psychological continuation theory, I think there is still a place for the narrative self-constitution view. This view can tell us about the formation and coherence of someone's character. If we adhere to a four-dimensionalist or reductionist view of persons, the narrative self-constitution view can also tell us something about how the concept of a person that endures over time can be constructed from individual time-slices, or how such a construct could be granted value and meaning. Even if, on Parfit's example, nations are nothing more than groups of people living in certain ways on some territory, it does not follow that there is nothing to be learned from a discussion of the history and nature of nations.

Bibliography

McTaggart, John Ellis (1908) "The Unreality of Time," *Mind: A Quarterly Review of Psychology and Philosophy* 17, pp. 456-473 <http://www.ditext.com/mctaggart/time.html>

Moran, Richard (2001) *Authority and Estrangement* (Princeton, NJ: Princeton University Press)

Nagel, Thomas (1989) *The View from Nowhere* (New York: Oxford University Press)

Parfit, Derek (1984) *Reasons and Persons* (New York: Oxford University Press)

_____ (1995) "The Unimportance of Identity," in: D. Chalmers (Ed),
Philosophy of Mind: Classical and Contemporary Readings, pp. 655-661
(New York: Oxford University Press)

Reid, Mark (1997) "Narrative and Fission: A Review Essay of Marya
Schechtman's *On The Constitution of Selves*," *Philosophical Psychology*
10.2, p. 211 <http://search.ebs cohost.com>

Schechtman, Marya. (1996) *The Constitution of Selves* (Ithaca, New York:
Cornell University Press)

BEYOND CARE ETHICS

Jessica Gonzalez

Developments in moral theory have brought us from the extreme of deontology and utilitarianism to another extreme: care ethics. Care ethics, though it proposes to be an enlightened moral philosophy, does not escape the very same criticism it casts at former theories. Care ethicists reject these former theories because their lack of consideration for feminine virtues is deemed harmful to women. Using Anne Fausto-Sterling's viewpoint on gender, I will show that care ethics, as specifically presented by Nel Noddings, is actually harmful to moral philosophy, and specifically to the cause of those whom it attempts to represent. In order to discuss my particular criticism of care ethics, I will first give a brief overview of the movement from utilitarianism and deontology to care ethics. Next, I will set out Fausto-Sterling's view of gender and apply this view to moral theory. Subsequently, I will present Noddings' version of care ethics. Finally, I will analyze Noddings' theory from a Fausto-Sterling perspective.

THE EVOLUTION OF MORAL THEORY

Traditional moral philosophy consists of *utilitarianism* and *deontology*. Utilitarianism was developed by theorists such as John Stuart Mill, Jeremy Bentham and Henry Sidgwick. Classic utilitarianism is concerned *only* with the consequences of one's action. To put it simply, it is the belief in optimizing the greatest good for the greatest number. The only moral principle held is that one should act in a way to do the least harm to the least people. There is no intrinsic moral nature of any act (Sinnot-Armstrong 2006). The same act could be deemed good in one case and bad in another. For example, Joe beats Gary to a pulp and Gary dies on his way to the hospital. In order for a classic utilitarian to make a judgment of

whether this is an acceptable act or not, they would need to know only what the consequences were. In case one, Joe beats Gary to a pulp because he is upset that Gary stole his girlfriend. In case two, Joe beats Gary to a pulp because Ralph, a masked gunman, promised to kill a dozen innocent customers at the bank he is holding up if Joe does not beat Gary up. The utilitarian would deem Joe's action in the first case unacceptable because there was needless harm caused to a person. In the second case, the utilitarian would praise Joe's actions because he has saved twelve innocent people.

Deontology, promoted by Immanuel Kant, is based on the Categorical Imperative (CI). This theory is based on the premise that one should only act according to the maxim that one can will to be a universal law. The CI is imperative because it commands us to exercise our wills in a particular way. It is categorical because it applies to us unconditionally, simply because we possess rational wills (Johnson 2004). Unlike utilitarianism, deontology is more concerned with the quality of the action rather than its consequences. When acting, we must consider whether we are doing something which would be good as a general rule. Concern for our own ends should not play into the morality of our actions. In fact, one strict premise of deontology is that one should not use any person as a means to an end. Instead, we should treat others as ends themselves.

Utilitarianism and deontology are very different systems of ethics. What they have in common, though, is that they did not consider women to be capable of the same moral virtue as men.[1] While the spirit of the theories is not necessarily exclusionary, proponents of these traditions have historically written women out of them. It is for this reason that they are rejected by proponents of *feminine ethics*. Many traditional moral philosophers provided a way for women to be virtuous; however, they were not considered to be capable of the kind of virtue attainable by men. Early feminists, like Mary Wollstonecraft, contended that women are not capable of this virtue because they are not given the opportunity. Though Wollstonecraft challenged mainstream thought during her time, she did not seek to change the system. Rather,

her purpose was to include women in traditional ethics. Many modern feminists, however, condemn the entire traditional system because these moral theories are "deficient to the degree that they lack, ignore, trivialize, or demean those traits of personality to the Freudian notion that whereas men have a well-developed moral sense, women do not" (Tong 2006). Here we can see that it is not enough for this version of feminism to accept that women are just as capable of virtue as men. Instead, this feminist theory rejects the entire moral system of these traditional theories. This will be further explored when I introduce care ethics.

Rejecting the Gender Binary

Now that we have discussed basic moral theory and its rejection by some feminists, let us pause so that I may introduce a specific view of feminism, namely the view of Anne Fausto-Sterling. This background will be important to understanding the argument that feminine ethics, specifically care ethics, is harmful. In considering Fausto-Sterling's particular feminist position, it will be important to examine her background as well as her views on science and society.

Fausto-Sterling is a biologist and a feminist. As a scientist, she studies molecular biology, examining living beings "from the perspective of the molecules from which they are built" (Fausto-Sterling 2000, p. 5). As a feminist, she is a social activist, working on "traditional feminist issues such as shelters for battered women, reproductive rights, and equal access for women in the academy" (Fausto-Sterling 2000, p. 259). Her intent on combining these two worlds is to show that scientists *create truths about sexuality*.

There is a cycle between biology and society such that biology is influenced by social norms and incorporates them into its "science." In turn, this "science" is taken as fact by society. Fausto-Sterling uses this analysis when discussing the difference between *sex* and *gender*. The two terms have been used for ages by academics, scientists and even the general population to describe separate phenomena: one physical and one social.

However, as Fausto-Sterling suggests, *sex* and *gender* are not as separate as was once believed. *Sex*, an assumed biological assignment, is not an exclusively physical category. Fausto-Sterling uses the example of Maria Patiño, an Olympic athlete, who was barred from competing on Spain's Olympic team because DNA analysis showed that she had a Y chromosome. This was a complete surprise to Patiño, who had external female genitalia and had always believed herself to be female. This begs the question, what does it mean to be female? Fausto-Sterling answers, "What bodily signals and functions we define as male or female come already entangled in our ideas about gender" (Fausto-Sterling 2000, p .5). In other words, our ideas and expectations of *sex* are influenced by our ideas and expectations of *gender*.

Fausto-Sterling's main argument for rejecting the gender binary is the existence of intersexuals. An estimated 1.7% of all babies born are intersexed (Fausto-Sterling 2000, p. 51). Though this may seem a minute figure, it is higher than the rate of more recognized phenomena, such as albinism. Intersexuality has several variations, including *pseudo-hermaphrodites* and *true hermaphrodites* (Fausto-Sterling 2000, p. 38). These variations result in people having chromosomes that are "opposite" of their genitalia, as well as "mixed" genitalia and "both" genitalia. Fausto-Sterling's analysis suggests that biological *sex/gender*[2] does not separate itself clearly into "male" and "female." Instead, biology operates on a continuum.

The most tragic consequence of this adherence to a fictional binary is the practice of infant sex assignment. If an infant is born with ambiguous genitalia in the US or Western Europe, a state of medical emergency is declared. Parents must work with doctors so that within twenty-four hours, the infant will leave the hospital "as a sex" (Fausto-Sterling 2000, p. 45). Many times, this results in a child being raised as one *sex/gender* then growing up to identify with the other. Parents are rushed into a decision that will change their child's life and body forever, and in a major way. This is perhaps the most severe result of the refusal of science to admit a *sex/gender* continuum and is one of the driving forces

behind Fausto-Sterling's push for change.

Biology has forced *sex/gender* into a binary rather than formulating a theory to reflect its continuous nature because biology is greatly influenced by society. This scientific acceptance of the binary is mostly a result of political and religious motives. From property rights to marriage laws, there are important issues at stake in this debate. Imagine a society in which a *sex/gender* binary did not exist. How could marriage laws discriminate against two males getting married? It could not because the term *male* would be nonsense.

It is important to consider the *progressive nightmare*[3] when discussing the implications of the rejection of the *sex/gender* binary. Fausto-Sterling clarifies her goal when she writes:

> I imagine a future in which our knowledge of the body has lead to resistance against medical surveillance, in which medical science has been placed at the service of gender variability, and genders have multiplied beyond currently fathomable limits… Ultimately, perhaps, concepts of masculinity and femininity might overlap so completely as to render the very notion of gender difference irrelevant. (Fausto-Sterling 2000, p. 101)

Clearly, Fausto-Sterling is not advocating that no one ever mention *sex/gender* again. If this were the case, it would be a disaster for all movements that have fought for the rights of intersexuals, lesbians, gays, bisexuals, and transgenderists. As a feminist, Fausto-Sterling is very aware of the need for such organizations. The revolution comes not in denying *sex/gender* but in understanding its lack of validity as a binary. This is a matter of education and is therefore not a regression but rather a progression into a world that has moved beyond the *sex/gender* binary.

APPLYING FAUSTO-STERLING TO MORAL THEORY

It is important here to point out that in applying Fausto-Sterling's perspective to moral theory, there will be much in common with

traditional modern feminist moral theory. As mentioned several times thus far, Fausto-Sterling is a feminist. Given her argument, it should be questioned whether this terminology—*feminist*—contradicts her entire notion of rejecting the *sex/gender* binary. Perhaps an argument could be made that a new term is needed. However, if one defines feminism as simply the belief and fight for equality regardless of sex/gender or sexual orientation, then this is indeed feminism, at its very core.

Fausto-Sterling paints the feminist picture when she writes that it "foresees a world of total equality…. Men and women would be represented equally, according to their abilities, in all walks of life" (Fausto-Sterling 1992, p. 207). She contrasts this perspective with the biological picture filled with gender myths. This picture describes women as naturally better mothers, whose "lack of aggressive drive ensures that they will always learn less" (Fausto-Sterling 1992, p. 207). Men, on the other hand, are described as being genetically predisposed to aggressive behavior and therefore naturally better in the work-place. For Fausto-Sterling, these two worlds are contradictory. We cannot keep our gender stereotypes *and* call ourselves feminists.

As the saying often goes, *stereotypes come from somewhere!* So why is it that we have these stereotypes to begin with? When answering a question like this, the decision is usually that it is either *nature or nurture*. Fausto-Sterling, (as well as some other scientists and social theorists) "no longer believes in the scientific validity of this framework" (Fausto-Sterling 1992, p. 7). This notion suggests that the answer to whether certain phenomena are a result of nature or nurture is not simply one or the other. Often times, it is a result of both. The point is that each case needs to be studied in a purely scientific manner, without bias. This is not contradictory to her rejection of the *sex/gender* binary; rather it explains that there is variation among all people on all points of the *sex/gender* continuum.

Since Fausto-Sterling is not a moral theorist, I will use an argument she does set forth to make a claim about what a basis for a moral theory in her style would be. In her discussion of the

history of the two-sex system, she brings up the transsexualism movement. Fully emerging as a type of person in the twentieth century, transsexuals are "individuals who have been born with 'good' male or 'good' female bodies. Psychologically, however, they envision themselves as members of the 'opposite' sex" (Fausto-Sterling 2000, p. 107). Many feel so strongly about their psychological gender that they seek medical aid to transform their body to match it. Fausto-Sterling points out that this movement, though seemingly progressive, in fact came at the price of reinforcing the *sex/gender* binary. Instead of challenging what it means to be *male* or *female*, transsexuals accept the dichotomy and just choose to switch sides. Society is left believing in this binary even more.

Fausto-Sterling then goes on to discuss the transgender movement. Transgenderists accept "kinship among those with gender-variant identities. Transgenderism supplants the dichotomy of transsexual and transvestite with a concept of continuity" (Fausto-Sterling 2000, p. 107). Here, we see that transgenderism is the preferred alternative since it allows for various *sex/gender* identities. More people are accepted with transgenderism because they are not forced to pick one extreme or another. It is comprehensive of all people.

A moral theory based on Fausto-Sterling's perspective would be inclusive of *all* people. It is my contention that a moral theory *should not* be exclusive, simply by definition. In comprising a moral theory, one should be determining what things *are* valued instead of what things *should* be valued. If we are only talking about what should be valued and not considering what people already value, we might as well be making a moral theory for Martians. This is not to say that a goal cannot be set. However, this goal should be something that is attainable by all people—not just some. It would not make sense, for example, to consider a moral theory that only applied to racially defined whites.[4] Such theories have only been able to do so by attacking the humanity of those excluded. Hopefully, we have at least passed this barbaric thinking. I do believe there is a general consensus that all humans

are equal in their humanity. Thus, any moral theory cannot be exclusionary without returning to this archaic way of thought.

In this sense, Fausto-Sterling would reject traditional moral theories because of their exclusion of women from true virtue. On the other hand, there would be no apparent reason for her to reject ideas like those set forth by Mary Wollstonecraft, that women are just as capable (if you give them the chance) as men in attaining virtue, in the traditional sense. The basic idea is that a moral theory that could be derived from Fausto-Sterling's perspective would necessarily be inclusive of all people, at all points of the *sex/gender* continuum. Relating back to the transgender discussion, it would *not* be enough to say that men who can think like women can be considered moral. This is how care ethics follows the same path as traditional moral theories. To say that there are exceptions where a man can think like a woman and therefore be moral, is no different from saying that a man can have the psychological identity of a woman, and therefore be a woman. Both these cases bring harm by reinforcing a dichotomy. This will be further explored once care ethics has been presented.

CARING AS A MORAL THEORY

Care ethics is a result of a feminine approach to moral theory. As mentioned earlier, this theory rejects traditional moral theories like utilitarianism and deontology. An entirely new foundation for morality is developed: care. This moral theory of care looks to relationships rather than consequences or a Categorical Imperative to guide action. I will specifically discuss the system of care ethics given by Nel Noddings. Noddings considers an ethic built on caring to be "essentially female" (Noddings 1984, p. 8). She believes that such an ethic arises out of one's experience as a woman. Traditional approaches to ethics are considered by her to be inspired by a masculine experience.

Noddings defines *care* as burdens or worries, as a stir of desire or inclination toward someone, or as having regard for someone's views and interests (Noddings 1984, p. 9). There are

three different ways of caring: the one-caring, the cared-for, and aesthetical caring. Noddings points out that "as I think about how I feel when I care, about what my frame of mind is, I see that my caring is always characterized by a move away from the self" (Noddings 1984, p. 16). Here, she is making the point that in a caring relationship, when an individual is the one-caring, he/she is moving away from oneself. Caring for others can happen in various intensities depending on the situation. This position requires that one exhibit *engrossment*. In order to illustrate what she means by engrossment, Noddings gives the example of a lover who has learned his beloved is ill. He cannot be at her bedside because they both are married to other people. However, he is in a mental state of engrossment because he feels the "deepest regard and, charged by his love with the duty to protect, he denies his own need in order to spare her one form of pain" (Noddings 1984, p. 10). The lover is selflessly putting the needs of his beloved over his own, understanding her situation first and foremost. Here we can see that the relationship is the basis for action. Instead of acting based on the consequences or on a maxim that should be some kind of universal law, the lover is considering the situation of his beloved and acting based on his engrossment.

The one cared-for is, above all, affected by the attitude of the one-caring (Noddings 1984, p. 20). If the one-caring seems inauthentic, the one cared-for is usually resentful. Ideally, the one-cared for and the one-caring will have reciprocal relationships. That is, both will care and be cared for. However, not all caring relationships are equal. Two examples of unequal relationships are teacher-student relationships and parent-child relationships. In a teacher-student relationship, it is important that the relationship stay unequal. This is because if the student were to attempt *inclusion* with respect to the teacher—that is, if the student were to attempt to discern the teacher's motives or think about what she is trying to accomplish—the student would be disadvantaged. In his student role, he must uphold his duty to be receptive to learning from the teacher. Once he crosses over into inclusion, he is disadvantaging both himself and his teacher for not letting her

teach. In parent-child relationships, the parent is responsible for keeping her side of the relationship. If, for example, a parent has one child who contributes to the caring relation by being receptive to her caring and by being responsive and caring as well, the parent must not favor this relationship over her other child, who is non-receptive and non-responsive. Noddings comments, "To demand such responsiveness is both futile and inconsistent with caring" (Noddings 1984, p. 72). Thus, the parent's role is to give both children equal caring so as to not give one an advantage over the other.

Noddings defines aesthetical caring as "caring about things and ideas" (Noddings 1984, p. 21). Like the simple receptivity of caring, there is no movement toward others that occurs in this instance. If one is engrossed in one's lawn or kitchen, there is no moving away from the self to another person. In fact, being engrossed in things and ideas can prevent one from taking on the role of the one-caring in other relationships. One is putting all one's energy into something that cannot receive or respond to one's caring.

MOVING BEYOND CARE ETHICS

In this section, I will give a Fausto-Sterling inspired criticism of Noddings' care ethics as well as point out ways in which the two perspectives could combine to form a new foundation for moral theory. The first point of criticism is that Noddings starts with the basic assumption that men and women think differently. As early as her book's introduction, she differentiates between the approach of the mother and the approach of the father. The approach through law and principle is the role of the "detached" father, while the role of the mother is rooted in receptivity, relatedness, and responsiveness (Noddings 1984, p. 2). She contends that men and women enter the domain of morality through different doors. Fausto-Sterling would clearly not be supportive of this claim. Even though Noddings concedes that it is possible for some men to embrace the feminine view and vice versa, this is *not*

enough by Fausto-Sterling's standards. Simply the language used implies that there are only two ways of thought and that these two ways of thought correlate to one's gender. This view suggests that one must either adhere to one's own gendered way of thought or cross over to adhere to the opposite view. Though there is perhaps less antagonism against this "reversed" identity than there is with transsexualism, the point is that both are seen as anomalies. This only reinforces a dichotomy between male and female.

Fausto-Sterling would also object to Noddings' oversimplification of the male and female mindsets. Noddings describes women as seeking to "'fill out' hypothetical situations in a defensible move toward concretization" (Noddings 1984, p. 36). Here, she is assuming that the feminine mind does not think deductively and rather focuses on hypothetical situations in order to decipher a moral dilemma. In a moral dilemma, Noddings suggests a "father might sacrifice his own child in fulfilling a principle; the mother might sacrifice any principle to preserve her child" (Noddings 1984, p. 37). Though she calls her own description too simplistic in this case, she also defends it by calling it *indicative* and *instructive*. These generalizations are not fair to either men or women. Even allowing for anomalies, this implies that the majority of men and women act in these respective ways. Not only is Noddings choosing not to criticize stereotypes, as Fausto-Sterling would encourage and perhaps even require of a moral theory, but Noddings is further perpetuating stereotypes.

It seems that in her theory, Noddings is giving women another role: martyr. This is a point which not only Fausto-Sterling but most feminists would criticize. We have established that Noddings is describing the female mindset when discussing the caregiver. The one-caring is clearly based on her stereotype of women. Even if she did use a male example and is suggesting that men take on this role as well, it is apparent that Noddings believes women will more than likely fulfill this role most of the time. There is no clear answer to how women will keep the ground they have gained politically or socially if they are put back in the role of the selfless caregiver.

One area that could foster a *relationship* between the Fausto-Sterling view and the Noddings view is Noddings' inclusion of thinking in the particulars. When she discusses right and wrong as concepts, she calls them "hard lessons—not swiftly accomplished by setting up as an objective the learning of some principle" (Noddings 1984, p. 93). Rather, we take a case by case basis to decide whether it was right or wrong to steal, lie, cheat, kill, etc. She argues that instead of giving universal principles, women give reasons for their acts that point to feelings, needs, conditions, etc. My question is, if Noddings can think in the particulars when considering acts, why is she so bent on universalizing gender-based morality? If she is willing to determine whether an act is right or wrong on a case-by-case basis, it is not a stretch to go on to determine whether a person is exhibiting a care-based mindset or virtue-based mindset on a case-by-case basis. With some adjustments to stereotypes and expectations, there could be some room here for common ground.

Perhaps care ethics is not a lost cause. There are definitely benefits to thinking in terms of relationships. In fact, the world could greatly benefit from just a little more care. However, in developing moral theories we must be careful that we do not move backwards in thought. Noddings' view, although there is a speck of light at the end of the tunnel, is very dark for feminists. Not only does it perpetuate a harmful *sex/gender* binary, but it goes as far as to impede the women's movement altogether. If Fausto-Sterling were to develop a moral theory, it may very well take a small part of care ethics. It would no doubt, however, keep at the foremost priority, the rejection of the *sex/gender* binary so that it could guarantee the most equal and comprehensive foundation.

Notes

1. There are notable exceptions to this. For example, J.S. Mill was a utilitarian who believed firmly in women's equality. However, the objection care ethicists have toward virtue ethics is that there is a general lack of consideration for women.

2. Since I have already argued that the concepts sex and gender are greatly influenced by each other and therefore not easily separable entities, I will refer to them together as sex/gender from now on.

3. I am borrowing this term from Naomi Zack, as she uses it in reference to the non-existence of race. In her argument that race does not exist, she considers the progressive nightmare, that if we rid our rhetoric of the notion of race, certain groups that have a history of being discriminated against would lose any ground they had attained through programs like Affirmative Action. Zack, Naomi. Philosophy of Science and Race. New York: Routledge, 2002. p. 8

4. I use the term racially defined as not to promote racialism. I only bring up the idea of race because it has been a basis of exclusion throughout history and therefore can parallel the case of gender as a basis of exclusion.

Bibliography

Fausto-Sterling, Anne. (1992) *Myths of Gender: Biological Theories About Women and Men* (New York: BasicBooks)

_____. (2000) *Sexing the Body: Gender Politics and the Construction of Sexuality* (New York: BasicBooks)

Johnson, Robert. (2004) "Kant's Moral Philosophy." *Stanford Encyclopedia of Philosophy*. (26 February 2004). 16 April 2009. <http://plato.stanford.edu/entries/kant-moral/#CatHypImp>.

Noddings, Nel. (1984) *Caring: A Feminine Approach to Ethics and Moral Education* (Berkeley: University of California Press)

Sinnott-Armstrong. (2006) "Consequentialism." *Stanford Encyclopedia of Philosophy*. (9 February 2006.) 16 April 2009. <http://plato.stanford.edu/entries/consequentialism/#ClaUti>.

Tong, Rosemarie and Nancy Williams. (2006) "Feminist Ethics." *Stanford Encyclopedia of Philosophy.* (29 November 2006). 16 April 2009. <http://plato.stanford.edu/entries/feminism-ethics/>.

On Objectification:
Oppressive or Liberating?

Melissa Foote

In the article, "Objectification," Martha Nussbaum responds to Catharine MacKinnon's position regarding the dehumanization involved with objectification. For MacKinnon, objectification is always a negative matter which only contributes to the desecration of gender relationships. Nussbaum discusses the different ways that an individual can be objectified, and "under…(some) specifications, objectification has features that may be either good or bad, depending on the overall context" (Nussbaum 1995, p.251). She also aims to discover whether there is a relationship between objectification and "equality, respect, and consent" (Nussbaum 1995, pp. 214-215). I will analyze Nussbaum's criteria for determining whether objectification is positive/liberating or negative/oppressive,[1] and point out the inconsistencies and impediments of her viewpoint.

She lays out seven senses of objectification and notes that while the instantiation of any one of them qualifies the individual as objectified, often many are combined in the same instance. The first notion is *instrumentality* (Nussbaum 1995, p. 257).[2] In order for one to be objectified in this way, she must be treated as a tool for the objectifier's purpose. Second, is the *denial of autonomy*. Here we find the objectified treated as though she lacked self-sufficiency and self-determination. Third on her list is *inertness*, which occurs when the objectified is treated as though she were "lacking in agency, and perhaps also in activity." Fourth is *fungibility*, which takes the objectified to be transposable with other objects (of either the same or different types). Fifth is the notion of *violability*, in which the objectified is treated as permissible to break, in a way devoid of boundary-integrity. The sixth notion is

ownership, which is (straight-forwardly) when the objectified is treated as something that is owned, to be bought or sold, etc. The final sense of 'objectified' is *denial of subjectivity*. An individual is objectified in this way when her "experiences and feelings (if any) need not be taken into account" (Nussbaum 1995, p. 257). Due to the variety of instantiations of 'objectification,' Nussbaum believes that "On the whole, it seems to me that 'objectification' is a relatively loose cluster-term, for whose application we sometimes treat any one of these features as sufficient, though more often a plurality of features is present when the term is applied" (Nussbaum 1995, p.258).

One of the most frequently discussed kinds of objectification is the first concept: instrumentalization of another human being (Nussbaum 1995, p. 265). MacKinnon has spent much time criticizing instrumentalization in pornography.[3] The problem with instrumentalizing an individual is that it denies her the status of being an end in, and of, herself (Nussbaum 1995, p. 265). However, Nussbaum offers examples of circumstances under which we can imagine an individual being used as an instrument without being instrumentalized. For example, if someone lays her head on the stomach of her lover, she is using her lover as an instrument, namely: a pillow (Nussbaum 1995, p.265). But it does not seem right to say that the lover is instrumentalized in the manner we are discussing. From this idea she concludes that, "… what is problematic is not instrumentalization per se, but treating someone *primarily* or *merely* as an instrument" (Nussbaum 1995, p.265). Here we see the importance of the notion of contextual relevance that Nussbaum advocates.

In order to assess what does and does not constitute objectification, we need to take into account the entire context of the relationship in question. If my lover and I have a mutually respectful relationship where we consistently treat one another as ends and not as means, there is no need to assume that any objectification[4] is occurring should I use my lover's stomach as a pillow. Nussbaum acknowledges that MacKinnon and Dworkin sanction the importance of context in regard to objectification by maintaining

that "…we assess male-female relations in the light of the larger social context and history of female subordination, and insist on differentiating the meaning of objectification in these contexts from its meaning in either male-male or female-female relations" (Nussbaum 1995, p. 271).[5] However, Nussbaum finds that they do not go far enough because they neglect the necessary evaluation of the details of an individual's history and/or psychology as relevant in the notion of objectification. Nussbaum advocates the analysis and consideration of the "overall context of the human relationship in question" (Nussbaum 1995, p. 271). While assessing society at large is important, we also need to take the specific relationship into account.

She uses this strategy of focusing on context when analyzing six different literary examples, in order to show how instrumentalization is not the only notion of objectification that requires contextually relevant standards. I will analyze the three examples that best demonstrate the usefulness and applicability of her theory. The first example is from *The Rainbow*, a novel by D.H. Lawrence:

> His blood beat up in waves of desire. He wanted to come to her, to meet her. She was there, if he could reach her. The reality of her who was just beyond him absorbed him. Blind and destroyed, he pressed forward, nearer, nearer, to receive the consummation of himself, be received within the darkness which should swallow him and yield him up to himself. If he could come really within the blazing kernel of darkness, if really he could be destroyed, burnt away till he lit with her in one consummation, that were supreme, supreme (qtd. in Nussbaum 1995, p. 215).

Next is a very graphic passage from *Isabelle and Veronique: Four Months, Four Cities*, a book by Laurence St. Clair (a pseudonym of the philosopher James Hankinson):

> She even has a sheet over her body, draped and folded into her contours. She doesn't move. She might be dead, Macrae

thinks…. Suddenly a desire to violate tears through his body like an electric shock, six thousand volts of violence, sacrilege, the lust to desecrate, destroy. His thumbs unite between the crack of her ass, nails inwards, knuckle hard on knuckle, and plunge up to the palms into her. A submarine scream rises from the deep green of her dreaming, and she snaps towards waking, half-waking, half-dreaming with no sense of self… and a hard pain stabbing at her entrails…. Isabelle opens her eyes, still not knowing where or what or why, her face hammed up against the cracking plaster…as Macrae digs deeper dragging another scream from her viscera, and her jerking head cracks hard on the wall…it seems as if he would tear the flesh from her to absorb it, crush it, melt it into his own hands…. And Isabelle…hears a voice calling out "don't stop"… and she realizes with surprise that it is coming from her mouth, it is her lips that are moving, it is her voice (qtd. in Nussbaum 1995, pp. 215-216).

The final example I wish to discuss is simply a caption found underneath three pictures in the April 1995 issue of *Playboy Magazine*. The photographs are of Nicollette Sheridan playing tennis at the Chris Evert Pro-Celebrity Tennis Classic, wearing a short tennis skirt that reveals her black underwear. The caption boasts, "Why We Love Tennis" (qtd. in Nussbaum 1995, p. 216). All three examples are deemed to be instances of objectification. Nussbaum clarifies that in each of her examples, "a human being is being regarded and/or treated as an object, in the context of a sexual relationship" (Nussbaum 1995, p. 216).

The example that I found most interesting is the first passage, from D.H. Lawrence. In it we have two lovers who are "burnt away till he lit with her in one consummation" (qtd. in Nussbaum 1995, p. 252). Nussbaum argues that within this passage we have a denial of autonomy and dehumanization. However, the notions of objectification that she laid out earlier were happening simultaneously. Moreover, there was a *mutual* denial of autonomy by the lovers in the story. They reduced one another (and themselves)

to genital organs, but this does not necessarily mean humanity is denied in the manner that Kant suggests.[6] Rather, "…the objectification is symmetrical and mutual—and in both cases undertaken in a context of mutual respect and rough social equality" (Nussbaum 1995, p. 275). We should not only hold this to be a negative example of objectification, but we should hold it as a positive set of events. Here Nussbaum is surprising in her praise of mutual objectification:

> Lawrence shows how a kind of sexual objectification-not, certainly, a commercial sort, and one that is profoundly opposed to the commercialization of sex-can be a vehicle of autonomy and self-expression for women, how the very surrender of autonomy in a certain sort of sex act can free energies that can be used to make the self whole and full (Nussbaum 1995, p.277).

Objectification can here be seen as a positive component of the sexual behavior of two consenting lovers. I find this interesting because it is not at all clear how we are to determine under what circumstances both people are truly consenting as rational agents, agents who want to be viewed and treated as dehumanized sexual accessories. Also, it is not at all clear how we can rule out this being the case in many of Nussbaum's other literary examples of alleged oppressive objectification (which we will turn our attention to later in the paper). When discussing this Lawrentian objectification, Nussbaum further explains, "objectification is frequently connected with a certain type of reduction of persons to their bodily parts, and the attribution of a certain sort of independent agency to the bodily parts" (Nussbaum 1995, p. 274). Within the literary passage in question (as well as other works by Lawrence) we see both parties putting their individuality aside in order to reduce themselves and one-another down to their bodily organs. It is because of this mutual denial of autonomy that Nussbaum argues we should not view this type of objectification as a negative aspect of a relationship. In fact, she suggests that we view it as precisely the opposite, by stating:

> … to be identified with her genital organs is not necessarily
> to be seen as dehumanized meat ripe for victimization and
> abuse, but can be a way of being seen more fully as the
> human individual she is. It is a reminder that the genital
> organs of people are not really fungible, but have their own
> individual character, and are in effect parts of the person, if
> one will really look at them closely without shame (Nuss-
> baum 1995, p.276).

She writes this in direct objection to MacKinnon and Kant's notions of such a reduction. Kant holds that such a reduction of an individual to their sexual organs entails the complete disdain of that individual's personhood. Nussbaum holds that with Lawrence's romantic rhetoric we can see how objectification can be a positive aspect in a relationship, much in the way that Sunstein claimed "objectification might be a wonderful part of sexual life" (sum. in Nussbaum 1995, p. 274).

I would now like to turn to another example that Nussbaum discusses in her assessment of the ramifications of objectification in the media: *Playboy Magazine*. The specific example (though, we can clearly see she is speaking of the magazine much more generally) that she cites is that of an actress playing tennis in a short skirt which reveals her underwear, with the caption reading, "Why we love tennis" (qtd. in Nussbaum 1995, p. 216). Here Nussbaum sides with MacKinnon and Dworkin, who have argued that soft-core and hard-core pornography are both parts of a collectively objectifying industry. Nussbaum holds that the photograph and its "demeaning" caption are giving the message that "whatever else this woman is and does, for us she is an object for sexual enjoyment" (Nussbaum 1995, p. 283). She takes this even further, asserting that through a photo and caption such as this, the male reader[7] is the one who has "subjectivity and autonomy," while the woman in the picture is presented as being a "delicious piece of fruit, existing only or primarily to satisfy his desire" (Nussbaum 1995, p. 283). Nussbaum attacks the instance of objectification that we find here (and on a regular basis) in *Playboy*

Magazine, and ultimately concludes that it is a negative display of objectification and a bad influence on men (Nussbaum 1995, p. 286). She attacks the central motive of *Playboy*, saying that it "depicts a thoroughgoing fungibility and commodification of sex partners, and, in the process, severs sex from any deep connection with self-expression or emotion" (Nussbaum 1995, p. 283). The critique that Nussbaum gives within this section is, I argue, wholly unjustified. Moreover, even if her conclusion is right (that *Playboy* is not a substantiation of positive objectification), her reasoning is flawed and inconsistent with assertions made elsewhere in the paper.

Nussbaum goes as far as to assert that the author of another one of her literary examples, Hankinson, would be justified in attacking *Playboy* as being worse than his own erotic novels, which are offensive, perverse, and deeply disturbing in their violent sexual content. She defends that Hankinson's work, writing:

> [Hankinson] at least connects sexuality to the depths of people's dreams and wishes and thus avoids the reduction of bodies to interchangeable commodities, whereas in Playboy sex is a commodity, and women become very like cars, or suits, namely, expensive possessions that mark one's status in the world of men (Nussbaum 1995, pp. 283-284).

The first thing I would like to bring attention to is that this is a defense of the novel that discusses a woman experiencing "a hard pain stabbing at her entrails" as a man forces his thumbs into her rectum. Moreover, the man committing this disturbing act experiences "a desire to violate," a "lust to desecrate, destroy," and is acting out of a "violence born of desperation and desire" (qtd. in Nussbaum 1995, pp. 252-253). Does this passage (which is much more graphic and offensive than the small amount I have excerpted) truly connect sexuality "to the depths of people's dreams and wishes," as Nussbaum advocates in the quote above? Furthermore, does it successfully avoid "the reduction of bodies to interchangeable commodities"? It seems clear that the dreams and wishes being fulfilled here are those of Macrae. It is necessary to

61

address the fact that Isabelle is yelling, "don't stop," but we should not be too eager to accept this as her deriving the same pleasure that Macrae is. We know that she is experiencing deep pain, and she is surprised to discover that the voice is her own. It could be the case that she obtains pleasure from her pain in a sadomasochistic sense, but it is not at all clear from this passage that Isabelle experiences sexual liberation from what is being done to her by Macrae.[8] Objectification *must* be considered more dangerous and negative when physical violence and abuse is involved. It seems ridiculous to assert that a substantiation of this graphic scene is "better than" a photograph of a woman in a short skirt (but otherwise fully dressed) playing tennis! Let us remember that one of Nussbaum's goals was to establish whether there is a relationship between objectification and "equality, respect, and consent" (Nussbaum 1995, pp. 214-215). Are Macrae and Isabelle equal in this story? One is painfully violated while the other is asserting his power and need for domination over her. She is shoved against a wall, and is screaming out at what is happening to her. Does this demonstrate the respect that Macrae has for Isabelle? Finally, she is asleep when he comes to her, and is "half-waking, half-dreaming with no sense of self" as he fulfills his need to desecrate her and cause her pain (Nussbaum 1995, p. 215). Consent is normally not granted when one of the individuals is asleep, and begins crying out in pain as she begins to wake. If there is indeed a relationship between objectification and equality, respect, and consent, we do not find it in this literary example.

Let us again consult Nussbaum's lengthy quote that compares these two sources of objectification, in particular, the notion that women become very much like cars (or objects) in *Playboy*. What Nussbaum has failed to take into account throughout this discussion is the complete *willingness* of women to place themselves in *Playboy* to be viewed as sexual *objects*. The women who pose in *Playboy* are acting entirely out of their own will (and for many of them, it is a dream come true!), and they are eager to be the source of sexual desire for thousands of men worldwide. I am not suggesting that this is a healthy goal or desire for women to have,

nor am I asserting that the only thing that needs to be taken into account when discussing objectification is the willingness of the individual. One can think of countless examples of horrific sexual behaviors, that women are willingly involved in, that are motivated by fear, desperation, financial need, etc. It seems obvious that we can identify a source of negative objectification without needing consent from the individual(s) being objectified. However, in the case of *Playboy*, it is precisely this desire to be represented as a sexual *object* that motivates many of the models to participate in spreads. For many women, *Playboy* is a liberating experience that allows them to be comfortable in their own sexuality, and to share their beauty with the world.[9] Ultimately, the models (and the thousands of hopefuls who never make it into the magazine but would jump at the chance) find a sense of satisfaction in the reduction of themselves to their sexuality. It is not at all clear to me why Nussbaum so unequivocally affirms that Lawrence's romantic novel is a positive example of objectification because the woman who was (to reiterate) "identified with her genital organs is not necessarily to be seen as dehumanized meat ripe for victimization and abuse, but can be a way of being seen more fully as the human individual she is" (Nussbaum 1995, p. 276). Yet, in the case of *Playboy* she does not take this possibility into account. Presumably, Nussbaum would respond to this objection by affirming that in the case of Lawrence's book, it was positive simply because *both* sexual partners were making the conscientious decision to deny their own autonomy, and one another's. With *Playboy*, we do not have the contextual analysis of the relationship that Nussbaum advocates in order to determine the level of positivity/negativity we should attribute to the objectification at hand. One (obvious) reason for this is because it takes place outside of a relationship. It involves one woman and thousands (if not millions) of men who will never meet her or be in a relationship with her.

However, I think we can view the context surrounding the situation, in a way independent from one-on-one relationships, by turning our attention to *Playgirl*. Nussbaum does not take the contra into account when discussing the dangers of *Playboy*.

Playgirl Magazine features nude men, posing in sexual positions, complete with clever little captions that are meant to demonstrate that this hot hunk of man just might be the man who shows up to fix your sink, or help you fix your tire, etc. It does not present men as individuals (as opposed to merely sexual objects) anymore than *Playboy* does for women. Moreover, *Playgirl* is potentially more objectifying than *Playboy* in the following sense: It encourages women to have the fantasy that these men are average. One of the criticisms that Nussbaum has of *Playboy* is that it, "strongly suggests that real-life women relevantly similar to the tennis-player can easily be cast in the roles in which *Playboy* casts its chosen few. In that way it constructs for the reader a fantasy objectification of a class of real women" (Nussbaum 1995, p. 284). It is not at all clear how showing an attractive woman playing tennis in a short skirt in any way suggests that men can deny the autonomy of their co-workers or neighbors. In fact, most women in *Playboy* are shown to be almost goddess-like, having nothing at all in common with the wife/girlfriend of the reader (which I assume, is part of the reason it is so popular). The average *Playboy* subscriber is fully aware that 99% of the female population does not look or act like *Playboy* models. It is precisely this which makes the entire process a fantasy for men. However, one can clearly see how *Playgirl* aims to display its models as normal men who have normal jobs in a small town just like the one its reader lives in. While we cannot view the context of a relationship between the objectifier and the objectified in this case like Nussbaum advocates, we can analyze the relationship between the male and female forms of the same type of objectification. If one is going to attack *Playboy* in the way that Nussbaum does, it seems entirely unfair to completely ignore the same objectification instantiated in *Playgirl*. I think it would be fair to analyze the objectification as being positive (or at least, in no way worse than the Lawrence examples) for the fact that members of both sexes are making a conscious decision to reduce themselves to their sexual organs for the gratification of themselves and others, a point that Nussbaum fails to acknowledge.

Another problem that needs to be addressed is the form of example that Nussbaum uses in her analysis. With the exception of the *Playboy* photograph, all of her remaining examples are taken from novels. Literature should not be taken as a reflection of the real struggles and oppression that take place. While they may be interesting to read, and give us sexy examples that will provoke emotional responses in the reader, they are in no way a mirror image of the objectification that surrounds women in real relationships within our patriarchal society. One of the difficulties in trying to use these examples to support Nussbaum's analysis of objectification is that we are given a completely independent and detached view of the instantiation of objectification in literature. With Nussbaum stressing the importance of the context surrounding the relationship in question, her theory weakens greatly when we attempt to transition from literature to society. Within Lawrence's novel, we are given a thorough framework of the relationship between Brangwen and Lydia (who are husband and wife). As in most pieces of literature, we are given a general overview of the relationship. We are aware of their backgrounds, social standing, quality of marriage, and care for one another. We know that Brangwen and Lydia respect one another because Lawrence tells us that they do. While we may be able to look at their sexual encounter and declare it a positive example of objectification, this does little to strengthen Nussbaum's argument. It is entirely unrealistic to suppose that we can analyze the relationship between two adults in the way that we can evaluate and dissect the relationship between two characters in a novel. In real life, we do not know if the members of a sexually objectifying occurrence have "equality, respect, and consent" for one another (Nussbaum 1995, pp. 214-215). It seems fairly obvious that observing a couple from outside their relationship is deceiving and inaccurate. How many people have surprised their loved ones by announcing that they were getting a divorce, or that one physically abused the other? It is entirely unrealistic for Nussbaum to suppose that we can determine the context surrounding the members of a sexual encounter in order to determine if the objectification is positive or

negative. We do not possess the ability to truly know and understand the details of anyone's relationship. At least in the case of *Playboy*, we know that the models are not in a dangerous relationship with the men who sexually objectify them as they appear on the page. If Nussbaum wishes for her theory to be applied to real examples of sexual objectification, then her fictional examples do nothing to strengthen or demonstrate the applicability of her viewpoint.

I appreciate that Nussbaum attempts to appeal to our intuitions regarding objectification. It does not seem that using my lover's stomach as a pillow is objectification in the way that pornography seems to be. A more extensive definition of objectification, as well as the notion that it could potentially be a positive experience/situation, is one that seems necessary within the boundaries of such a complex concept. However, it is not at all clear how one is to go about substantiating such a principle, and there are many cases of objectification that do not seem to be obviously positive or negative in this way. Moreover, it is dangerous to place the final declaration of objectification into the hands of a part-time witness outside of the relationship or situation in question. One cannot truly know if the relationship involves mutual respect for the other person's autonomy by having a social friendship with one (or both) members of the relationship. It also seems equally problematic to rely upon the individuals in question who may not be acting in their own best interest. One can certainly think of an instance where a woman claims to be content, while being horribly and violently objectified, because she has never known anything better, or desperately needs money, etc. Certainly, more work needs to be completed on the subject, but I appreciate the interesting contribution that Nussbaum has advanced in this paper. Overall, the criteria for determining what we should classify as objectification seems effective, but the contextual relevance that Nussbaum advocates, in order to determine whether objectification is positive/liberating or negative/oppressive, is ultimately ineffective for a normative theory.

Notes

1. Throughout the paper I will be using these terms interchangeably.

2. All seven notions of objectivity taken from Nussbaum (1995, p. 257).

3. See MacKinnon (1987) for a detailed account of her theory and the utilization of the notion of instrumentalization.

4. This would certainly not be an example of negative/oppressive objectification that we are concerned with.

5. This is a summary of MacKinnon and Dworkin in Nussbaum's own words.

6. For a more detailed analysis on Nussbaum's objection to Kant in this sense, see Nussbaum (1995, p. 275).

7. She has not taken into account that some women enjoy *Playboy* magazine, but we all must agree that the demographic to which the magazine is (very successfully) marketing itself is male.

8. The full passage, quoted in Nussbaum (1995, pp. 215-216), does nothing more to illuminate the amount of pleasure/pain that Isabelle receives from Macrae's actions.

9. Nussbaum briefly addresses the notion that *Playboy* can be viewed as a feature of the women's liberation movement, but quickly dismisses it as a "profound betrayal of… the Kantian ideal of human regard" (1995, p. 283). I think her response is extremely aloof and unsatisfying given the extensive nature of her attack on Playboy.

Bibliography

MacKinnon, Catharine. (1987) *Feminism Unmodified* (Cambridge, MA: Harvard University Press)

Nussbaum, Martha C. (1995) "Objectification," in: *Philosophy and Public Affairs* Vol. 21 No. 2, pp. 249-291

LEVINAS AND RICOEUR:
"BEING-ABANDONED-TO-THE-OTHER"

Joshua Allen

INTRODUCTION

For both Levinas and Ricouer, moral agency and what constitutes the ethical life are issues central to their respective philosophical projects. In this paper, I examine to what extent their ethical views resemble one another, and how they ultimately diverge from each other. The similarities and dissimilarities between Levinas's and Ricoeur's ethical projects turn primarily on the ways in which each thinker theorizes the role of alterity in the ethical relation. So the focus of this paper will be on how the concept of "alterity" operates in Levinas and Ricoeur. Because this paper focuses on both the convergences and divergences between Levinas and Ricoeur, it is composed of two parts. (1) For the claim of convergence, I argue that Levinas and Ricoeur are most similar in how they think about moral agency; for both thinkers, to be an ethical subject is always to be involved in "an originary being-abandoned-to-the-other."[1] In other words, both Levinas and Ricoeur agree that *the very subjectivity of the moral agent is constituted by alterity*. (2) Despite this convergence in their theories of moral agency, Levinas and Ricoeur ultimately understand what it *means* to live the ethical life in radically different ways. According to Ricoeur, because Levinas's conception of alterity, as a necessary condition for any notion of ethics, is so radically beyond any comprehensible context, beyond "the language of relation," Levinas unwittingly renders any relation between the self and the Other impossible; thus, Levinas, in turn, makes impossible any ethical relationship at all. Yet, in his explication of the notion of "solicitude," Ricoeur presupposes precisely what Levinas denies, namely a reciprocal relation between the self and the Other based on mutuality, where

the desire for "the good life" is predicated on the need for Others. For Levinas, this suggests that Ricoeur's ethical project destroys the very thing needed for the ethical relation, the absolute alterity of the Other. Therefore, I conclude that, due to the ways in which Levinas and Ricoeur understand the role of alterity for the ethical relation, they may ultimately be talking at cross-purposes.

Although a number of comparisons between Levinas and Ricoeur have been made by contemporary commentators, none to my knowledge have attempted to show that both Levinas's and Ricoeur's formulations of the ethical subject involve what I am calling, following Bernhard Waldenfels, "an-originary-being-abandoned-to-the-other." Other commentators, particularly Richard Cohen and Patrick Bourgeois, have provided insightful analyses of the complementary aspects as well as antagonisms between Levinas and Ricoeur, but neither of them explicitly claim that, due to the fact that Levinas's ethics denies what Ricoeur's presupposes, Levinas and Ricoeur seem to have fundamentally misunderstood each other.

METAPHYSICAL DESIRE AND THE ABSOLUTE ALTERITY OF THE OTHER

In Levinasian terminology, the "Other" connotes a kaleidoscope of related terms—"alterity," "exteriority," "excess," "the Infinite," "The Good," etc—all of which are supposed to invoke the idea of "beyond" and/or "disrupting" the totality of Being (ontology). In the history of Western philosophy there are two major figures who, in their own ways, have approximated this notion of the "Other" as absolute alterity, transcendence, or beyond Being: Descartes, with his idea of the Infinite, and Plato, with his idea of the good beyond Being. Built into Levinas's notion of the "Other" are both the Cartesian and Platonic models of the "more" than Being.

In the *Third Mediation*, Descartes conceives of the idea of the Infinite as that which cannot originate in a finite (imperfect) mind, nor can its *ideatum* be contained by the idea itself. Unlike the correspondence that transpires between the idea and

its *ideatum* in cognition, the Infinite "overflows" its idea. In principle, the thought that "thinks" the Infinite is unable to acquire its *ideatum*, and thus fails to establish a bi-polar relation between the "I" and the Other (Levinas 1969: 48-49). Properly speaking, then, there is no adequate "relation" to the Infinite; the idea of the Infinite is *non-adequation*, indeterminacy par excellence. This indeterminate distance that opens up between, and separates, *ideatum* and idea, "constitutes the content of the ideatum itself," which is precisely the Infinite (Levinas 1969: 49). The positive "content"[2] of the Cartesian idea of the Infinite is characterized, then, by transcendence, by the thought surpassing itself in thinking of that which remains absolutely exterior: the Other.

The germ of what Levinas calls "metaphysical desire for the absolutely Other" is also found in the structure of the Platonic Good. In the Platonic idea of the Good, Levinas finds evidence of a kind of thought that already "thinks" within the ethical. The Good is the "highest idea" which, posited beyond the totality of essences, engenders a kind of thinking that issues from a being already fulfilled, who lacks nothing, but, despite this fulfillment, moves "beyond" this self-satisfaction, this "plentitude," toward the transcendent. In this way, Levinas sees in the Platonic idea of the Good a "disinterested" desire for the "transcendent" not motivated by lack or need, which gives it its ethical structure.

Levinas adopts from both Descartes and Plato this notion of the desire for the absolutely Other, and calls it "metaphysical desire." Unlike ordinary desire, metaphysical desire is for Levinas distinctive because it's not motivated by *need*. Unlike *need*, metaphysical desire cannot be sated by acquiring the object of its desire—the Infinite. This is because the Infinite "exceeds" any idea that one may have of it.

According to Levinas, the significance of metaphysical desire hinges on the way in which it preserves alterity between the Same and the Other. For Levinas, alterity is a necessary condition for ethics, because it ensures that the Other's otherness is not assimilated into the Same, i.e. totalized.[3] In a very qualified sense, Levinas claims that the subject (the Same) and the Other

must paradoxically remain separated in their "relation" in order to guard against collapsing the Other's alterity into the totality of the Same. According to Levinas, the structure of the metaphysical desire for the absolutely Other is able to preserve alterity because it satisfies the condition of separation, which turns on two further sub-conditions for the ethical "relation": (1) Interiority, and (2) Exteriority.

SEPARATION: INTERIORITY OF THE SAME, EXTERIORITY OF THE OTHER

The metaphysical desire for the absolutely Other comes to pass in a subject who is already fully at home with him/herself, completely self-satisfied. In *Totality and Infinity*, Levinas's phenomenological descriptions of basic human existence as interiority, solitude, economy, and enjoyment constitute a primordial and eminent sense of subjectivity, one which draws attention to the substantiality of the subject who "lives from" what Levinas calls "the elemental"—food, other people, the environment in general. As "interiority," the subject lives within *the economy of need*, where the "I" enjoys its existence by virtue of its dependence on, and mastery of, the things from which it lives:

> The human being thrives on its needs; he is happy for his needs. The paradox of "living from something," or as Plato would say, the folly of these pleasures, is precisely in a complacency with regard to what life depends on—not a mastery on the one hand and a dependence on the other, but a mastery in this dependence. This is perhaps the very definition of complacency and pleasure. Living from … is the dependency that turns into sovereignty, into happiness—essentially egoist (Levinas 1969: 114).

This egoism, which thrives by appropriating or "internalizing" the Other, characterizes the "way" of the Same.

> If enjoyment is the very eddy of the same, it is not ignorance

but exploitation of the other. The alterity of the other the world is, is surmounted by need, which enjoyment remembers and is rekindled by; need is the primary movement of thesame (Levinas 1969: 115-16).

Happiness, as the enjoyment of "living from" others, is "exploitative" if self-maintenance is predicated on mastering one's dependence on the "nutriments" the other provides. Within this economy of interiority, as "independence through dependence," as "the transmutation of the other into the same," the subject's separation is consummated (Levinas 1969: 115). In this sense, "the absolutely intransitive element" of subjectivity is the fact of *existing*; it is principally the *egoism of life* that individuates the self from Others, for I cannot share my pains and pleasure with another. As Levinas says: "Subjectivity originates in the independence and sovereignty of enjoyment" (Levinas 1969: 114). Thus, for Levinas, happiness is essentially "a principle of individuation," which is necessary for the condition of separation: "In the happiness of enjoyment is enacted the individuation, the auto-personification, the substantialization, and the independence of the self […]" (Levinas 1969:147, *passage altered*).

Levinas emphasizes that a finite being must be wholly content in its interiority, satisfied with a kind of solipsistic enjoyment of life, in order for the idea of the Infinite to remain in its exteriority and be absolutely *transcendent*. If the desire for the absolutely Other issued from the will of the subject, that is, if the Other was anticipated by the subject's desire for it, then the Other's alterity would be reduced to the subject's intention, which would in turn destroy it (Levinas 1969: 35-6). In this sense, metaphysical desire cannot begin with the initiative of the subject, as an intentional act, rendering the Other an objective theme for consciousness. If it were to do so, the Other would be reduced to the subject's will and remain within the scope of totality. Understood in this sense, metaphysical desire desires the Infinity of the absolutely Other non-intentionally or "*passively*."

But how are we to understand this notion of an unintentional,

non-voluntary, and passive desire, one which does not issue from *need*? Regarding the issue of *passivity*, Levinas claims that metaphysical desire is actually initiated by the Other; in other words, metaphysical desire is first produced by experiencing the Other's Face![4] According to Levinas, one concretely experiences the Infinite in one's proximity to the Other person.[5] Levinas claims that one's *immanent* perception of the Other's face is not exhaustive of the Other as an individual, since the consciousness distinctive of any individual is not found in any perception of a person's body. So, for example, when I perceive a person's Face, the Face is *immanent* to my perception of it. However, what is indicated by the Face—an Other consciousness—remains transcendent to me. In this sense, the face always indicates "more" than what I can incorporate into my perceptual horizon.[6] So, because metaphysical desire is initiated by the expression of the Other as "Face," as the idea of Infinity or transcendence, the Other's alterity remains absolutely *exterior* to the subject's totalizing interiority. In its transcendence, the Other maintains an exteriority from the self-Same because it opens up an irreducible distance in "relation" to the one who desires it. Thanks to its absolute exteriority, the Other's

> transcendence designates a relation with a reality infinitely distant from my own reality, yet without this distance destroying this relation and without this relation destroying this distance, as would happen with relations within the same; this relation does not become an implantation in the other and a confusion with him, does not affect the very identity of the same […] (Levinas 1969: 41-2).

The *transcendent movement* of the metaphysical desire for the absolutely Other produces a "relation" that "does not affect the very identity of the same," and does not "confuse" the Same with the Other. Thus, in metaphysical desire the subject remains "closed" in upon itself in interiority, yet is, at the same time, somehow "open" to a "relation" with the exteriority of the absolutely Other (Levinas 1969: 148-49). But how exactly does this occur? How is it that the Same and the Other can be in "relation"

while remaining infinitely "separated?"

ALTERITY AND THE FACE OF THE OTHER

From our discussion above, we already know that Levinas models the ethical "relation" on both Descartes' idea of the Infinite and Plato's Good beyond Being. He calls this "relation" "trans-ascendence," which signifies "excess," "height," "absolute alterity," "separateness," "exteriority," and so on. All of these adjectives connote "the beyond" of transcendence—infinite distance. What Levinas wants to capture with such hyperbolic terminology is the idea that the Other is "more than" what thought can represent it as (Levinas 1969: 51). But, then, one might ask: If the Other surpasses all thought, exceeds all "relations" that one might establish with it, how would it be possible to experience or think the Other?[7] The answer, as we noted above, is found in the expression of the other person's "Face":

> The way in which the other presents himself, exceeding *the idea of the other in me*, we here name face. This *mode* does not consist in figuring as a theme under my gaze, in spreading itself forth as a set of qualities forming an image. The face of the Other at each moment destroys and overflows the plastic image it leaves   me, the idea existing to my own measure and to the measure of its *ideatum*—the adequate idea (Levinas 1969: 50-51).

The Face of the Other is a "mode" that "destroys" and "overflows" any idea I might have of him/her. The Face is not a "theme," nor an "image"; rather, it is "a signification without a context." As "transascendance" the Other expresses the surplus of signification (Levinas 1969: 23). The impossibility of integrating the Other's Infinity into one's thematic scheme is experienced as a shock or "traumatism." The properly ethical moment occurs in this traumatism, when the Other person faces and judges the self-Same. Taking the form of persecution, accusation, or injunction, the judgment of the Other's Face expresses an *asymmetrical* "rela-

tion" which it establishes with the self-Same, (Levinas 1969: 215-16).

> Its [i.e. the Other's] critical intention then leads it beyond theory and ontology: critique does not reduce the other to the same as does ontology, but calls into question the exercise of the same. A calling into question of the same—which cannot occur within the egoist spontaneity of the same—is brought about by the other. We name this calling into question of my spontaneity by the presence of the Other ethics (Levinas 1969: 43).

The ethical significance of the Face issues from the Other calling into question the spontaneous freedom and egoism of the "I." The face of the Other expresses absolute nakedness, "disengaged from every form, but having meaning by itself" (Levinas 1969: 74). This "meaning by itself," without form or context, is primordial signi-fication, the idea of the Infinite (Peperzak 1993:165). Revealed in the nudity of the Other's Face is the revelation that my egoistic existence is unjust. In this sense, the Other's Face as injunction "measures me" (Peperzak 1993: 116). And since the Face of the Other expresses the idea of the Infinite, it is precisely *Infinity* that I am measured by. Here, the face-to-face relation between the Other and the Same is strictly *asymmetrical*, and totally irreversible, which Levinas characterizes as heteronomy.

The "height" of the Other is first and foremost his/her authority as judge, a "master of justice," who calls me to good-ness. Interrupting my complacent existence and summoning me to respond, the judgment of the Other declares me *infinitely responsible* for the Other's very being. Why an *infinite* responsi-bility? Because the ethical requirement of the good goes to *infinity* (Peperzak 1993: 222, 192)! In this sense, no matter what I do for the Other, it is never enough; if we agree with Levinas that the Other's alterity signifies the Infinite, then the demands of good-ness are, in principle, unable to be satisfied.[8] Thus, I am always already *infinitely* guilty before, and *infinitely* responsible for, the Other! In *Otherwise than Being*, Levinas even goes so far as to say

that *I am guilty prior to my freedom*, that is, prior to any commitments I may have made or to any deeds I may have done.[9] My guilt simply issues from my existence; in my *Dasein* (my *being-there*) I have always already taken away the Other's possibilities of existence, even without intending to do so (Peperzak 1993: 116). This is why, in the language of *Otherwise than Being*, Levinas says that in the proximity of the Other I am a "hostage"; by the Other's election, she *makes* me responsible, and this responsibility I can never initially assume; it does not originate in my consciousness. Thus, my responsibility is entirely *passive*.

It is the initiative of the Other, then, which allows me to become an ethical subject.[10] It is the idea of the Other "in" me as infinite goodness which disrupts my formerly complacent existence and "obsesses" me to the point of substituting myself for-the-sake-of-the-Other. My (metaphysical) desire for the absolutely Other, that is, for the Infinite, for Goodness itself, goes against my will, against my better judgment, so to speak.

Recall that metaphysical desire is "disinterested." By this, Levinas means that it is always "despite oneself"[11] that one desires the Other. "This Desire is a desire in a being already happy: *desire is the misfortune of the happy* […]" (Levinas 1969: 62, my emphasis). Desire is "misfortunate" because it is self-dispossession, a "fundamental inversion" of the subject's very exercise of being, which "suspends its spontaneous movement of existing" in its egoistic interiority (Levinas 1969: 63). In a paradoxical sense, desire, which turns into an *obsessive responsibility* in the discourse of *Otherwise than Being*, is a *superlative passivity* where the subject is increasingly "inspired" by the Other's proximity, and so much so that this "inspiration" goes all the way to my substitution for-the-Other: "To give, to-be-for-another, despite oneself, but in interrupting the for-itself, is to take the bread out of one's own mouth, to nourish the hunger of another with one's own fasting" (Levinas 2004: 56). According to this ethical extremism, it is in my patient *suffering* for-the-Other as a "hostage" that I properly become an ethical subject, a "*sub-jectum*," assigned as "responsible for everything" prior to my freedom. Hence, "[t]o undergo

from the other is an absolute patience only if by this from-the-other is already for-the-other. This transfer, other than interested, 'otherwise than essence,' is subjectivity itself" (Levinas 2004: 111). Here, "I am 'in myself' through others" (Levinas 2004: 112). Speaking of this *radical passivity* which constitutes the ethical subject in *Otherwise than Being*, Levinas states:

> The recurrence of the self in responsibility for others, a persecuting obsession, goes against intentionality, such that responsibility for others could never mean altruistic will, instinct of "natural benevolence," or love. It is in the passivity of obsession, or incarnated passivity, that an identity individuates itself as unique, without recourse to any system of references, in the impossibility of evading the assignment of the other without blame. […] For under accusation by everyone, the responsibility for everyone goes to the point of substitution. A subject is a hostage (Levinas 2004: 112).

For the Levinas of *Otherwise than Being*, the subject's very subjectivity arises from his/her *asymmetrical* "relation" with the Other, where, under persecution before the Other's Face, his/her election as responsible is that which individuates his/her identity. "The identity aroused thus behind identification is an identity by pure election" (Levinas 2004: 145). Hence, for Levinas, alterity is at the core of the ethical self.[12]

With this established let us turn now to a discussion of Ricoeur, and to his critique of Levinas. Beginning with Ricoeur's critical assessment of Levinas will allow us to work backwards to a fuller picture of Ricoeur's "little ethics," and to arrive finally at an understanding of alterity as it relates to the Ricoeurian subject.

RICOEUR READS LEVINAS

Ricoeur's critique[13] of Levinas centers on Levinas's excessive use of hyperbole, which Ricoeur claims progressively intensifies from *Totality and Infinity to Otherwise than Being*. Ricoeur argues that there are a number of interrelated problems connected to the way

in which hyperbole is strategically employed as a method of philosophical argumentation in Levinas's discourse. Let us paraphrase these "interrelated problems" here before analyzing them in more detail below.[14]

1. Since the subject is defined by excessive interiority, by absolute separation, how is s/he to experience, to hear the call of, the Other? Because there can be no such "relation" established between the Same and the Other, they remain completely out of touch.

2. The absolute infinity of the Other, which makes him/her wholly exterior to the Same, ends with the same consequence: it puts the Same and the Other out of relation; i.e. in an "irrelation."

3. Ricoeur claims that, because of this irreducible exteriority separating the Same and the Other, coupled with Levinas's refusal to grant the subject any form of initiative, the only alternative left open to Levinas for breaching this "irrelation" is violence (i.e. "traumatism," "shock," etc.). Ricoeur thinks his interpretation of Levinas is legitimized by the asymmetry of the so called ethical "relation" where initiative is granted only to the Other: "This is why the Other… has to storm the defenses of a separate "I." (Ricoeur 1992: 190).

What ultimately concerns Ricoeur is the strict *asymmetry* of this "irrelation." Not only is it a question of *how* this strange "irrelation" is established, but the way in which its *asymmetry* precludes a response from the subject. Recall that Levinas refuses the ethical subject any recourse to his/her own powers in order to prevent a return to "the certainty of the ego that rejoins itself in freedom" (Levinas 2004: 118). Utterly *passive* and "disinterested," Levinas's ethical subject is for-the-Other all the way, resulting in the self-divestment that characterizes "substitution." Commenting on Levinas's inability to account for a responsive moment by the subject in the ethical relation, Ricoeur says:

> The question is then whether, to be heard and received, the injunction must not call for a response that compensates for the dissymmetry of the face-to-face encounter. Taken literally, a dissymmetry left uncompensated would break off the exchange of giving and receiving and would exclude any instruction by the face within the field of solicitude. But how could this sort of instruction be inscribed within the dialectic of giving and receiving, if a capacity for giving in return were not freed by the other's very initiative? (Ricoeur 1992: 189).

For Ricoeur, in order to not "break off the exchange" between the Same and the Other, the Other's movement towards the subject must be compensated by an opposing movement coming from the subject to the Other; in a word, there must be *reciprocity*. But, Ricoeur asks, how can "the self make itself available to others" when doing so implies, for Levinas, a return to the un-ethical life of egoism? For Ricoeur, the responsible subject must first be capable of receiving and recognizing the superiority of the Other's call. In this sense, "[o]ne has to grant a capacity of reception to the self that is the result of a reflexive structure, better defined by its power of reconsidering preexisting objectifications than by an initial separation" (Ricoeur 1992: 339). But, because "E. Levinas's entire philosophy rests on the initiative of the Other in the intersubjective relation," his ethics lacks the theoretical resources needed to explain how and why the subject is able to respond to the Other. Ricoeur contends that the initial "trauma" suffered by the subject in its experience of the Face must, at some point, be recovered, must become the subject's "conviction, a conviction equal to the accusative" coming from the Other, so as to avoid the result that, in order to establish any "relation," the Other "has to storm the defenses of a separate 'I'" (Ricoeur 1992: 190, 339). Ricoeur is pressing the point that the Face's injunction, as "meaning without a context," must be contextualized by an understanding subject in order for it to be not only heard but responded to. Thus, on Ricoeur's account, Levinas's "systematic practice of

excess in philosophical argumentation" unwittingly results in a paradoxical situation where the *irreversibility* of the *asymmetry* of the Face-to-Face "relation" can be read as "an inhumane condition called upon to say the ethical injunction" (Ricoeur 2004: 12).

> The self occupies the place of the Other without having chosen or wished to do so. The "despite-oneself" of the hostage condition signifies the extreme passivity of the injunction. This paradox—of an inhumane condition called upon to say the ethical injunction—should be shocking. The non-ethical says the ethical solely by virtue of its excess. If substitution must signify something irreducible to a will to suffer, in which the Self would recover mastery over itself in the sovereign gesture of the offering, of obligation, then it must remain "expulsion of self outside itself…the self emptying itself of itself" (OB, 110-111). In short, it must be by its "very malice" that "persecuting hatred" (OB, 111) *signifies* the "subjection through the Other" of the injunction under the aegis of the Good. I wonder whether Levinas's readers have assessed the enormity of the paradox that consists in having malice say the extreme degree of passivity in the ethical condition (Ricoeur 2004: 12).

In Ricoeur's mind, the excessiveness of "the hostage condition" belies a fundamental inhumanity at the core of Levinasian ethics. Because the self's initiative, as an ethical subject, is never his/her own, but always first *sub-jected* to the Other's command, the Other, Ricoeur tells us, is no longer only the "master of justice," but also the "offender," "executioner," or "oppressor" (Ricoeur 1992: 339). Here, Ricoeur's point turns on the fact that, according to Levinas, the Face does not "appear"[15] to the subject; the subject's "obsession" for the Other is "not consciousness, nor a species or modality of consciousness" (Levinas 2004: 86-87). And this begs the question: How, then, does the subject distinguish the "master of justice" from the oppressor? How does the subject recognize the reasons motivating the Other's command? On what basis is the Other's injunction justified? According to Ricoeur, due to the

excessive *asymmetry* that defines the ethical "relation," Levinas is unable to give a satisfactory answer to these questions.

It is here that Ricoeur claims Levinas's discourse, in *Otherwise than Being*, "reaches its paroxysm." It is in the theme of *substitution* as "hostage taking" that Ricoeur finds evidence for "a sort of reversal of the reversal performed in *Totality and Infinity*" which is accomplished by the culminating force of hyperbole in the discourse of *Otherwise than Being* (Ricoeur 1992: 340). By "reversal of the reversal," Ricoeur means the transition from the assignment of responsibility found in *Totality and Infinity* to the complete self-divestment of the subject found in *Otherwise than Being*.

> The assignment of responsibility, stemming from the summons by the Other and interpreted in terms of the most total passivity, is reversed in a show of abnegation in which the self attests to itself by the very movement with which it removes itself. Who, in fact, is obsessed by the Other? Who is hostage to the Other if not a Same not longer defined by separation but by it's contrary, substitution (Ricoeur 1992: 340)?

The moment of "paroxysm" occurs in this self-abnegation where the question "who" no longer matters. For Levinas, to ask the question "who?" is to enter the realm of reflection, consciousness, intentionality, objectivity, and *conatus*, which designates a return to "the certainty of the ego." In substitution, the "who" is emptied of itself for-the-sake-of-the-Other. In contrast, Ricoeur wants to claim that the accusative form of the injunction issuing from the Other cannot remain "non-assumable" by the responsible subject, as Levinas needs it to be, but must be received and converted into the self's conviction in order to avoid being stripped of all ethical significance (Ricoeur 1992: 340). Because of this, Levinas's ethics, in the end, can say nothing meaningful about responsibility as an ethical theme (Ricoeur 2004: 13). On Ricoeur's account, the notion of substitution, which is the effect of the strategic practice of hyperbole in Levinas's philosophy, leads

to an "impossible ethics," insufficient in its ability to account for a responsibility that does not begin and end in an "inhumane condition."

RICOEUR, AIMING AT THE "GOOD LIFE": SOLICITUDE

Now I would like to turn to Ricoeur's "little ethics," found in *Oneself as Another*, to investigate how he adjusts Levinas's thought in order to "compensate for the initial dissymmetry resulting from the primacy of the other in the situation of instruction, through the reverse movement of recognition" (Ricoeur 1992: 190). Undoubtedly, Ricoeur intends to preserve the primacy of the other found in Levinas's philosophy, but he also wants to provide a theoretical basis to account for a tenable ethical relationship between the subject and the Other, which he formulates as "*solicitude*." In my view, by investigating the place of *solicitude* within Ricoeur's ethico-moral framework we will be able to show how the Ricoeurian ethical subject, like Levinas's conception, involves "an originary being-abandoned-to-the-other." However, because Ricoeur's formulation presupposes precisely what Levinas's denies, I make the additional claim that Levinas and Ricoeur are, in the end, talking at cross-purposes.

To begin, let us note that Ricoeur interprets Levinas's privileging of the Other over the Same as an instance of the moral law (as injunction) arising on the ethical horizon too soon. According to Ricoeur,

> it is possible to dig down under the level of obligation and to discover an ethical sense not so completely buried under norms that it cannot be invoked when these norms themselves are silent, in the case of undecidable matters of conscience. This is why it is so important to us to give solicitude a more fundamental status than obedience to duty" (Ricoeur 1992: 190, my emphasis).

For Ricoeur, the ethical must take precedence over the moral in

order to allow for an adjustment to norms when they are unable to resolve ethical issues that arise in particular situations. If we return to the context of the injunction in Levinas's ethics, we can understand why Ricoeur privileges the ethical over the moral; the ethical (intention) restores to the subject the *capacity* to receive, recognize, and discern the call coming from the Other; in Ricoeur's words, it makes possible "the self's recognition of the superiority of the authority enjoining it to act in accordance with justice" (Ricoeur 1992: 190). In reference to Levinas, restoring the primacy of this "ethical sense" makes possible not only the distinction between the Other as the "master of justice" and the Other as the offender, but also the determination of whether or not the Other's call for expiation is at all justified.

Ricoeur's formulation of this "ethical sense" is comprised of "three components": (1) Aiming at the "good life", (2) with and for others, (3) in just institutions (Ricoeur 1992: 172).[16] Here, this "ethical sense" is construed by Ricoeur as essentially an "aiming," that is, as intentionality. Internal to this aiming at the "good life"[17] are both self-esteem, "understood as a reflexive moment" of the desire to live well, and *solicitude*, which is "not added on to self-esteem from the outside," but rather unfolds within it as its "dialogical dimension" (Ricoeur 1992: 180). For our purposes, the significance of this "ethical sense" resides in the way in which *solicitude* functions as a moment of self-esteem in its *lack* and *need*. The "dialogical dimension" internal to self-esteem, Ricoeur tells us, refers to the fact that what we *lack* and *need* are essentially others; we *need* others to esteem us in order for us to esteem ourselves. As Ricoeur puts it: "If self-esteem does indeed draw its initial meaning from the reflexive movement through which the evaluation of certain actions judged to be good are carried back to the author of these actions, this meaning remains abstract as long as it lacks the dialogic structure which is introduced by the reference to others" (Ricoeur 1992: 172). In other words, in order to become "good," my actions must refer to the judgment of others. "I cannot myself have self-esteem unless I esteem others as myself" (Ricoeur 1992: 193). According to Ricoeur, this

"as" structure[18] built into *solicitude* does not simply reduplicate the other person as a figure of myself, as an alter ego, but rather preserves the other's alterity in *mutuality*. As Ricoeur says in *The Course of Recognition*: "[T]he one is not the other at the very heart of the *alleloi*, the 'one another'" (Ricoeur 2005:152). For Ricoeur, the alterity between the self and the other is maintained within the reciprocal relationship of *mutuality*.[19]

This brings us to the major point of contention between Levinas and Ricoeur, which can be brought into focus by examining the role of *desire* in Ricoeur's "little ethics" vis-à-vis Levinas. With Levinas, we can ask: What is the status of alterity in Ricoeur's conception of the desire for the "good life"? Since the desire of Ricoeur's ethico-moral philosophy is prompted by *lack* and *need*, doesn't this formulation, by presupposing fulfillment and adequation, already destroy the absolute alterity needed for the ethical relation? Because Ricoeur's ethical thought is constructed within a theoretical framework that privileges equality, universality, reflection, consciousness, capacity, and so on, does it not render alterity *merely relative* to the capacities of an acting and suffering subject? As we mentioned above, Ricoeur thinks he can preserve alterity within the "as" structure of *mutuality*. However, from a Levinasian perspective, it seems clear that Ricoeur's ethical project may return in the end to the totalizing tendencies of the Same.

Ricoeur, of course, disputes this. He claims that, unlike Levinas, his philosophy has never assumed "an ontology of totality," which defines the identity of the Same as the will to closure (Ricoeur 1992: 335). On Ricoeur's reading, Levinas's radical formulation of the identity of the Same, to which the otherness of the Other is diametrically opposed, is much too simple, for it is unable to account for the subtle distinction that Ricoeur makes between the two sorts of identity in his hermeneutics of the self: *idem-identity* (sameness) and *ipse-identity* (selfhood). Addressing this alleged shortcoming internal to Levinas's conception of the identity of the Same, Ricoeur says: "It results that the self, not distinguished from the I, is not taken in the sense of the self-

designation of a subject of discourse, action, narrative, or ethical commitment" (Ricoeur 1992: 335). To Ricoeur's mind, this dual sense of identity allows him to circumvent the charge of relapsing into a Levinasian-style totality precisely because *ipse-identity* (selfhood) is not the "I," it is not the self-Same. "To say self is not to say I. The I is posited—or is deposed. The self is implied reflexively in the operations, the analysis of which precedes the return toward this self" (Ricoeur 1992: 18). In this sense, *the meaning of selfhood is fragile* because it is dispersed throughout the multiple reflexive detours and analyses prompted by the question "who am I?"[20] And this "implied self," to which one attempts to return, is always already permeated by various experiences/meanings of "otherness." As such, the other is always already *implicated* in the constitution of selfhood. This is precisely why Ricoeur claims that the identity of the subject (as *ipse-identity*, as a self) can never assume the extreme solipsism of a self-enclosed totality, which defines Levinas's identity of the Same. For Ricoeur, because self-understanding is an unending process of self-interpretation mediated by multiple analytic detours, the question of the "who?" of selfhood always remains open.

So, for Ricoeur, the "I" as *idem-identity* (sameness) is on par with Levinas's conception of the Same. But *ipse-identity* (selfhood) is of a different order, a different mode of being, than *idem-identity* (sameness). This difference between *idem-identity* and *ipse-identity* is based on "the twofold valence of permanence in time" (Ricoeur 1992: 318). Briefly put, *idem-identity* designates the immutability of *character*, which "assures at once numerical identity, qualitative identity, uninterrupted continuity across change, and, finally, permanence in time which defines sameness" (Ricoeur 1992: 122). Selfhood or *ipse-identity*, on the other hand, signifies *self-constancy* expressed as promising, which, Ricoeur tells us, "cannot be inscribed, as character was, within the dimension of something in general but solely within the dimension of 'who?'" (Ricoeur 1992: 123). Contra *idem-identity*, *ipse-identity* "does indeed appear to stand as a challenge to time, a denial of change: even if my desire were to change, even if I were to change

my opinion or my inclination, 'I will hold firm'" (Ricoeur 1992: 124). In this sense, "the ipse poses the question of its identity without the aid and support of the *idem*" (Ricoeur 1992: 124). Hence, the fragility of *ipse-identity* (selfhood) is made evident when one poses the question "who am I?" without recourse to the surety of identity provided by sameness. Self-constancy becomes self-attestation[21] by virtue of the assurance that, despite any sort of changes I may undergo, I will remain who I am so that others can count on me. We can see that it is in the mode of *ipse-identity* (selfhood) as self-constancy that we arrive with Ricoeur at the level of ethics.[22] And it is here that selfhood is mediated by otherness at its very core; the self is who s/he is only in relation to the context of promises kept or broken, that is, by the self's response to "being-enjoined" by others. Commenting on the ultimate structure of selfhood as *being-enjoined*, Ricoeur says:

> Because someone is counting on me, I am *accountable for* my actions before another. The term 'responsibility' unites both meanings: 'counting on' and 'being accountable for.' It unites them, adding to them the idea of a *response* to the question 'Where are you?' asked by another who needs me. This response is the following: 'Here I am!' a response that is a statement of self-constancy" (Ricoeur 1992: 165).

Self-constancy, produced by attestation in promise keeping, brings to a halt, even if only tentatively, the barrage of endless answers that could be given in response to the question "who am I?" Self-constancy is attested to in the conviction "Here is where I stand," as the promise that I will "be the same today as the one who acted yesterday and who will act tomorrow" (Ricoeur 1992: 295). To be sure, attestation remains nonsensical apart from the other; for attestation must pass through the test of whether or not one has lived up to one's promises, making good on the idea that, before the other, I am who I claim to be, and *this* one cannot do without recourse to the judgment of others. In this way the constancy of my self is inextricably bound to my responsibility before others; in remaining the same for-the-sake-of-the-other, I,

in turn, attest to who I am. The constancy of the one who attests to his/her own identity, then, is predicated on its reference to others. Thus, we can say that selfhood as attestation is configured *dialectically*, in the mediation between the self and the other. Here, on the ethical plane, the question "who am I?" receives its answer in the context of *being-enjoined* by the other, where keeping one's word signifies the assurance of being one's self in the act of attestation. And, it is in this sense that Ricoeur's conception of the self (*ipse-identity*) can be appropriately designated as "an originary being-abandoned-to-the-other."

LEVINAS AND RICOEUR: PARTING WAYS…

It is clear that both Levinas and Ricoeur theorize the ethical subject as "an originary being-abandoned-to-the-other." However, they do so in such radically different ways that, in the end, their respective ethical positions wind up at odds with one another. My claim here is that this antinomic character between Levinas and Ricoeur turns on the different ways in which both desire and alterity function within their respective ethical projects.

Let us recall that for Levinas a desire rooted in *lack* and *need* is ethically irrelevant. But this is precisely how Ricoeur construes the role of desire within the ethical relation. For Ricoeur, desire originates in an acting and suffering subject who desires the "good life." Integral to this "desire" is the *need* for others. Since desire, for Ricoeur, is motivated by *lack* and *need*, this suggests that Levinas would be quick to claim that Ricoeur's "little ethics" reduces *absolute* alterity to the status of being merely *relative* to the needs of a desiring subject.

Now, from Ricoeur's perspective, since Levinas defines the subject by absolute interiority and the Other by absolute exteriority, this raises the question of how a relation between them can even be established. On Ricoeur's reading, the only "relation" that Levinas allows between the Same and the Other is that which is established *asymmetrically*, via the injunction issuing from the Other's face. But, as we have seen, Ricoeur argues that Levinas's

position is paradoxical in the sense that the ethical response the Other demands from the Same begins in an already "inhumane condition." Due to the excessiveness of the condition of separation, which defines the "relation" (or "irrelation") between the Same and the Other in Levinas's ethical discourse, and coupled with the fact that Levinas only grants the power of initiative to the Other, Ricoeur believes that the Same can only be *made* to hear the Other's injunction if the Other "storms the defenses of a separate I." Thus, Ricoeur claims that Levinas's is an "impossible ethics," which unwittingly ends in the discourse of malice.

In the final analysis, we can say with Ricoeur that Levinas's ethics ends in "scandal" because the conditions that supposedly make possible the ethical "relation" are too excessive; in fact, the condition of separation turns on a double excessiveness traveling in opposite directions: interiority as absolute inwardness and exteriority as absolute alterity, always absolving itself from this inwardness. Thus, for Ricoeur, the infinite distance separating the self and the Other precludes the possibility of establishing any ethically significant relationship between them. With Levinas, we can say that Ricoeur does not conceive of the self and the Other as being separate enough, and, because of this, ends up destroying the very thing that is necessary for the ethical relation: the absolute alterity of the Other. It is precisely in this way that Levinas and Ricoeur seem to be talking at cross purposes.

Notes

1. I'm borrowing this phrase directly from Bernhard Waldenfels, found in his excellent essay "The other and the foreign." B. Waldenfels (1996), pp. 11-124. Indeed, I am following Waldenfel in his description of the Ricoeurian self being constituted by alterity at its very core, but I modify his description by showing how it intensifies and makes most sense at the ethical level of attestation.

2. The idea of the Inifinte, Levinas insists, is not merely a negation of being; its "content" is not negative (Levinas 1969: 40-42).

3. "Totality" is Levinas's term for characterizing the violence inherent in Western philosophy's approach toward the Other, evinced by its exercise of universal Reason, which, in its representational structure, reduces the Other's alterity to the neutral inter-signification of its conceptual schemes.

For Levinas, this "totalizing" feature of Western ontology is ethically problematic (i.e. violent) because it violates the notion of alterity. According to Levinas, alterity, or human otherness, exemplifies the irreducible difference between self and Other. Throughout all of his works, Levinas's primary argument is that absolute alterity is a necessary condition for the ethical relation because it guarantees an infinite distance separating self and Other, which, in turn, ensures that the Other is unable to be totalized by being reduced to the self's conceptual scheme. (Levinas 1969: 21-30).

4. The notion of the Other's Face will be dealt with in detail in the following section.

5. The Infinity of the Other is perhaps one of Levinas's most difficult notions to grasp. For a detailed discussion of this notion see *Totality and Infinity,* pp. 48-52.

6. More precisely, Levinas says that it is in language as "discourse" that the transcendence of the Other is concretely experienced. When the other person speaks, the meaning of her speech is immanent to my perception of it. However, what is indicated by the face as "discourse"—an(other) absolutely unique consciousness—remains transcendent to that which is discursively expressed. In other words, the significance of saying, of signification, is irreducible to that which is said. See *Totality and Infinity*, pp. 194-212.

7. Levinas actually poses this question at the beginning of *Totality and Infinity*, when he asks: "But how can the same, produced as egoism, enter into relationship with an other without immediately divesting it of its alterity? What is the nature of this relationship" (Levinas 1969: 38)?

8. Commenting on this aspect of infinite responsibility in Levinasian ethics, Peperzak says: "The life of freedom discovering itself to be unjust, the life of freedom in heteronomy, consists in an infinite movement of freedom putting itself ever more into question. This is how the very depth of inwardness is hollowed out. The augmentation of exigency I have in regard to myself aggravates the judgment that is borne on me, that is, my responsibility increases these exigencies. In this movement my freedom does not have the last word; I never find my solitude again—or, one might say, moral consciousness is essentially unsatisfied, or again, is always Desire" (Peperzak 1993: 117). This constitutes the formal structure of what Levinas calls "bad conscience."

9. "The responsibility for the other can not have begun in my commitments, in my decision. The unlimited responsibility in which I find myself comes from the hither side of my freedom, from a 'prior to every memory,' an 'ulterior to every accomplishment,' from the non-present par excellence, the non-original, the an-archical, prior to or beyond essence" (Levinas 2004: 10).

10. "It is in this very concrete sense that the judgment that is borne upon me is never assumed by me. This inability to assume is the very life, the essence, of

conscience. My freedom does have the last word; I am not alone" (Levinas 1969: 101).

11. Speaking of the desire for the good "despite-oneself," Levinas says: "It [the face of the Other] provokes this responsibility against my will, that is, by substituting me for the other as a hostage. All my inwardness is invested in the form of despite-me, for-another. Despite-me, for-another, is signification par excellence. And it is in this sense of the "oneself," that accusative that derives from no nominative; it is the very fact of finding oneself while losing oneself" (Levinas 2004: 11). See also pp. 51-53 of *Otherwise than Being*.

12. All throughout *Otherwise than Being* Levinas describes, in various formulations, the subjectivity of the ethical subject as the "breakup," "undoing," or "suspension" of essence. For instance: "This breakup of identity, this changing of being into signification, that is, into substitution, is the subject's subjectivity, or its subjection to everything, its susceptibility, its vulnerability, that is, its sensibility. Subjectivity, locus and null-site of this breakup, comes to pass as a passivity more passive than all passivity" (Levinas 2004: 14).

13. Ricoeur's critique of Levinas is found primarily in two texts: *Oneself as Another* and *Otherwise: A Reading of Emmanuel Levinas's Otherwise than Being or Beyond Essence.*

14. This summary is taken from Richard Cohen's essay "Moral Selfhood" in Ricoeur as Another: the ethics of subjectivity (2002) pp. 127-160.

15. Levinas rules out the Other's "appearing" to the self as a phenomenon because "vision is essentially an adequation of exteriority to interiority: in it exteriority is reabsorbed in the contemplative soul and, as an adequate idea, revealed to be a priori, the result of Sinngebung" (Levinas 1969: 295).

16. The triad composing this "ethical sense" is the heart of Ricoeur's "little ethics." Compressed within and parallel to this "ethical sense"—which Ricoeur also calls "benevolent spontaneity"—are three hypotheses which constitute the teleological structure of Ricoeur's ethical enterprise: (1) the primacy of ethics over morality, (2) the necessity for the ethical aim to pass through the sieve of the norm, and (3) the legitimacy of recourse by the norm to the aim whenever the norm leads to impasses in practice (Ricoeur 1992: 170).

17. The notion of "the good life" remains a limiting idea for Ricoeur. Commenting on this notion, he says: "With respect to its content, the "good life" is, for each of us, the nebulas of ideals and dreams of achievements with regard to which a life is held to be more or less fulfilled or unfulfilled" (Ricoeur 1992: 179-80).

18. Speaking of this "as" structure, Ricoer says: "This 'as being' (as being what the other is) averts any subsequent egotistic leanings, it is constitutive of mututality" (Ricoeur 1992 183-84). Moreover, in the beginning of *Oneself as Another*, Ricoeur stresses the point that the particle "as" in the formulation

"oneself as another" is to be understood in two senses: (1) As a comparative (wie), and (2) as an implicative (als). Since alterity for Ricoeur is not added on to selfhood from the outside, but constitutes the self from within, he stresses the implicative sense: "Oneself as Another suggests from the outset that the selfhood of oneself implies otherness to such an intimate degree that one cannot be thought of without the other…" (Ricoeur 1992: 3).

19. In *Oneself as Another*, Ricoeur analyzes *mutuality* into three basic elements: reversibility, non-substitutibility, and similitude (Ricoeur 1992: 192). Although the linguistic roles (personal pronouns) between speaker and listener can be reversed within the context of interlocution, the individuals performing these roles are themselves irreplaceable (nonsubstitutable). It is this feature of irreplaceability, "whereby each person is *irreplaceable* in our affection and our esteem," that first marks the inequality across the relational exchange between the agents and patients of an action. Finally, similitude, as "the fruit of the exchange between esteem for oneself and solicitude for others," signifies the equalization "of all the initially unequal forms of the bond between oneself and the other" (Ricoeur 1992: 193). Hence, "[i]t is this search for equality in the midst of inequality" that "defines the place of solicitude along the trajectory of ethics" (Ricoeur 1992: 192). In this way, *mutuality*, as "the esteem of the *other as a oneself* and the esteem of *oneself as an other*," contains the paradox of equalizing unequals, or, as Ricoeur says in *The Course of Recognition*, of "comparing incomparables" (Ricoeur 1992:194; 2005: 161).

20. In *Oneself as Another*, Ricoeur claims that in introducing the problematic of the self by way of the question "who," he has "opened the way for the genuine polysemy inherent in the question itself: Who is speaking of what? Who does what? About whom and about what does one construct a narrative? Who is morally responsible for what" (Ricoeur 1992: 19)? Here the question of "who am I?" becomes contingent upon which sort of question is be posed about the self. In this way the polysemy inherent in the question "who?" becomes correlative to the polysemy given in the answers about the self. Hence, the meaning of the self is polysemous through and through.

21. "[A]ttestation is the assurance—the credence and the trust—of existing in the mode of selfhood" (Ricoeur 1992: 302).

22. "The properly ethical justification of the promise suffices of itself, a justification which can be derived from the obligation to safeguard the institution of language and to respond to the trust that the other places in my faithfulness" Ricoeur 1992: 124).

Bibliography

Bourgeois, P. (2002) "Ricoeur and Levinas: Solitude in Reciprocity and Solitude in Existence," in: R. Cohen and J. Marsh (Eds), *Ricoeur as Another: The Ethics of Subjectivity*, pp. 109-126 (Albany: SUNY)

Cohen, R. (2002) "Moral Selfhood: A Levinasian response to Ricoeur on Levinas," in: R. Cohen and J. Marsh (Eds), *Ricoeur as Another: The Ethics of Subjectivity*, pp. 127-160 (Albany: SUNY)

Levinas, E. (1969) *Totality and Infinity: An Essay on Exteriority* (Pittsburgh: Duquesne University Press)

_____. (1996) *Emmanuel Levinas: Basic Philosophical Writings* (Bloomington and Indianapolis: Indiana University Press)

_____. (1999) *Alterity & Transcendence* (New York: Columbia University Press)

_____. (2004) *Otherwise Than Being: Or Beyond Essence* (Pittsburgh: Duquesne University Press)

Peperzak, A. (1993) *To the Other* (West Lafayette, Indiana: Purdue University Press)

Ricoeur, P. (1974) *The Conflict of Interpretations: Essays in Hermeneutics* (Evanston: Northwestern University Press)

_____. (1992) *Oneself as Another* (Chicago & London: The University of Chicago Press)

_____. (2004) "Otherwise: A Reading of Emmanuel Levinas's 'Otherwise than Being of beyond Essence,'" *Yale French Studies*104, pp. 82-99

_____. (2005) *The Course of Recognition* (Cambridge: Harvard University Press)

Waldenfels, B. (1996) "The other and the foreign," in: R. Kearney (Eds), *Paul Ricoeur: The Hermeneutics of Action*, pp. 111-124 (London: Sage)

The Exteriority of Accidents: The Edifying Effect of Foucault's Genealogy as a Critique of Norms

Anthony Ristow

> "My role … is to show people that they are freer than they feel, that people accept as truth, as evidence, some themes which have been built up at a certain moment during history, and that this so-called evidence can be criticized and destroyed." (Foucault 1982, *TSP*, p. 10)

Michel Foucault's genealogical project emerges with his essay "Nietzsche, Genealogy, History." The piece uses Nietzsche's concept of genealogy while focusing the historico-critical attitude they share into a more tangible and localized context. Foucault's genealogy is not destructive solely for the sake of destroying, as Kant advocates reason as an end in itself. Nor is genealogy a method that is explicitly ethical in any way. Genealogy "is to discover that truth or being does not lie at the root of what we know and what we are, but the exteriority of accidents" (Foucault 1977, *FR*, p. 81). Unlike Kant's commitment to normative qua practical reason or Heidegger's claim of disclosive truth, genealogy is "the history of an error we call truth," and it depends on this history to "dispel the chimeras of the origin" (Foucault 1977, *TFR*, p. 80). As Hubert L. Dreyfus and Paul Rabinow frame it in their book, *Michel Foucault: Beyond Structuralism and Hermeneutics*:

> They (Kant & Heidegger) both wanted to provide a universal theory and to know the sources and legitimate uses of the concepts presupposed by their predecessors. Foucault accepts this project but rejects the attempts to find a universal grounding in either thought or Being. Analytics today must find a way of taking seriously the problems and conceptual

tools of the past, but not the solutions and conclusions based on them. (Dreyfus & Rabinow 1983, *MF*, p. 122)

Although Heidegger is skeptical about the Kantian, representational concept of normativity, they both agree that there is some true and valid source of that normativity. In this paper, I have two aims. I will attempt to establish Foucault's genealogy as a critical enterprise that "criticizes and destroys" the validity of normative practices by rejecting any universal grounding behind them. If we accept his thesis, Foucault's method of genealogy exposes norms and practices as inherently arbitrary, selected and created on what amounts to *no good reason*. That is, Foucault mounts a critique against *de facto* cultural practices that, despite this absence of reason, operate and function under the presupposition of some such rationality. My second aim is to assert that Foucault's genealogical project has an edifying affect for this very reason. Not only is genealogy edifying in spite of its destructive diagnosis, but this affect is also predicated on that very diagnosis alone. I argue that by collapsing the distinction between norms and "accidents" genealogy is a potentially enlightening critique rather than something systematic in a Kantian sense or "fundamental" in a Heideggerian sense. Genealogy is showing us something: it doesn't have to be this way, yet Foucault does not advocate any alternative way that it *should* be.

I begin this paper with a contrast by providing a fairly non-controversial conception of normative practices in order to frame and argue for Foucault's genealogy as an accurate critique of such practices. In section (1) I offer a preliminary sketch of Kant's basic moral philosophy for the contrary view that norms make up a valid category distinct from mere facts. In section (2) I will present Kant's topical essay, "What is Enlightenment?" for his view of his own philosophy as a critical enterprise and then discuss Foucault's own essay on Kant's "What is Enlightenment?," in which he suggests that the normative Kant and the critical Kant can be divorced from one another. With section (3) I will then entertain an alternative view of everyday norms, which I find

more plausible, in the philosophy of Martin Heidegger. Yet, while it is my belief Heidegger's account shows Kant's to be deficient, it does not exempt Heidegger from the same problem of origin that both philosophies suffer. Indeed, Heidegger's belief in the deep, hidden truth of authentic Being as the source of all norms is more implausible. These accounts then make it possible in section (4) to do an analysis of Foucault's genealogy as presented in his essay, "Nietzsche, Genealogy, History." In looking at Heidegger's accurate depiction of everyday norms and practices, as well as his and Kant's dubious attempt to explain their origin, I hope to show Foucauldian genealogy in all its inglorious dissent.

Finally, in section (5), with this framework in place, I will be able to assert my thesis that genealogy is both a critical and edifying enterprise predicated on the denial of normativity. I conclude by comparing the potentially freeing affect of this project favorably with that of Heidegger's failed attempt to show his readers their potential freedom in *Being & Time*.

I.

1. Kant's Moral Philosophy

Immanuel Kant's normative moral philosophy is best explained in *The Foundations of the Metaphysics of Morals*. In this work, Kant claims to find the foundations of morality and tries to convert the everyday, seemingly intuitive standards of morality to a metaphysics of morals. The resulting moral judgments should be in accordance with what normal, sane adults would generally accept in some form of consensus. However, Kant is explicit that he is in search of the foundational principles of morality *a priori* and not something empirical, or factually based in acts.

In his concept of duty, according to which we are obligated as rational autonomous beings to act in conformity with moral law, Kant says that hoped-for outcomes of acts can have no bearing on dutiful conduct. Our dutiful acts must be motivated out of conformity to the *a priori* moral law alone. One can enact some type of duty but if it is motivated by anything other than a good will, it is

morally bereft no matter how virtuous or positive the outcome. In other words, for an action to have any moral worth it must have a pure intention regardless of its external manifestation (Kant 1959, *FMM*, p. 19 [402]). He explicitly states, "the purposes we may have for our actions and their effects as ends and incentives of our volition cannot give the actions any unconditional and moral worth… It can lie nowhere else than in the principle of the will, irrespective of the ends that can be brought about by such action" (Kant 1959, *FMM*, p. 16 [400]). In short, Kant believes the rightness and wrongness of an act is in the intended will, as governed by reason. Kant derives his ultimate, singular moral law from this strict concept of duty and unqualified good: the "Categorical Imperative." This is Kant's moral philosophy in a rather small nutshell but for present purposes all that is required is to show that Kant views norms as making up a category distinct from contingent facts.

Derived from the categorical imperative, the kingdom of ends is a "realm" entirely made up of rational beings that act according to universal maxims in adherence with the CI. The thought experiment reveals Kant's implausible intuition that a "systematic union of different rational beings," which is therefore good-willed, would be capable of agreement in coming up with the same, common laws for one another to follow as an egalitarian, legislative community (Kant 1959, *FMM*, p. 51 [433]). Thus, the prevailing norms would be based on good reason and clearly not incidental.

Reason is a universal faculty, so it makes sense that homogenous values would arise and become ever availed to normalization, insofar as rationality was chosen as advantageous. Conversely, Kant views heteronomy of the will as the clear source of any specious morality that spurs reason. "If the will seeks the law which is to determine it anywhere else than in the fitness of its maxims for its own giving of universal legislation … heteronomy always results" (Kant 1959, *FMM*, p. 59 [441]). In other words, when one wills outside of universal law or acts upon anything other than that which is determined by a proper universal maxim,

the resulting behavior is morally questionable.

According to Kant, moral norms are grounded in and validated by human reason; this is irrespective of, and clearly distinct from, consequential facts. Facts are mere outcomes whereas norms in general represent their own distinct category derived from the authority of universal reason. This supposedly valid authority would also explain their very normativity.

2. "What is Enlightenment?"

(a) Kant's Critical Enterprise

Let me now turn to "What is Enlightenment?," a minor essay in Kant's body of work, but one that garnered special interest from Foucault because of what he believed to be its significant departure in intellectual history. In this small essay, Kant defines enlightenment as "man's release from his self-incurred tutelage. Tutelage is man's inability to make use of his understanding without direction from another" (Kant 1959, *WIE*, p. 85 [35]). In Kant's examples, this vast majority is in arrested development, and do not bother to think or trouble themselves with autonomy because they have books that understand for them, pastors who have a conscience for them, and so on. These "placid creatures" do not take the "step to competence" because the guardians have made their "domestic cattle dumb" and afraid to learn to walk on their own. The motto of enlightenment is a challenge to this majority: "Have courage to use your own reason!" (Kant 1959, *WIE*, p. 85 [35])

Kant says enlightenment requires nothing but freedom. However, "the public can only slowly attain enlightenment," specifically through "freedom to make public use of one's reason." More specifically, he means the freedom to argue. But Kant finds that freedom is restricted everywhere. Everywhere, he hears, "Do not argue!…but drill…but pay…but believe!" In not so veiled terms, Kant points out his contemporary ruler, Frederick II, as the only prince reasonable enough to allow public use of one's reason by saying, "Argue as much as you will, and about what you will,

but obey!" (Kant 1959, *WIE*, p. 86-87 [36-37])

Here in the text, Kant makes a key distinction between the use of public and private reason. As we've established, public use of reason is predicated on the potentially trite claim that you must obey, but you are free to complain as much as you like. There is private use of one's reason that is restricted to "a particular civil post or office that is entrusted to him." In that context of the role one plays in some particular office, one is not free to make public use of her reason. "Here argument is certainly not allowed—one must obey." However, this is not a blind obedience; insofar as one is a member of a greater community or the "society of world citizens," she is allowed to argue, through her writings, by assuming the role of a scholar before the public. For example, Kant believes it would be disastrous for a military officer to debate her orders with a superior. Yet, she is welcome to complain and judge "as a scholar" at some later date. "The citizen cannot refuse to pay the taxes imposed on him…. But the same person nevertheless does not act contrary to his duty as a citizen when, as a scholar, he publicly expresses his thoughts on the inappropriateness or even injustice of these levies." (Kant 1959, *WIE*, p. 87-88 [37-38])

With a firm belief that human nature's proper destination lies in the progress of general enlightenment, Kant links this progress to the authority of the monarch—"for his lawgiving authority rests on his uniting the general public will to his own" (Kant 1959, *WIE*, p. 90 [40]). A head of state who allows his subjects the freedom to make public use of reason will find it advantageous to treat men as "more than machines, in accordance with their dignity," because it will only increase the civil obedience of his now free thinking constituency. Kant is advocating a reasoned obedience founded on free, public use of reason. He actually argues, paradoxically, that "a lower degree of civil freedom" in this context allows man to reach his full potential. One must assume this is because of the fact that, and on condition that, less civil freedom translates to greater freedom of thought, which Kant obviously deems preferable. "But only one who is himself enlightened [a monarch a la Frederick II], is not afraid of shadows, and has a numerous and

well-disciplined army to assure public peace, can say: 'Argue as much as you will, and about what you will, only obey!'" (Kant 1959, *WIE*, p. 91-92 [41-42]). This priority of public reason over private reason reveals a critical element in Kant's formulation of enlightenment. He is invoking a conditional relationship with the state, and the civil roles thereby entailed, dependent upon reason. For Kant, we are individually rational and therefore, autonomous. In order for the state to exercise legitimate power over us, the state must submit itself to a critique of practical reason.

(b) "The Contract of Rational Despotism" & "The Critical Task"

In his own essay titled "What is Enlightenment?" Foucault, not without irony, conceives of his genealogical project as a continuation of the Kantian notion of enlightenment. Foucault is not only analyzing how we have been conditioned by such a notion of enlightenment but he also appropriates Kant's critical enterprise in the process.

Foucault understands the concept of enlightenment not as an epoch or event but an attitude of modernity. To characterize modern philosophy is to ask the very question, what is enlightenment? In Kant's seemingly inconsequential piece *Was ist Aufklarung?* two centuries ago, Foucault finds him trying to answer this question in an entirely different way, "an almost entirely negative way." Kant characterizes enlightenment [Aufklarung] as an "exit" [Ausgang], or a "way out."

This "way out" as we have established in our reading of Kant is man's release from self-incurred tutelage, or what Foucault characterizes as the process which releases us from "immaturity"—the state of accepting someone else's authority when one's own reason is called for, a self-incurred unburdening of oneself into the hands of another - best illustrated in the "blind obedience" of "Don't think, just follow orders" (Foucault 1984, *FR*, p. 34). But Foucault is concerned by Kant's notion of reason here because it is politically and institutionally loaded. Mankind's mature status is not the opposite of immaturity, "Don't follow

orders, just think," nor is it simply, "Think." Rather, "Humanity will reach maturity when it is no longer required to obey, but when men are told: 'Obey, and you will be able to reason as much as you like.'" (Foucault 1984, *FR*, p. 36)

It is this politico-critical element of Kant's treatment that piques Foucault's interest as something of philosophical consequence. The public use of reason for Kant is free, as in the scholarly role before the public, reasoning as "a member of reasonable humanity," whereas the private use of reason is "submissive," like that of the subordinate soldier on duty, where man is like a "cog in a machine"—"a circumscribed position, where he has to apply particular rules and pursue particular ends" (Foucault 1984, *FR*, p. 36). With misgivings, Foucault then asks, "how the use of reason can take the public form that it requires…while individuals are obeying as scrupulously as possible?" On this view, enlightenment is not merely "an obligation prescribed to individuals," but also a political problem:

> And Kant, in conclusion, proposes to Frederick II, in scarcely veiled terms, a sort of contract—what might be called the contract of rational despotism with free reason: the public and free use of autonomous reason will be the best guarantee of obedience, on condition, however, that the political principle that must be obeyed itself be in conformity with universal reason. (Foucault 1984, *FR*, p. 37)

Considering this formulation of Enlightenment, with its emphasis on reason, "now it is precisely at this moment that critique is necessary;" for critique is that which determines the legitimate uses of reason, which presumably produce norms, and the illegitimate uses of reason which give way to "heteronomy." Foucault then proposes a hypothesis in which he believes this particular text, *Was ist Aufklarung?*, to be a found at the "crossroads of critical reflection and reflection on history"—a reflective critique by Kant on his time and his philosophical enterprise. (Foucault 1984, *FR*, p. 37)

Foucault's emphasis in this essay is twofold. First, he wants

to show how "a type of philosophical interrogation—one that simultaneously problematizes man's relation to the present," to history, and to "the constitution of the self as an autonomous subject"—has its roots in the Enlightenment. And while Foucault endorses enlightenment in this sense, he harbors very serious reservations about the Kantian notion of reason; a notion which has been proposed as the governing principle in this philosophical interrogation of ourselves. And so, secondly, Foucault emphasizes that our connection to the Enlightenment is not a fidelity towards any doctrine or value such as reason in particular, but rather what he calls the "permanent reactivation of an attitude"—which is to say, a disposition of "permanent critique of our historical era" (Foucault 1984, *FR*, p. 42). Just why he believes this will have to wait until we have conducted further investigation into the origins of normative practices and how they relate to such reason.

3. Inauthentic Dasein & "The Call of Conscience"

> "For me Heidegger has always been the essential philosopher…. My entire philosophical development was determined by my reading of Heidegger. But I recognize that Nietzsche prevailed over him." (Foucault 1985, *FI*, p. 8)

In the philosophy of Martin Heidegger we encounter an account of prevailing norms that is more plausible than Kant's distinct category derived from the authority of universal reason. Heidegger captures a more accurate picture of how we are *de facto*, rather than *de jure*, normatively inclined 'in the world'—by virtue solely of the inauthentic cultural practices of what he calls "the They." Heidegger's project in *Being & Time* is for the most part a description how typical, everyday beings are in the world. He believes that there is an authentic way of being and an inauthentic way of being—a dual self that is simultaneously both authentic and inauthentic. However, the way we are in the world, our norms, and the dealings that we busy ourselves with are all essentially evasions of the authentic self and our anxiety over death. In a sense, being in the world is interpretation all the

way down. There is no doubt that Heidegger was an influence on Foucault in this way, but what lies at the bottom of all this interpretation is a point of contention.

For Heidegger, human being, or Dasein, is always already in the world with others, for others, and determined by others. Therefore, all the facts about us, along with our values and norms, are not of our choosing, nor are they within our control; they are thrust upon us, or as Heidegger puts it, we are "thrown" into the world of "the They"—or in Nietzschean terms, "the herd." "The They" [Das Man] is Heidegger's concept for the inauthentic, public realm with others - "The 'who' is not this one, not that one, not oneself, not some people, and not the sum of them all. The 'who' is the neuter, *the 'they.'*" Most importantly, in "the They," "every Other is like the next" (Heidegger 1962, *BT*, p. 164). In other words, you are no more than an average representative of a corporation of norms—which you were "thrown" into without choice or control—and of which, you have become a collaborator and participant. What you can do, want to do, should do, and, for the most part, will do, is determined by the "dictatorship of the 'they,'" which unfolds in its "ready-to-hand" "inconspicuousness and unascertainability." One does not typically contemplate their "thrownness," rather, they simply go about their dealings in the world, taking norms for granted and conforming. "Thus the particular Dasein, or individual, in its everydayness is *disburdened* by the normalization of the 'they.'" (Heidegger 1962, *BT*, p. 165)

In this sense, it is similar to Kant's idea of immature, self-incurred tutelage, but, unlike Kant, Heidegger believes that norms are not the result of autonomous legislators, but a product of a contingent social situation. According to Heidegger, we are "thrown" into cultural practices, which means we do not create them through reflection and obedience to our faculty of reason. Rather, we encounter our moral norms as always already within the world in a "ready-to-hand" context—as given in practical use (Heidegger 1962, *BT*, p. 98). As inauthentic Dasein, we are not rational in the Kantian sense whatsoever; we do not contemplate our reason as "present-at-hand"—through, intentional, reflec-

tive and otherwise representational contemplation. In fact, for Heidegger, this type of Kantian rationality, which makes reason "present-at-hand," is a derivative corruption of "readiness-to-hand" and how we typically act in the world. If I am reflecting on my moral obligations as a soldier, making them "present-at-hand," then I already know how to march, shoot, fight, etc. If I am making use of my public reason by complaining about taxation, I have long known how to be a citizen absorbed within a society. "Readiness to hand" is prior to "presence-at-hand" in its "non-thematic circumspective absorption in references and assignments." (Heidegger 1962, *BT*, p. 105-107)

Heidegger says that we simply assume without reflection, without making "present-at-hand" any reason, and conform to the social norms into which we were "thrown." For Heidegger, the universal, autonomous reason Kant is advocating as the legislator of norms is derivative of practical, external manifestations in the world and therefore, such norms cannot be valid by Kant's pure, intentional standards. Good will is simply a *post facto* addition in reflection. As a matter of fact, norms are historical, contingent features of everyday, inauthentic life determined by the "They." They are not a distinct category from facts. So it would seem that in order to exercise autonomy in the Kantian sense, we would have to somehow become liberated from "the They." By Kantian lights, however, Heidegger's way out of "the They" is indeed even more irrational.

According to Heidegger, the only way we can encounter something that actually matters to us, and liberate ourselves from "the They," is through the quintessential mood of "anxiety." Typically, for Heidegger, we "flee" from anxiety because it discloses our authenticity to ourselves as an apparent, "undisguised" possibility and because we flee we are "fallen." This fallenness is exercised as "care," which is simply our dealings with things and others, and it is the state of the inauthentic, "they" self. In short, Dasein "clings" to the world and cares as a consequence of its running away from the prospect of its own demise. (Heidegger 1962, *BT*, p. 230)

Heidegger holds that, ultimately, the reason one cares, or has care in the world, is to cope with the futural prospect of death. "The urge 'to live' ('at any price') is not to be annihilated" (Heidegger 1962, *BT*, p. 240). And given the temporal characterization of Dasein as care, death poses a problem for Dasein in that death is the limit of possibility. In Heidegger's words, "Death is the possibility of the impossibility of Dasein. Thus death reveals itself as that *possibility which is one's ownmost* [exclusively your possibility], *which is non-relational* [disconnected from anyone else], *and which is not to be outstripped* [the limit, no going beyond it]" (Heidegger 1962, *BT*, p. 294). Thus, an honest relationship to and towards death, through anxiety, is what authenticates and sets one apart as an autonomous being.

Since death is non-relational and cannot be outstripped, the authentic, existential conception of it is a future self that lies where possibility is no more: "Being-towards-death." This type of Being does "not evade," or "cover up this possibility" by fleeing from it; rather, it "comports itself authentically towards its end" (Heidegger 1962, *BT*, p. 304). As such, it is through this *ownmost* possibility that Dasein can realize its authenticity—"Here it can become manifest to Dasein that in this distinctive possibility of its own self, it has been wrenched away from the 'they,'" for "death *lays claim* to it as an *individual* Dasein" (Heidegger 1962, BT, p. 307-308). Heidegger prescribes a type of liberation through embracing and listening to anxiety which actually attains for us "**freedom towards death**—*a freedom which has been released from the Illusions of the 'they,' and which is factical, certain of itself, and anxious.*" (Heidegger 1962, *BT*, p. 311)

However, it remains to be seen why it must be anxiety that discloses this to Dasein. Heidegger simply asserts that this is the case. It is certainly conceivable that Heidegger is privileging the mood of anxiety over all other moods, and the fear of one's own death over all other fears. Is it truly a necessary condition of Being that Dasein fears death over all others, and that this can only be learned through an examination of one's own particular anxiety, especially when we consider the dubious premise that this

constant and exclusive anxiousness is so unique that it individual-izes us from our "fallen" status?

Regardless of these contrasting views of normativity, Heidegger and Kant are in agreement that norms do in fact have a source; for both thinkers norms are derived from a specific, authoritative voice. In Kant this source is the universal human faculty of reason in the form of the categorical imperative, and, in *Being & Time*, Heidegger gives us the "call of conscience." In fact, when talking about "the call of conscience" and the state of "Being-guilty," Heidegger invokes the categorical imperative in his philosophy. Despite his wholesale rejection of anything like a Kantian metaphysics of morals and intentional rationality, Heidegger does indeed advocate a true, authoritative groundwork for our prevailing norms.

Heidegger believes that without this authentic "Being-towards-death," which is always there as an inherent part of Dasein—whether disclosed, understood, misunderstood or ignored completely—one would not "care" in the world. With this relationship, the inauthentic self owes, or is indebted to, the authentic self for such care. This debt, for Heidegger, is commu-nicated in the "call of conscience." "*In conscience Dasein calls itself*"—the authentic self pursues Dasein against its will, "like an *alien* voice," as a threat to "the lostness in which it has forgotten itself" (Heidegger 1962, *BT*, p. 321). Furthermore, Dasein flees the "call of conscience" which then manifests itself in the world as care. For example, conscience is interpreted by the "they" as some kind of perversion, a "universal conscience" which further becomes "exalted to a 'world conscience,'" which, according to Heidegger, is a "dubious fabrication" (Heidegger 1962 *BT*, p. 323). The call could be misconstrued into something within the world akin to Kant's categorical imperative.

"Being guilty" is the result of fleeing the "call of conscience," as communicated from the abandoned, indebted, authentic self. "This essential Being-guilty is, equiprimordially, the existential condition for the possibility of … morality in general and for possible forms which this may take factically" (Heidegger 1962,

BT, p. 332). In short, this is Heidegger's origin story for why and how we have morality and norms at all. With the notion of such a call, Heidegger is attempting to explain why we care and feel obligated, through guilt, to follow norms; for these are norms into which we are thrown and have no authentic or chosen significance for us.

As accurate as Heidegger's depiction of inauthentic, everyday Dasein in its "thrownness" may seem, his sweeping claims about ethical normativity in general are not persuasive. The call of conscience and being-guilty are not only unnecessary conditions of moral norms, but if they were essential they would indicate arbitrary and contingent norms at best. He simply gives an account of why humans have norms (fleeing the call of conscience) and why they feel the inclination to follow those norms (feeling guilty towards the abandoned, authentic self), but there is nothing in this account that says anything of traditional moral worth. There is no discernible distinction between a morally weighty norm like "thou shalt not kill" and, to use Heidegger's famous example of the craftsman, how one ought to use a hammer. All he's actually saying is that we feel guilty towards our authentic self, and, as a result, we make do with ethical norms. That's an interesting notion but the fact remains that those norms could be absolutely anything; they are the result of historical accidents, entirely subject to change. They are incidental of a general sense of guilt, and how they manifest is historically contingent. Therefore, it need not be a morally relevant claim at all.

But Heidegger seems to imply a normative payoff in this discovery, that there is something to be learned from the fact that our norms are indirect consequences of the authentic self. What's more, he suggests that by getting in touch with our anxiety, our death, and therefore our authenticity, this will impart some kind of new, "unthrown" vantage point that gives way to more authentic norms. By getting in touch with the authentic self, one is in a sense *not guilty*, or no longer *cares*, in the world, which has been disclosed as insignificant. Rather, they are now freed up to engage in "authentic care." However, this begs the question, why should

they? And what is the criterion for this authentic normativity Heidegger alludes to? In a bizarre turn, Heidegger seemingly advocates a Kantian type of ethics, an ethics he has no recourse to after thoroughly and effectively arguing the utter contingency of, and lack of reflective reason for, norms in general.

What's worse is that these claims are not only insubstantial, but they are unsubstantiated as well. The very basis of Heidegger's origin story is implausible. One can easily be persuaded by Heidegger's description of inauthentic Dasein, but this in no way entitles him to the existence of an additional, authentic self. It is much more apparent that he has illustrated how Dasein is and acts in the world typically rather than unearthed any hidden truth. Perhaps one can Be-towards-death (in touch with one's own impossibility), which it seems *is* what Heidegger considers authentic, but how could this way of Being plausibly be outside the world, beyond the historically determined circumstances and values of one's time? And this new attitude towards death still wouldn't make a chronic, lurking fear of death, embodied in this authentic self, any more plausible as the grand motivator of all life. It would seem more plausible if this dual concept of authenticity/inauthenticity were just a *potentially* enlightening reaction to the discovery of one's own contingency and thrownness.

So if one throws out the notion of authenticity or sees it as a reaction rather than an "existentiale," there is obviously no authentic self to which Dasein owes any debt whatsoever, and, therefore, there is no call of conscience or necessary condition of being guilty. It as though Heidegger is prescribing a supra-historical perspective towards life from the authentic standpoint, but the fact remains that if there is no authentic self, the resulting norms, whatever they may be, are empty and do not correspond to any deep truth. (And even if they did, they would be entirely contingent upon each Dasein's individual choice, whether authentically chosen or not.) Heidegger tells us that "the call undoubtedly does not come from someone else who is with me in the world;" the call comes from the authentic self (Heidegger 1962, *BT*, p. 320). But how could it not come from someone in the world? It comes from

Heidegger himself, and it serves as yet another story, in a long line of narratives, which peddles an incarnation of essentialism. A more persuasive argument for the whither and why of contingent moral norms, and care in the world in general, is that of Nietzsche and Foucault, according to which the origin of moral normativity is a historical accident by process of power as an end in itself.

If we engage Foucault's genealogy in this respect, there is no authentic self. The genealogist would reject this notion altogether. Consequently, in a Heideggerian sense, we are only our inauthentic selves, and necessarily thrown into our factical, and therefore essentially meaningless normative situation. In accordance with Heidegger, the genealogist rejects Kant's notion of autonomous reason, but by the same token, cannot accept Heideggerian authenticity either. There is not only discomforting groundlessness behind all our interpretation within the world of the "they," but a lacuna waiting beneath the shallow interpretation of our entire existence. So there is no essential authenticity or freedom derived by Being-towards-death. It is far more plausible that one can never be outside the world; we are "always limited and determined" (Foucault 1984, *FR*, p. 47). One cannot be held out in the "nothing" because it is nothing. And as Foucault tells us, "we are nothing but our history." (Dreyfus & Rabinow 1983, *MF*, p. 122)

II.

4. "Nietzsche, Genealogy, History"

Foucault's work is based on the theory that the human and social sciences are made up of dubious artifact concepts that grip no matter of fact, and thus express no stable extensions. That is, there are two categories of sciences: those which Foucault calls "sciences which have passed the threshold of scientificity" with "relatively stable practices and objects," such as physics, and "dubious disciplines like the human sciences," such as psychiatry (Dreyfus & Rabinow 1983, *MF*, p. 116). In this cultural context, concepts like "morality," "knowledge," "justice," "duty," even "truth" are all constructed, artifact concepts. There is no *good*

reason to adhere to these concepts and they amount to fictional stories that we tell ourselves. Foucault's earlier project, archaeology, was focused on the study of these normative concepts as a confluence—but determined by arbitrary rules for *selecting* certain objects over others; a kind of history of representation. However, in the 1970s, Foucault's work took a decidedly Nietzschean turn, replacing arbitrary *selection* with an analysis of the productive and creative exercises of power, in his seminal essay, "Nietzsche, Genealogy, History." Those concepts, which account for normativity in cultural practices, still lack any true essence for the genealogist, but rather than merely being selected for dubious reasons, they are actually created and appropriated as a function of impersonal power.

There are no authors behind the movements and permutations of history—this is what Dreyfus and Rabinow call "strategies without strategists" (Dreyfus & Rabinow 1983, *MF*, p. 109). There is an impersonal nature to this concept of power that only emerges through tactics, practices, and "meticulous rituals" rather than individual actors exercising autonomous reason. Whereas conventional thought might conceive of "substantial entities" as the condition for a relationship of struggle, the genealogist sees these entities as created by the very emergence of that struggle of forces. "Subjects do not first preexist and later enter into combat or harmony. In genealogy subjects emerge on a field of battle and play their roles, there and there alone. The world is not a play which simply masks a truer reality that exists behind the scenes. It is as it appears. This is the profundity of the genealogist's insight" (Dreyfus & Rabinow 1983, *MF*, p. 109). By recognizing these superficial emergences as such, the genealogist practices "effective history." He understands that "knowledge is made for cutting." Kant advocates a Kingdom of Ends but "the world of effective history knows only one kingdom, without providence or final cause, where there is only 'the iron hand of necessity shaking the dice-box of chance.'" (Foucault 1977, *FR*, p. 88-89)

Foucault concludes "Nietzsche, Genealogy, History" by posing Nietzsche's concept of 'the will to knowledge' as a potential

problem in this way. All knowledge, for Nietzsche and Foucault, "rests upon injustice," yet it masquerades as a neutral endeavor committed only to the search for objective truth, which it has no right to. In reality, historical analysis reveals this "rancorous will to knowledge" as amplifying risk and danger for those subject to it (Foucault 1977, *FR*, p. 95). "Where religions once demanded the sacrifice of bodies, knowledge now calls for experimentation on ourselves; calls us to the sacrifice of the subject of knowledge." Foucault's concern here, by Dreyfus and Rabinow's account, is that the genealogist understands "knowledge is thoroughly enmeshed in the petty malice of the clash of dominations" but not only does it not "offer a way out," knowledge exacerbates and multiplies the dangers we face as subjects (Dreyfus & Rabinow 1983, *MF*, p. 114). In the wake of Enlightenment or Nietzsche's death of god, so to speak, traditional value systems and religions are unhinged. This is a danger because desire for knowledge is a passion that "fears nothing but its own extinction." (Foucault 1977, *FR*, p. 96-97)

The "endless deployment" of the ruinous "will to knowledge," these "strategies without strategists," gives way to what Dreyfus and Rabinow call "procedural reason." It seems Foucault's concern with the problem of knowledge, which is more specific than Nietzsche's, is the exponential tendency of norms towards all-consuming totalization. But, for Foucault, these norms, which carry with them great consequences, only amount to haphazard, meticulous rituals of power:

> We try to ground our norms in reason, but it is as if reason, which for the Greeks corresponded to static natural kinds, has become unmoored and no longer corresponds to anything beyond itself. As Kant argued in The Critique of Pure Reason, scientific rationality, once cut off from things in themselves, must seek ever more general principles under which to subsume more and more phenomena, and ever more refined categories into which to subdivide the phenomena. Thus reason becomes procedural, the demand for greater

and greater systematization for its own sake. (Dreyfus & Rabinow 1983, *FM*, p. 258-259)

Once "unmoored and no longer corresponding to anything beyond itself," normativity becomes reason, to use a Nietzschean phrase, "at any price." Foucault, attempting to establish this dangerous trend in normativity, the problem of "the sacrifice of the subject of knowledge," utilizes a genealogy of our norms to see just how they did become affixed to this faith in procedural reason. This leads us back to Kant. According to Dreyfus & Rabinow, Foucault finds precisely this connection and point of departure in Kant's topical essay *Was ist Aufklarung?*, and his Enlightenment challenge—"Can humanity reach its maturity by using its reason to overcome its subservience to anything but its own rational capacities? Kant argues that the culture will gain maturity when the state, in this case Frederick the Great, takes over the task of assuring the onward march of reason in every sector of society" (Dreyfus & Rabinow 1983, *FM*, p. 259). But Kant believed that mankind was imbued with a universal, autonomous reason that would correspond to norms *de jure* and govern political principles. Yet, if one accepts genealogy and Heideggerian hermeneutics, there is a gaping void where reason once adhered to something objective, universal, and true. "Examining the history of reason, [the genealogist] learns that it was born in an altogether 'reasonable' fashion—from chance" (Foucault 1977, *FR*, p. 78). Foucault sees Kant proposing that we fill in this lacuna with what amounts to entirely empty procedural reason, as administered by the state.

As we've established, Kant tells us that mankind will achieve mature enlightenment when it is no longer required to thoughtlessly and blindly obey, but when there is a ruler who stipulates, "Obey! And you will be able to reason as much as you like." However, what Foucault calls the "contract of rational despotism with free reason" presupposes that the "political principle that must be obeyed" is itself "in conformity with universal reason." Obedience is conditional upon justified, reasonable orders. That is, the "rational despot" must, in fact, be rational. But Foucault's

genealogical project denies objective, universal reason altogether. Therefore, no monarch, no State, no political principle can be in conformity with universal reason. That would be a breach of contract. If there is no rationale or justification for the State, or its principles, then it is, in fact, based on no *good reason*. So by this account one does not have to obey. One does not *have* to do anything. Rather, Foucault proposes an active disobedience or "hyperactive pessimism" (Dreyfus & Rabinow 1983, *MF*, p. 264). For Foucault, the appropriate contemporary response to the Kantian notion of Enlightenment is transgression.

"The point … is to transform the critique conducted in the form of necessary limitation [the Kantian notion] into a practical critique that takes the form of possible transgression" (Foucault 1984, *FR*, p. 45). Consequently, criticism is no longer practiced on the assumption of or search for foundations or universals. Instead, criticism turns into a "historical ontology of ourselves" based on events and facts. By engaging Kant, Foucault asserts that although the historical event that was the Enlightenment did not make us into "mature adults" and he is dubious whether the "critical task still entails faith in the Enlightenment" at all, he does believe that Kant's "critical interrogation on the present and on ourselves" was meaningful; there is still work to be done on our problematized selves. "I continue to think that this task requires work on our limits, that is, a patient labor giving form to our impatience for liberty." (Foucault 1984, *FR*, p. 49-50)

5. "To Think Otherwise"

"And this critique will be genealogical in the sense that it will not deduce from the form of what we are what is impossible for us to do and to know; but it will separate out, from the contingency that has made us what we are, the possibility of no longer being, doing, or thinking what we are, do, or think." (Foucault 1977, *FR*, p. 46)

However destructive Foucault's genealogy proves to be as a critical enterprise, it also has an edifying affect. Foucault's critique carries with it the ability to impress the mind and move

the feelings. Much aligned with Heidegger's portrayal of inauthentic Being, Foucault offers a penetrating look into our historically determined world via our normative beliefs and practices. And this has the potential "to show people that they are freer than they feel." Like Heidegger, genealogy shows us that his notion of the "that-it-is-and-has-to-be" of being "thrown," born blind into the tyranny of our norms, is actually a case of that-it-is … *but-does-not-have-to-be*. Edith Wyschogrod calls this similarity the "emancipatory askeses" that both Heidegger and Foucault share—"disciplines of liberation in which each may be seen as engaged in the freeing of knowledge and truth from embedding context of repressive epistemological constraints and their ancillary ethical implications, a freeing through which a certain release is attained." (Wyschogrod 2003, *HFA*, p. 276)

However, I would argue that genealogy achieves a freeing affect in relation to one's "fallenness" into all-pervasive normativity without relying on a false notion like the essentialism of uncovered authenticity. As Dreyfus and Rabinow observe, while genealogy, similarly to Heidegger, might offer "some sort of liberation" or perhaps "increased flexibility" as a result of facing the unsettling truth - the "realization that nothing is grounded and that there are no guidelines"—Foucault's project does not confide in nor depend on "some authority which has already seen the truth" to "lead the self-deluded participant to see it too. (In Being and Time this authority is called the voice of conscience.)" (Dreyfus & Rabinow 1983, *MF*, p. xxvi)

Rather, Foucault's genealogy is edifying by virtue of its unique, brute, undeniability—"What convictions and, far more decisively, what knowledge can resist it?" (Foucault 1977, *FR*, p. 82). That is, genealogy is irresistible. It has the uncanny ability to "separate out, from the contingency that has made us what we are, the possibility of no longer being, doing, or thinking what we are, do, or think." The payoff for the genealogist is that there is no deep, hidden truth in the authentic self and no universal reason governing and validating our norms through a categorical imperative. If we accept this genealogical critique we are in a position to

acknowledge our own historical contingency; the fact that only facts remain. Implicit in this is freedom from our own norms.

Bibliography

Dreyfus, Hubert L. & Rabinow, Paul. (1983) *Michel Foucault: Beyond Structuralism and Hermeneutics* (Chicago: University of Chicago Press)

Foucault, Michel. (1977) "Nietzsche, Genealogy, History," in: Rabinow, Paul (Editor), *The Foucault Reader*, pp. 76-100 (New York: Pantheon Books)

_____. (1982) "Truth, Power, Self: An Interview with Michel Foucault," in: Martin, L.H. et al (1988), *Technologies of the Self: A Seminar with Michel Foucault*, pp. 9-15 (London: Tavistock)

_____. (1984) "What is Enlightenment?," in: Rabinow, Paul (Editor), *The Foucault Reader*, pp. 32-50 (New York: Pantheon Books)

_____. (1985) "Final Interview," *Raritan* (summer 1985), p. 8

Heidegger, Martin. (1962) *Being & Time* (New York: Harper & Row)

Kant, Immanuel. (1959) *Foundations of the Metaphysics of Morals* (Indianapolis: ITT Bobbs-Merrill Educational Publishing Company, Inc.)

_____. (1959) *What is Enlightenment* (Indianapolis: ITT Bobbs-Merrill Educational Publishing Company, Inc.)

Wyschogrod, Edith. (2003) "Heidegger, Foucault, and the Askeses of Self-Transformation," in: Milchman & Rosenberg (Eds), *Foucault and Heidegger: Critical Encounters*, pp. 276-294 (Minneapolis, Minn.: University of Minnesota Press)

MOTHERHOOD

Paige Harrison

Motherhood is a cornerstone of our society. In some shape or another, mothers have played an important role in the construction of both the propagation of the human race and in how we humans evolve from helpless infants to independent adults. The significance of motherhood seems obvious. This paper will explore whether or not motherhood is best viewed as a social or biological construction. I will discuss what it means to be a natural kind or a socially constructed kind. I will then discuss how motherhood might be viewed through the lens of those definitions. Finally, I will present my views as what *kind* motherhood is.

As a biological entity, mothers not only play a pivotal role in the conception of life, they also gestate that life until the person is prepared to live outside the womb. This gestational process is not something to be taken lightly. Women who carry a healthy baby to term generally make large sacrifices for the health of the unborn baby by refraining from activities that might bring harm to that child. Traditionally, how motherhood is defined has been deemed important for both legal and policy reasons. Susan Feldman writes, "Thinking of maternity in a way which emphasizes the gestational role, and backing it up with appropriate social policy, would have beneficial effects on the health of newborns" (Feldman 1990, p. 100). I would propose we establish the right sort of philosophical metaphysics first and worry later about what sort of political implications are drawn.

MOTHERHOOD AS A NATURAL OR BIOLOGICAL KIND

According to Alexander Bird and Emma Tobin in the Stanford Encyclopedia of Philosophy, "to say that a kind is *natural* is to say

that it corresponds to a grouping or ordering that does not depend on humans."[1] From that definition, we can loosely imagine what is captured by the term and what isn't. Rocks and planets presumably don't depend upon us, but forks and television do. These are material examples and thus more straightforward than a concept like race or ethics. *Prima facie*, it isn't a huge leap to conclude that motherhood is a natural kind. While mothers are clearly human, the state of motherhood is biological in its roots. Mothers are the genetic and gestational bearers of children. The physicality that is required to produce a child naturally makes us think that motherhood has to be biological. As many writers on this topic have noted, there is some history to this line of thought.

Aristotle thought the maternal biological role was limited to gestation. In what Caroline Whitbeck terms the "flowerpot theory", mothers were thought to only be the carriers of the fetus, without actually contributing any of their own genetic material. Aristotle thought men were the sole genetic determinants of human life. A faintly more sophisticated take on this view is a blueprint theory. Here, the same idea is at play, only the female is given slightly more than human incubator status. The idea here is that once the genetic material is in place, the fetus simply develops as programmed by the genetic code. The blueprint theory grants that the female is responsible for half of that genetic material. Yet, this is the only credit she is given. The growth of the fetus is assumed to be determined by the genetic material and the genetic material alone.

Susan Feldman argues that this fails to appreciate a very special way in which mothers are a biological kind. She argues that we should think of gestation as actually being composed of two different entities, there is a physical act of allowing one's body to provide for a growing fetus, and then there is "the work, {mental}, conscious and automatic, that a nine month gestation requires. This work, as we have seen, is an important factor in determining the actual nature, well being and state of the newborn. Pregnancy is work whose quality strongly affects the newborn" (Feldman 1999, p. 99). Pregnancy is a component of mother-

hood that is obviously a natural kind. Pregnancy is governed by biological rules that were not created by humans. Clearly human behavior is necessary for this state to emerge, but humans did not create the notion of pregnancy.

To say that a natural kind is not created by humans both seems to simplify and confuse the situation. From a semantic perspective, everything is created by us. Take the celestial object Pluto. Pluto is a referring term provided by us, by humans. Specifically, it was provided by an eleven year old English schoolgirl, who suggested the name Pluto. That term at one time referred to a planet. Now it refers to a small object in the Kuiper Belt. But 'Pluto' is not a mass of ice and rock. 'Pluto' is just a word that was invented by us. More importantly, the entire concept of 'a planet' was invented by us. So we need a more robust definition of a biological kind. Generally speaking, the following criteria, adopted from the Stanford Encyclopedia of Philosophy[2] are suggested for a kind to be classified as "natural":

1. The members of the kind share some sort of natural property.

2. Natural kinds have "inductive inference." What this means is that when we see a certain property of a natural kind, we should expect to see that it is included in the set of natural kinds.

3. Natural kinds participate in laws of nature.

4. Members of a natural kind should form a kind.

5. Natural kinds should form a hierarchy.

6. Natural kinds should be categorically distinct.

From this loose rubric, there are various ways of considering a natural kind. I will discuss three attempts to explain biological kinds: essentialism, cluster concepts, and natural kind realism. I've arranged these attempts from the strictest to the more liberal. I will then revisit this rubric and discuss how motherhood meets these criteria.

BIOLOGICAL ESSENTIALISM

A cornerstone of the motivation for believing in natural kinds is essentialism. Not all proponents of natural kinds are essentialists. But for those who are there must be an essence to a natural kind. This essence really boils down to a necessary condition, and what we are really talking about is a condition that must be present for a kind to be allowed set membership. In this case, the set is the set of natural kinds. For a kind P, it must possess some property that makes it is essential that P is included in the set Q. This property may not be evident or observable, but it must be present. Kripke and Putnam both hold views that inspire belief in the concept of biological essentialism. Kripke's view is semantic in that it questions what we mean by essence and what we think we are referring to when we claim that a kind has an essence. For Kripke, we don't need to be referring to an identifiable essence. We can refer to an essence that is not visible. For example, water necessarily has the chemical structure H_2O. The naked eye can't see these molecules, but we can still refer to an essence that we can't see. Moreover, Kripke believes that the essence exists, so he is also making a metaphysical claim about natural kinds. Putnam's argument is in the form of a thought experiment. We can imagine an Earth exactly like ours in every way; call this Earth "Twin Earth". There is a clear, odorless, drinkable fluid on Twin Earth that seems in every way to be like what we call water. Yet, when scientists examine the chemical components of this liquid, it turns out not to be H_2O but XYZ instead. If one shares Putnam's intuition that nothing that is not H_2O can be considered water, such that the stereotypical properties of water shared by XYZ are insufficient for counting XYZ as water, then one should believe in essential properties for natural kinds.

Does motherhood have an essential property? It is tempting to say yes. But what could be considered a biologically necessary property of motherhood? If there is such a property, I think the ovum is about as good as any. The ovum is a cell that contains the potential for fertilization. Note that this is not a sufficient prop-

erty. Clearly, there must be sperm and a gestational host, but the viable ovum only needs to be a necessary property to qualify as a biological essence. But what if the ovum were donated? If this seemingly essential property came from an anonymous donor, then ceteris paribus it seems the donor can be considered as the mother at least as much as the eventual host. But intuitively that does not seem right.

BIOLOGICAL KINDS AS CLUSTERED

It might seem that the view for essentialism is too strict. Yes, we are allowed to forgo sufficient conditions, but we still are required to have necessary conditions. To ease the admissions requirements a bit, some have proposed that kinds can have more than one natural property that will render them a natural kind. To make the deal even sweeter, clusters are not inert, meaning that the individual properties might change over time. According to Boyd, on this view, cluster kinds have properties that are indeterminate. In his words, natural kinds are "homeostatic property-cluster definitions" (Boyd 1999). What this means is that the properties that make a kind a *natural kind* are neither sufficient nor necessary. Boyd is suggesting that in biological or natural settings, properties will have a connection to one another that is imperfect. The properties that collectively make a kind a natural kind today may not necessarily be the same cluster properties that make the same kind a natural kind in the future. Boyd is relying on the indeterminacy of science to make a semantic point about how we can make biological realism work.

NATURALIZED NATURAL KINDS

Boyd has provided substantial wiggle room for natural kinds. But in doing so, he stipulates that the properties of natural kinds have some sort of causal relationship with other kinds, and he isn't clear about the parameters of change granted to these properties. To simplify everything, we could do away with essences and cluster

properties and just use common sense. Pretend we see a raven and it's black. We then see a thousand more ravens and they are all black. Meanwhile, we have not see a raven that is any other color. We naturally would draw the conclusion that all ravens are black. This is a very casual means of noting a property of a natural kind. Quine thinks this is exactly the way we should think about natural kinds; or rather this is how we should do science. Quine might be considered slightly inclined to throw out the idea of kinds altogether. He thinks we have an innate sense of inductive inference about the natural world and we don't need a list of criteria by which to establish how to define a natural kind. Quine not only thinks this ability is just a natural by-product of being human, he also thinks we have the ability to self-correct. So a person who sees a thousand black ravens and also sees a thousand crows might draw the conclusion that all birds are black. But someday, he sees a blue-jay and then a parrot. At this point, he will just naturally correct his view and conclude that while all ravens and all crows are black, not all birds are black. This ability to adjust our beliefs contributes to our ability to predict the future. So beyond knowing the color of a raven, our bird-watcher will be able to predict that the next time he sees a raven, the bird will be black.

Employing a Quinean-type view, motherhood seems pretty easy to cash out. We just go out in the world and we start watching. We see a female with a bulging stomach, she must be a mother. We see a woman without a functioning reproductive system, and we assume she is not a mother. We witness a woman breast-feeding, and we naturally infer that she is a mother. However, this type of inductive inference isn't always truth-preserving. In a later section I will discuss how Quine's view gets tricky with respect to determining motherhood.

MOTHERHOOD AS A BIOLOGICAL KIND

Before I began to discuss some of the problems associated with the aforementioned conceptions of natural kind, I'd like to revisit the rubric described at the onset of this discussion. I will consider

this list of criteria as a litmus test for natural kind motherhood.

1. *The members of the kind share some sort of natural property.* What would this be? It can't be some sort of reproductive biology for that would exclude adoptive mothers and count surrogates as mothers.

2. *Natural kinds have "inductive inference."* What this means is that when we see a certain property of a natural kind, we should expect to see that it is included in the set of natural kinds. By this reasoning, we would expect all pregnant women to be mothers and all women not able to conceive as non-mothers.

3. *Natural kinds participate in laws of nature.* Which laws of nature would these be? This seems to not consider things like abortion or fetal surgery.

4. *Members of a natural kind should form a kind.* Given the various scenarios of mothers I've presented, it hard to imagine all those variations forming their own kind.

5. *Natural kinds should form a hierarchy.* It might be the case that mothers do form a hierarchy, but if it is, it seems it would be of a social one and not of a biological nature.

6. *Natural kinds should be categorically distinct.* I think there are two ways to think about this. One, it seems incredibly difficult to figure out who is a mother in the first place so that she can be categorized. Secondly, even if you could, this seems as though it would follow the same pattern as the previous criterion, this would be a social category and not biological.

Problems with Natural Kinds

The entire discussion of motherhood as a natural kind has rested upon the assumption that mothers are *only* the "genetic and gestational bearers of children." Most mothers fit this criteria but this

assumption is not a litmus test for motherhood. Clearly, women can be either of those things. We do not need to imagine this, for modern science has made it a reality that women can be genetic bearers solely by donating ovum to another woman. Conversely, a woman can be a gestational bearer *only* by performing as a surrogate. Furthermore, women who have no genetic or gestational link to a child can be a mother. A woman could be both a genetic and a gestational bearer but elect to surrender the caretaking and parental responsibility of that child by offering the child up for adoption. Finally, we can imagine a scenario in which a woman who has no genetic or gestational contribution to a baby becomes the adoptive mother. That I can posit all these scenarios, and because more are viable, it seems clear that the idea of motherhood as a biological kind is seriously flawed. If motherhood were a biological kind, an adoptive mother who raises a child, provides emotional, spiritual and financial support and performs exactly the same duties as some birth mother could not be considered a mother at all. This conclusion seems questionable at best. Furthermore, as discussed previously motherhood fails to possess an essence. There isn't a property that can be identified as necessary to allow a certain kind to be admitted to the set of motherhood.

One might argue that biological essentialism is the wrong way to consider a natural kind. You might propose that such a theory is too limiting and natural kinds can have clusters of properties. So as long as all kinds in the natural category of motherhood share one of these properties, they would be considered a natural kind. But, that still doesn't admit any woman without a biological link to the child. The godparent who has absolutely no familial ties to a child, and who assumes the role of caretaker, is clearly a mother. But, how does she fit into the cluster kind? She doesn't, because she fails to possess any sort of biological property associated with motherhood.

Finally, there is the Quinean type view which is a liberal take on natural kinds. Certainly, if motherhood is going to be considered a natural kind, it needs a very liberal admission policy. It seems immediately obvious that the *prima facie* inductions about

mothers are wrong. A pregnant woman may be a surrogate, a woman without the ability to conceive might adopt and finally, wet-nurses, while no longer *au courant* are certainly a phenomenon that could lead an observer astray. The Quinean response (this is for argument's sake, as I highly doubt that Quine would have defended motherhood as a biological kind) would be to say that naturalizing natural kinds allow for these sorts of mistakes and further observation will allow for self-correction. But this makes no sense. Motherhood is such a mixed bag of biology that it is impossible to predict who is a mother and who isn't. Recall that in the initial discussion of how we define a biological kind, I mentioned that the view started out strict and became more liberal. That might have seemed enticing in that it allowed the possibility of more things admitted to the metaphysical club of biological realism. But that might actually be a bad view. Furthermore, it seems to be a bad way to do philosophy. We don't want a view of motherhood that is so loose that everything and anything that looks like a mother gets counted as a natural kind. It seems as though we are stuck between making the definition either too loose or too strict. For all of these reasons, it seems unlikely that motherhood is a biological kind at all.

MOTHERHOOD AS A SOCIAL KIND

Given the failure of natural kinds to properly capture motherhood, the next possibility is to conceive of motherhood as a social construct. In many ways, this might seem intuitively pleasing. A social construction of motherhood avoids the problems associated with biological essentialism. On this interpretation, anyone could be a mother. Birth mothers, adoptive mothers, aunts or grandparents who are responsible for primary care-giving, really any person that society wants to bestow the term 'motherhood' upon will qualify. Motherhood as social construction also alleviates some of the thorniness that arises in conjunction with newer medical procedures that complicate the question of who can be a mother. For example: egg donation, in which the genetic material

may or not be that of the eventual care-giver. The eggs may be donated in order for another woman to carry and then upon birth given back to the original egg producer. Or perhaps, the eggs are donated by someone who has no desire to participate in mother-hood, and are gestated by another woman who may or may not be the eventual care-giver. In one example of medical technology, I've identified three different women all of whom could fulfill the role of mother or could not. By using a social construction model of motherhood, we avoid the problems associated with biology forcing the hand of nomenclature, and allowing the term 'mother' to be applied to whoever ends up in that role and not to an anony-mous egg donor or a surrogate.

This approach is advantageous because it casts a wide net, but we want to be cautious in our optimism, for if the net is cast too wide, then the term is rendered meaningless. So there must be some criteria that society uses for its attachment of the term 'mother' to a particular woman. We need some guidelines, so while women who are mothers under non-birth circumstances are awarded their proper semantic status, we also don't end up calling anyone with an operative uterus or a nurturing hand a 'mother'. Furthermore, our social construction must be able to navigate cases where a woman who may contribute genetic or gestational material is not called mother if she doesn't want to be. A social constructivist version of motherhood should preserve the status of women and mothers given each unique and potentially complex biological and logistical situation that arises in birthing and caring for a child.

If the following test can be met, then motherhood can safely be considered a social construction. If it can't, then we must admit that motherhood fails as a social construction. To begin, we need to unpack the concept of "mother-like duties" that I've been using. Because it is a socially constructed definition that we're after, the concept "mother-like duties" can be anything that we dream up. Unlike biological essentialism, we are not bound to any natural property. We might come up with some seemingly socially accepted notions of what it means to be a mother. A 'mother'

meets all of the following four criteria (she may range across the spectrum of each criterion, but at some level she meets the bare minimum):

1. A formal initiation with the child. This could be birth or adoption or some other type of indication that a particular woman is to assume the role of 'mother'. The term 'formal' denotes formality between mother and child and does not suggest that the initiation be sanctioned by law. We can easily imagine a case where a child is taken care of by another family member without official government sanction. This is not to suggest that an initiation without a child's consent or under illicit circumstances meets this criterion.[3]

2. A vested interest in the guidance of a child's upbringing. The mother both declaratively and privately wishes to guide the child through infancy and childhood without malevolence.

3. A certain amount of resources are devoted to the child. As resource possession differs from mother to mother, this could be financial, temporal, mental, emotional, etc. The central point is that some personal resources of the mother are given up for the child.

4. Intent to pursue a life-long relationship with the child. The notion is that the mother is interested in and committed to a history and a future of a personal relationship with the child.

We can imagine that these guidelines are easily translated amongst cultures and societies. These criteria are not specific to one society. It certainly seems as though these are norms that have been practiced for as long as societies have been in the business of raising children. This social constructivist view is significantly more open than the biological view in that it allows for more flexibility. It is also more descriptive. Instead of pointing to one essential or cluster of properties, we can more fully tell the story of motherhood.

Now, the question remains, do these guidelines actually do the work that we want from our social theory? Let's revisit what

we wanted from a social constructivist account:

1. Consider a woman a mother even if she is not the birth-mother, but performs as one.

2. Consider a woman not to be a mother who performs "mother-like" duties but is not a mother.

3. Consider a woman not to be a mother if she does perform "mother-like" duties but does not wish to be called a mother.

4. Preserve the status of the average women who gives birth and performs the role of mother to that child.

To some degree my social construction project has succeeded and to some degree it has failed. It certainly seems as though the positive account is protected. It seems pretty straightforward that the average woman who looks like a mother and acts like a mother, will be considered to be a mother by society. What about those women who perform some of those roles or variations of those roles and do not actually have dependents, who are not actually mothers? This negative case does not seem to be addressed. Furthermore, it seems as though it would be terribly difficult to attempt to address this problem and formulate a way in which non-mothers are semantically protected in a social construction.

PROBLEMS WITH SOCIAL KINDS

Before dissecting some of the issues associated with motherhood as a social kind, we should re-examine the concept of a social kind. As Haslanger notes, social construction is a relationship between the constructor and the constructed. X constructs Y. In this case, society is constructing motherhood. Motherhood is not an object or an artifact, but an idea. It is an idea constructed by society. But social constructivism isn't always right just because it has more flexibility than biological realism. Certainly the term, 'mother' carries enormous social weight. One might even argue that it carries social pressure. But a social constructionist account could clearly be blind to some parts biology that are *real*. The

problem with motherhood as a social kind lies more with the notion of social construction than with motherhood itself.

Haslanger acknowledges the problems with social constructivism when she says "…we must be attentive to the possibility that the terms we use are defined by and in the interest of dominant social groups. But from this it does not follow that the only function of judgment is the social one of perpetuating useful stories or that our point of view on the world is always socially conditioned; but there is no reason to conclude that the world we have a point of view on is likewise socially conditioned" (Haslanger 1995). From this, it seems that as much as we might have the right intentions in mind, we might be misguided in believing that social constructivism is the correct route to motherhood.

PROBLEMS WITH THE DICHOTOMY

Now, we turn to the question of whether we choose one of our faulty explanations or construct a different interpretation altogether by viewing motherhood as an altogether different kind, neither biological nor social (or possibly a combination of both). Biological realism runs the risk of reducing women to baby-making machines and also gives rise to the worry that the responsibility of the child-producing body is too great. We may begin to think that we should legislate that body and regulate it. We see this in the form of restrictions on women's choices to terminate or keep a pregnancy, who to offer fertility treatments to, and who is offered pre-natal care.

Motherhood as a social construction has its own set of problems. One, it ignores the significance of the biological contribution made by mothers. More importantly, social constructions of motherhood tend to be normative, meaning we start talking about how mothers "ought" to be. As I've shown, there are so many different ways that a woman become a mother, it hard to imagine that a normative account of motherhood would ever hold up. A third problem arises in that we are never consistent in our view. Society wavers between considering mothers as biological kinds

and sometimes we think of them as social kinds.

These problems hint at a sort of collective confusion. Yet, I don't believe it is the answer that is confused; I think it is the question. When I asked the question, *Is motherhood socially or biologically constructed?*, I was asking not so much the wrong question, but a question that is bound to set the answer up to fail. My original question seems to generate one of the following three answers:

1. Motherhood is both biologically real and socially constructed.

2. Motherhood is neither biologically nor socially real, but something else entirely.

3. Motherhood is an example of supervenience. The social supervenes on the biological.

It might seem that the proper response would be 1, given the data I've presented on the various ways in which motherhood seems to be realized both biologically and socially, and fails to stand up to scrutiny as solely one or the other. But 1 fails to capture something that I think is important. It isn't just that motherhood has both of these components, biologically real and socially constructed. When we say that an entity has both A and B, it seems that we are suggesting A and B are wholly discrete parts that happen to be part of the sum total of something else, let's call C. In this case, the biology of motherhood is not just a separate construction from the social piece. Rather, they seem to coalesce, separate, and then intermingle. These two conceptions are not always distinct. The way in which I want to capture the nature of motherhood must preserve this unique conflation of biological reality and social construction.

I've shown that motherhood is not viable as a strictly biological or social kind. I've argued that no account of either view will work. I've also shown that there are elements of both that are critical to understanding the concept of motherhood. At this point, there can only be some sort of hybrid theory or supervenience. I

think supervenience is a more accurate picture. The type of super-venience I propose looks like this: the social properties of mother-hood supervene upon the biological properties. An exact reading of supervenience would entail that any mother who has a social property has some biological property such that any person with that biological property has that social property. But I don't have this sort of precision in mind. I would propose we entertain a type of naturalized supervenience. By this I mean that motherhood is a relationship of properties; social properties that are of a hierarchal nature, but still dependent upon the biological properties and that the indeterminacy of the variances in these property relationships has some degree of naturalism to it. The naturalism that I have in mind is intuited by the notion of inductive inference or as Justice Stevens said much more succinctly, "I know it when I see it." I imagine he knew a mother when he saw one.

To say that something is a natural kind is really just to say that it has evolutionary properties. If we change those properties, then we have just changed the very nature of the kind itself. Through this process, we have become a social agent of change, which means even more that the social supervenes on the biological. It might sound odd to suggest that we can change evolutionary properties, but in the case of motherhood, that is exactly what we have done. Modern medicine has changed the biological rules as to who can be a mother. But, *what* is modern medicine? A better question might be *who* is modern medicine? The answer is *us*. We are. Suddenly it becomes very difficult to imagine motherhood as a natural kind that exists outside of humans. Such a scenario is certainly impossible in today's world, although I seriously doubt it ever was possible.

Notes

1. http://plato.stanford.edu/entries/natural-kinds/
2. http://plato.stanford.edu/entries/natural-kinds/
3. Note that birth is one option for the initiation event. There are some who would argue that birth while under the influence of medication, or non-vaginal birth is not really giving birth. This is a classic example of how value

judgment based upon a biological assumption of motherhood dictates who is a "real" mother and who is not.

Bibliography

Bird, Alexander & Emma Tobin (2008) *Natural Kinds* (Stanford Encyclopedia of Philosophy) http://plato.stanford.edu/entries/natural-kinds/

Boyd, Richard (1988) "How to be a moral realist," in G. Sayre-McCord (Ed), *Essays on Moral Realism*, pp. 181-228 (Ithaca: Cornell University Press)

Donchin, Anne (2004) *Feminine Bioethics* (Stanford Encyclopedia of Philosophy) http://plato.stanford.edu/entries/feminist-bioethics/

Feldman, Susan (1992) "Multiple Biological Mothers: The Case for Gestation," in *Journal of Social Philosophy*, pp. 98-104

Haslanger, Sally (1995) "Ontology and Social Construction," in *Philosophical Topics* 23:2 pp. 95-125

Quine, W.V.O. (1969) "Natural Kinds," in *Ontology and Other Essays*, pp. 114-138 (New York: Columbia University Press)

Assessing the Use, Usefulness, and Justifiability of Moral Intuitions

Theresa Cheng

John Kekes articulates what he considers the primary mechanism for acting in accordance with our moral tradition: our moral intuitions. He characterizes and defines moral intuitions, specifying that they should only be considered useful in a limited way—to govern our everyday actions. But moral intuitions are not the sole motivation for our everyday actions, even those sanctioned by our moral tradition. I argue that people do not rely as heavily on moral intuitions to perform everyday actions as Kekes suggests. Instead, people perform everyday actions for diverse reasons, including non-intuitions and non-moral intuitions. I also comment on Kekes' views on the fallibility and justifiability of moral intuitions, ultimately arguing that Kekes' criteria of justifiability of moral intuitions, the soundness of the moral tradition from which the moral intuitions are derived, is problematic.

The purpose of developing a view which accurately reflects the degree to which moral intuitions are used, useful, and justified is to contribute to a greater meta-ethical discussion; articulating and assessing Kekes' view of moral intuitions serves as one way to verify or reject the claim that moral intuitions and the moral tradition from which they are derived are the central part of moral life, and therefore represent the appropriate primary area of interest for moral theorists.

Kekes describes moral intuitions as sudden realizations not requiring conscious reflection (Kekes 1986, p. 84). He characterizes intuitions as immediate, and thinks that they are distinguishable from other kinds of thought because their conclusions are derived in the absence of deliberation or reflection. There are two theoretical steps involved in intuiting something: (1) identifying

relevant facts about the situation, and (2) interpreting the pattern of facts in a particular way (Kekes 1986, p. 86). Although later I discuss some ways in which Kekes' definition of a moral intuition is flawed, he and I could agree on some examples of intuition in general.[1] A cumulative difference in a friend's behavior or mannerisms might lead a person to intuit that the friend is sad, or angry, or under the influence of stress, without conscious analysis from the person or explicit statement from the friend. Another case is when those with some training in logic or mathematics intuit that something is wrong with a proof before having clearly identified and demonstrated the error. These cases also demonstrate that intuitions may not ultimately be true or justified—the friend may not be upset, and the proof may be correct.

With regard to what makes a moral intuition moral, Kekes specifies that he follows a more or less utilitarian view of the issue by defining "moral" in terms of human benefit and loss caused by human agency.[2] He views moral intuition as a clear subdivision of intuition involving immediate moral insight proceeding from non-moral information about a situation, for "…what we intuit is not a moral fact over and above non-moral facts, but an interpretation of non-moral facts" (Kekes 1986, p. 86). Moral intuitions are unique because unlike sensory, aesthetic, or other types of intuitions, they are strongly tied to action; we experience them as imperatives which motivate specific actions, though the actions might not be carried out. Kekes also states that moral intuitions strongly influence behavior, and tend to be accepted unquestioningly by the person experiencing them.

In defining "moral intuition," Kekes pits himself against old-style intuitionists who believe that moral intuitions are a largely infallible way to access general moral principles that determine the correct response to a given situation. According to him, old-style intuitionists are moral realists, i.e., they believe there are moral facts that can be known, and that moral intuitions are the means to accessing them. Kekes argues for moral cognitivism, the belief that we hold rational moral intuitions. While moral realism is a kind of moral cognitivism, not all moral cognitivists are moral

realists. He is the latter type of moral cognitivist, and strives to establish the rationality of moral intuitions in order to strengthen the moral cognitivist perspective.

Even so, Kekes acknowledges that moral intuitions can seem irrational. For him, this irrationality is only apparent because moral intuitions can be explained in rational terms within a moral tradition. He thinks that there are three different components of intuition: the cognitive, emotional, and volitional. Kekes argues that there is evidence that the cognitive component of moral intuitions is stable, because rational beings within a moral tradition frequently reach the same conclusions. This is cognitive not because the moral tradition is devoid of emotional components, but because there is a "right" answer within a moral tradition that we are purportedly able to access. He states that the fallibility of moral intuitions stems from the emotional and volitional components, which are affected by factors like an individual's past experiences and personality. However, he thinks that even the emotional and volitional components can be explained in rational terms, not because they accurately represent independently existing moral facts about the world, but because they are open to justification and criticism. This is consistent with his assertion that moral intuitions are interpretive rather than descriptive, and presumptive rather than completely reliable.

Because of their fallibility, Kekes thinks moral intuitions should be used in a limited way. Rather than relying on them in all situations, their use is justified only in situations which are uncontroversial. He says that moral intuitions are primarily used and useful in responding to routine situations, like paying bills, and being decent and courteous to others. Using moral intuitions in this limited way is meaningful because it grounds everyday actions in rationality, establishing their meaningfulness against the claim that they are only arbitrary.

Having appropriate moral intuition presupposes our education in a moral tradition. Thus, moral intuitions are only justified insofar as the moral tradition on which they depend on is justified. Kekes' justification of a moral tradition is, again, utilitarian:

a moral tradition is justified if it encourages conduct which maximizes the goods and minimizes the harms for the group of people to which it belongs. Moral traditions help make possible forms of a good life by providing certain goods, including a means for self-direction, intimate relationships, and decency for its members. A sound moral tradition prevents chaos by facilitating interaction with others which would otherwise be difficult. Because it provides goods and promotes social harmony, a moral tradition "is not an enemy of individuality, but a necessary condition of its development" (Kekes 1986, p. 90). A moral tradition can be unintrusive, Kekes argues, because human activities range from the private to the public. On one end of the spectrum are actions of personal significance, whereas the other end involves actions with great public/social significance. Sound moral traditions, Kekes says, have the least impact on our lives at the personal end, and the greatest impact at the impersonal end. While sound moral traditions generate a plurality of personal life paths within society, they tolerate few differences of opinion at the social end.

As a result of pluralism, there is individuality, which breeds conflict. Conflict, too, is considered a product of a healthy moral tradition. In cases of conflict, Kekes recommends that moral intuitions cease to be useful. He asserts that we will know when our moral intuitions have gone astray, because if they have, we will encounter people who challenge them successfully. From that point on, he thinks we should use reasoning, reflection, and rational deliberation, leaving the realm of moral intuition entirely.

In Kekes' view, people fail to interpret situations in accordance with their moral tradition when selfishness causes them to misrepresent others' interests, values, or intentions. Kekes claims that our tendency toward self-centeredness is a major cause of the fallibility of moral intuitions. But this can be remedied. We can realize our error, because we come into contact with other members of our moral community who will disagree with selfish intuitions. Once aware of this, he believes that we should sort through the conflict using rational thought rather than continuing to rely on our intuitions. Selfishness and other factors which create

flawed moral intuitions can then be corrected by dialogue with others from within the moral tradition.

In sum, Kekes believes that moral intuitions are immediate, routine, interpretive, forceful, rational, and presumptive. He drives a middle path between old-style intuitionists who believe that moral intuitions are a means to accessing moral truth, and others who believe that moral intuitions are utterly irrational, hopeless, and have no place in our lives. Tying moral intuitions to our moral tradition, he argues, justifies using moral intuitions in a limited manner to serve as useful motivators of everyday actions.

Kekes argues that moral intuitions are at the core of our moral lives because they drive everyday actions. Recall that Kekes thinks that decency, common courtesy, and paying our bills, are among the everyday actions which are motivated by moral intuition. He argues that moral intuitions are the primary mechanism by which moral tradition guides everyday action. I will argue that he overstates the use, and usefulness of moral intuitions for two reasons. First, because moral intuitions are meaningful beliefs, it cannot be the case that moral intuitions motivate us to perform the majority of our everyday actions, because many everyday actions are not motivated in this way. When I use the term "non-intuitions," I am referring to the class of motivations which are either non-spontaneous, non-meaningful, or are not held as beliefs. Second, I think that immediate moral concerns are not the chief motivator of our everyday actions, but are one factor among many. When I use the term "non-moral intuitions," I am referring to the class of intuitions which are not moral because the intuition does not involve an evaluation of good or harm to humans. By weakening the relationship between moral intuitions and everyday actions, I mean to clarify when moral intuitions are used and useful from when they are not, as well as to indirectly attack Kekes' claim that moral intuitions are the primary mechanism which mediates between the moral tradition and everyday action.

If we have good reasons for thinking that moral intuitions must be meaningful beliefs, then moral intuitions cannot also be the primary motivation for the majority of our everyday actions,

even those sanctioned by our moral tradition. This is because we would then hold the unlikely view that people are driven to perform the majority of their everyday actions through spontaneous meaningful (and moral) beliefs, to the exclusion of other motivations. Such beliefs may take the form of recognition of a previous belief, or newfound realization. To clarify, I do not mean that we must be conscious of the logical steps leading up the belief, nor do I mean that it necessarily takes the form of an explicit, verbal statement. So, an experience like awareness of uneasiness in response to a disturbing situation could count as a meaningful belief.

Why should we think that moral intuitions are meaningful beliefs? Intuitive responses differ in some way from physiological reflexes. When your health care provider taps on your knee and it rises, we do not say that this was the result of intuition. Intuitions can, of course, propel us to take certain actions, e.g., when you have the intuition that you have lost or forgotten something, you might be motivated to retrace your steps, but this is not strictly required.

Furthermore, if the claims made by moral intuition are believed by the person experiencing them, they have content and differ from gestures which are not meant to indicate anything, as well as habitual statements which are not sincere. This is what I mean when I say that moral intuitions are meaningful. For instance, let's say that you immediately reply to the inquiry "How are you?" by saying "Good." You might have rapidly assessed your state of being, and responded accordingly, or you may have merely responded in accordance with the appropriate response of our moral tradition. If you do not hold a sincere belief about your state of being, then you have not had an intuition about it and therefore cannot actually be said to believe that you are good. In the latter case, it is possible that you have had a moral intuition, and also possible that you have not. If you were motivated to respond because of a belief that you should not burden others with your problems, then Kekes and I would agree that you have had a moral intuition. But if you habitually responded to the situation, then the response did not involve a moral intuition because

there is an absence of meaningful belief. In this explanation, everyday actions can be motivated by moral intuitions, but often they are not. Analysis of this scenario also demonstrates that while people sometimes perform actions because they have had a moral intuition, there are a wide variety of legitimate motivations for a person in a moral tradition to respond to a given situation.

Kekes' key claim is that moral intuitions actually serve as the primary motivator of everyday actions. His main method of arguing this is by pointing out the existence of various everyday situations in which members of a moral tradition proceed in an inevitable way which make sense to other members of the tradition. However, I doubt that the identification of common motivators during the retrospective rationalization of our behaviors serve, in the moment, as our actual motivation. The experiences of followers of self-help advice highlight the difference between the way we are usually motivated to act, and the experience of somebody who is spurred to action by a meaningful belief. The self-help industry is constantly encouraging people to improve their lives through changing their actions. The fact that people struggle to implement recommendations about productivity, relationships, and resolving addictions suggests that people need to work to connect their beliefs to their everyday actions, i.e., that this is not our typical mode of behavior. However, once we have incorporated good practices into our lives, then they become easy, effortless, habitual, and may no longer be driven, in the moment, by meaningful beliefs.

It may be the case that in some instances our everyday actions are motivated by spontaneous moral beliefs. However, even when we are motivated to perform everyday actions by beliefs, these beliefs may not be moral. Kekes is not meticulous about what qualifies as moral and what does not, but believes that moral intuitions are interpretations of situations with human harm or benefit in mind. So in analyzing whether a belief is moral, I will allow for a wide range of possibilities regarding what might make it so. By moral it might mean that the content of the belief involves a value judgment or evaluation of the action, a moral emotion, or

the recognition of the moral significance of the action's consequences.[3] Sometimes we are motivated by these considerations, but at other times this is not the case. For instance, it is possible to be strongly motivated to vote because of a belief with moral content, such as the belief that you are a valuable participating member of a democracy. But it is also possible that your motivation to vote is to escape the negative stimulus of a nagging spouse. Clearly, we can be motivated to perform seemingly moral actions by simple conditioning, pragmatic reasons, peer pressure, the desire to appear a certain way, or a mixture of all of these reasons and others. Indeed, it might even be possible to have an immoral motivation for performing an action which is widely perceived as moral.[4] Again, the widespread ability to identify the morally acceptable reason for performing an action in hindsight is not evidence that we were actually motivated in this way.

Finally, Kekes might think that everyday actions are motivated by something moral because if somebody were to tell us that we were behaving wrongly, we might have a strong reaction. In the face of opposition, it might become clear that we are very committed to the action that we have performed. In this case, we might immediately think that our action was right. We might also have, upon further reflection, some pretty good reasons for thinking that we are right. But the criteria of having been motivated by an intuition ought to be a discussion of actual motivation rather than a discussion of retrospective explanation.[5]

I hope to clarify my two objections through a final illustration. Students' lives are seriously affected by classroom attendance and the tasks of completing reading, homework, assignments, and projects. For students, these activities constitute a significant part of their everyday existence. Some students are motivated to go to class for a variety of morally interesting reasons. They might feel compelled to gain basic skills to become a productive member of society, and believe that education is a good way to achieve this end. My first objection to Kekes suggests that students, even those with this objective in mind, are probably not driven to perform individual acts of classroom attendance through moral intuitions

about classroom attendance, because attending class becomes a habitual practice. In my second objection to Kekes, I mean to point out that students might not have any moral interests in mind when pursuing any particular action, or they may have several motivations, some moral and others not.

Essentially, I have not denied that moral intuitions might be useful, but I think that their use in everyday situations is more limited and less uniform than Kekes suggests. So far, we can modify his description of moral intuitions by being more specific about what counts as a moral intuition and what does not, in a way that directly responds to my comments. First, it is necessary, but not sufficient, that moral intuitions engage meaningful beliefs. Second, I think that the content of the belief should be moral, and not merely its possible extended implications, or something else. While moral intuitions still occur and might motivate some everyday actions, they cannot be cited as the clear underlying motivation for a majority of everyday actions. This greatly weakened relationship between moral intuitions and everyday actions casts doubt on Kekes' claim that moral intuitions are the primary mechanism by which moral traditions govern everyday actions.

Perhaps Kekes would say that I am missing the point: what's interesting to him about moral intuitions is that they explain the profound way our moral tradition shapes our everyday activities. He believes that moral intuitions are important because they successfully maintain order and a well-functioning society. Are Kekes and I are just having a quibble over words? I want to be more specific than Kekes about what qualifies as moral intuition, and point out other factors involved in motivating everyday action. But because he pits himself against old-style intuitionists, he is clear in his desire to establish a view of intuition that is in lieu of what they have offered. He is not merely describing a phenomenon and then positing that we call it a moral intuition, because he is making a claim about what a moral intuition is. This makes it reasonable for us to compare his concept with our own thoughts on what a moral intuition is.

If it is habits that motivate our actions in everyday situa-

tions, then I propose that we use moral intuitions in novel, non-routine situations. Again, even within a moral tradition, motivations for actions vary, and I am opposed to overgeneralization: not all everyday actions are motivated by habits only, and not all reactions to novel situations are motivated by moral intuitions. However, I think that clarifying when moral intuitions are used is an important step in asking various other questions about them. This is a challenge to Kekes' terminology rather than to his view, because I think that he mistakenly lumps routine and non-routine situations together. Though he describes many actions as routine, some of the situations he presents might be better described as novel. Early on in his paper, Kekes provides four examples of routine moral intuition at work. These include helping an elderly person who has stumbled, experiencing outrage from witnessing a coworker lie, paying a bill, and feeling regret upon discovering that a former mentee committed suicide. Earlier, I pointed out that the instance of paying a bill might not be the result of a moral intuition. Regarding the other examples, I am puzzled as to why we would call them routine. For most people, these events do not occur on an everyday basis. One reason that Kekes calls them routine is to contrast them with difficult moral situations, including unlikely conundrums that philosophers have thought up specifically to generate moral conflict, (e.g. divided train tracks, sinking ships, and the like). But this way of distinguishing between routine and non-routine is not helpful in understanding what a moral intuition might be if all that is signified is that some situations generate difficult moral decisions, and some do not. Kekes does indeed suggest that this is one of the major differences between the routine and non-routine when he says, "Thus we are led to another property of intuitions: they occur *routinely*…. They are, therefore, not to be sought in cases of moral conflict, as many philosophers suppose, but in the innumerable spontaneous acts of the morality of every day life…." (Kekes 1986, p. 85). Kekes' claim might even be fairly weak: moral intuitions occur *at least* routinely. But it does not follow that they should thereby be sought out in these situations. He argues that routine situations in which

the use of moral intuition is justified are unlike non-routine situations, in which issues such as self-direction and intimacy cries out for extended reflection.

I propose that moral intuitions are used and useful in intermediate or simple novel situations. These situations are frequently social. Again, the situation or action itself is not able to provide an explanation of motive, although it can suggest a probable one. Simple novel situations are those which cannot be adequately attended to with sheer habit, situations in which it would be difficult to claim that the impetus for action was the fact that the action is what we did yesterday. But simple novel situations also do not cry out for the deliberate application of our problem solving abilities. In situations of complexity, which tend to provoke conflicting intuitions within and between individuals, moral intuitions might be used, but they may be less useful. When we encounter simple novel situations, we act rapidly, but also consciously and with meaningful intent, because the type of situation tends to provoke us to have a basic level of mental engagement with the situation or its outcomes.

Finally, we reach the issue of evaluating moral intuitions. There are two standards which I think can be used to do so: whether they are fallible, and whether their use is justified. Some argue that fallibility is enough to dismiss moral intuitions. Jonathon Baron argues that there is a parallel between the fallibility of intuitive theories about the world and the fallibility of moral intuitions. Psychologists have discovered ways in which intuitions are systematically incorrect: Baron cites failed naïve explanations for phenomena which are perfectly plausible but also patently false (1995, p. 37). In one study, students said that seasonal differences in temperature could be explained by saying that the distance from the earth to the sun was shorter in the summer and longer in the winter, when in actuality seasonal differences are a result of the tilt of the earth's axis.

Another type of problematic type of intuitive thinking mentioned is the use of heuristics, or speculative problem solving (Baron 1995, p. 37). These include phenomena like the status-quo

effect, in which people are more likely to think of the status quo as valuable and work to preserve it, failing to rationally pursue better outcomes or recognize that certain outcomes are identical. The status-quo effect may impact both our moral and non-moral intuitions. Baron points out that it may be the source of our justification of intuitively distinguishing between harm caused by action (changing status quo) versus omission (maintaining status quo).

Baron's argument for the fallibility of moral intuition relies on the idea that if some kinds of intuition are fallible, then we should also be worried that moral intuitions are also fallible. Baron's work suggests that moral intuitions might be fallible, but only indirectly. He does not provide sufficient reason to believe that moral intuitions are fallible, because it is possible that there are different kinds of intuitions, and one way moral intuitions could differ from these other kinds is by being infallible. Furthermore, Kekes agrees that moral intuitions are fallible. For him the more pressing question is whether moral intuitions can be considered rational. Baron assumes that if moral intuitions are fallible, then we should discard them. However, Kekes argues that moral intuitions may be important if we can establish their rationality, because we take rationality seriously even though it is not infallible.

According to Kekes, the test of the justifiability of moral intuitions is the soundness of the moral tradition from which they are derived. He thinks justified moral intuitions will emerge from a sound moral tradition. He sketches out the characteristics of such a sound moral tradition: facilitation of individual self-direction, space to develop intimate personal relationships, and the absence of imposition on personal life choices. Recall that Kekes specifically proposes that there is a continuum of human actions ranging from the personal to the impersonal. This spectrum serves to demonstrate that a moral tradition may be highly influential at the impersonal end, while leaving sufficient room for self-determination on the personal end. And Kekes is right; within the same moral tradition, we are usually able to meet new people daily with little friction, and we have complex social expectations for

certain situations that people understand well. Fulfillment of these obligations facilitates social stability and allows us to be productive human beings. Our moral tradition permeates our lives, and moral intuitions are one way that it does so, by regulating simple novel social exchanges. We have already questioned the idea that moral intuitions are the primary mechanism which controls our everyday actions. But a moral tradition might still influence our behavior through other means, like by determining the types of moral commitments, moral reasoning, and "moral" habits that we develop.

If Kekes' explanation of the way that a sound moral tradition impacts human lives is false, then it detracts from his argument because it would be impossible for his idea of a sound moral tradition to exist. However, he may have a weaker claim about what the soundness of moral tradition relies on. This would be the claim that the soundness of a moral tradition is due to the provision of goods resulting from the moral tradition, as compared to a smaller amount of harm. He states, "The justification of a moral tradition is that it fosters conduct that leads to a favorable balance of benefit over harm for its members as possible given the context" (Kekes 1986, p. 89). Kekes' concept of the soundness of moral traditions and his subsequent reasons for justifying the resulting moral intuitions is systematically problematic.

It appears that Kekes thinks that our moral tradition (whatever that may really be—presumably the predominant, modern, Western tradition) is sound, and that our resulting moral intuitions are justified. In dense urban populations, it seems as though we move across moral traditions from one ethnic neighborhood to the next. While some behavioral expectations stay constant, the type of behavior (ranging from major moral controversies to those matters which are simple, routine, and everyday) which is considered acceptable changes with geographic location, country of origin, between households, etc., even within Los Angeles County. Kekes doesn't say that intuitions need to be universal, but only that they need to be widely accepted within a moral tradition in a manner which facilitates positive interactions among those

who might come into contact with one another. He is committed to the view that major monolithic moral traditions can be identified, but it is not at all clear that this is the best representation of moral traditions today.

Second, if a moral tradition were to provide more good than harm overall, moral intuitions in accordance with the moral tradition may be unjustified, even on utilitarian grounds, because the resulting moral intuitions might be from a part of the moral tradition which unnecessarily perpetuates harms. This is because those intuitions which result from the parts of the moral tradition which provide goods or more good than harm would be justified, while the parts of the moral tradition which inflict harms or more harm than good are unjustified. And these latter parts of the moral tradition which are both indefensible and harmful must be admitted to exist, unless it can be said that we have a pristine and perfect moral tradition. This argument is a challenge to Kekes' claim that moral intuitions are justified. It points out the possibility of bad moral intuitions, not because of some deficiency on the part of the individual, but due to systematic flaws in the structure of the moral tradition as a whole. However, this is an argument against the justifiability of particular intuitions. Kekes might respond that the use of moral intuitions in general remains justified. I think that this would be true if unjustifiable moral intuitions resulting from imperfect moral traditions were truly negligible, but moral traditions are not benign providers of goods. The moral tradition has historically included expectations of personal behavior which have been extremely limiting. In particular, I am thinking about what common social decency entailed during the Jim Crow era, but this can also be seen through a discussion of the evolution of civil rights, as well as in remaining prejudices against any number of groups which exist today. In these cases, expectations about personal and social behavior were significantly intertwined and mutually enforcing. Also, those practices were informed historically, that is, they were affected by past iterations of the moral tradition. My argument is not that some moral traditions have been unsound. Kekes would respond that moral traditions can and have

been corrected through dialogue, and that we can recognize faulty intuitions when they are called into question. Instead, I want to cast doubt on the structural possibility of a moral tradition which provides social expectations and goods without affecting personal choices, because of the very nature of the way that culture establishes behavioral norms, i.e., the degree to which attitudes in our moral tradition determine the lives of its members.

Third, tenets of moral traditions which are the source of greater good can also entail negative consequences. This is different from Kekes' view, in which it is implied that there are good and bad parts of moral traditions which can be isolated from one another, and then sifted through so as to pick out the bad and preserve the good. However, tenets of moral traditions are more complicated and multi-faceted than this view allows. For instance, many arguably sound moral traditions emphasize a celebration of the group to which its members belong, and so encourage expressions of kindness and good will to those who are recognized as one of the group in some way. This tradition makes good sense when it comes to treating our families and those close to us with great significance, or uniting us with other citizens in pride or reverence. Conversely, this same tradition may encourage a different code of behavior toward those who don't belong or who are perceived to live on the fringes. The resulting behavior may merely be breeches of etiquette towards the so-called out-group, but it may also be more profound: in the study of violence, the permissibility of a different code of behavior with regard to an identified out-group can be a part of internalized justification for gross injustices. Thus, the portrayal of the central struggle in ethics as "society's moral traditions versus individual selfishness" is utterly inadequate, partly because the American moral tradition, as it is commonly identified, encourages special consideration for the groups to which we belong (as well as individualism: the special consideration of ourselves).

While I have discussed several problems with Kekes' way of justifying moral intuitions, I think that it is neither possible nor necessarily desirable to eliminate their use. Having moral

intuitions seems inevitable, not because they are ubiquitous, but because they are used particularly, as I have proposed, as motivation for responding to novel situations. However, they are also sources of moral conflict. Kekes is optimistic about the potential of individuals to mentally distance themselves from intuitions and instead use reasoning and reflection when encountering such conflict. Given the strength of moral intuitions, I doubt that for most people this would be a discussion which results in profound changes in behavior or thought. Instead, the conversations that follow might be more like a description or systematization of existing intuitions, which—although not a bad discussion to be having—highlights the difficulty in divorcing intuition from reasoning or reflection. Indeed, Baron laments that philosophers, too, have fallen into merely describing and systematizing existing intuitions. Baron advocates that we reject altogether the use of moral intuitions as primary data in constructing moral theories. I am uncertain about this possibility. However, I think there is an important purpose for moral theories which provide novel, cross-cultural ways of thinking about morality and address the ever-pressing need to reach understanding considering the challenges personal, social, or global conflict. Thus, I think Kekes' recommendation to shift from conflict-centered moral discussions to the morality of everyday life (within our own moral tradition) is an inadequate approach if we think that one purpose of moral theory is to help us address real world problems in which we navigate and negotiate between two or more moral traditions. But importantly, deciding whether conflict or agreement is the most important part of moral theory seems like a false dilemma: understanding the motivation behind human behavior at the level of everyday decision-making could provide insight into better understanding sources of conflict, and vice versa.

Notes

1. In these early examples, I consider what speakers of American English would consider cases of intuition, as opposed to a technical or psychological definition.

2. I assess Kekes' views from a utilitarian standpoint because he himself argues on utilitarian grounds. I am not endorsing utilitarianism in particular.

3. We may also consider the possibility that the implications of the action rather than something about our experience of the intuition make it moral. This is the view that no matter what you experience as motivating your everyday actions, they are moral if they result in human benefit or harm, whatever the quantity. This is not Kekes' view. If this were the case, then too many realizations become moral. For instance, when you consider driving a vehicle, there are possible harms that emerge from failure to distinguish between red and green, but we don't think of being able to distinguish between red and green as moral.

4. For example, someone might vote right after committing a crime in an attempt to establish an alibi. (This example was provided by my reviewer.)

5. Utilitarian theories, as opposed to deontological ones, tend to separate action and motivation/intent, i.e., an action can be moral even if its motivation is not. I do not take a stance on this issue here. Instead I distinguish motivation from how an action is perceived by others, and also from retrospective explanation of ones' own motivations.

Bibliography

Kekes, John. (1986) "Moral Intuition," *American Philosophical Quarterly* 23.1, pp. 83-93

Baron, Jonathon. (Spring 1995) "A Psychological View of Moral Intuition," *The Harvard Review of Philosophy*, pp. 36-40

Context is Everything: A Brief Examination of Ethical Contextualism

Phil Lollar

In addition to being one of the major areas of philosophical study, ethics can also be defined as how we live with and behave toward others, society, and ourselves. Each of us makes ethical decisions and choices on a daily, and even hourly, basis. That something so central to our humanity should be talked about and understood by all seems obvious. And yet, the great ethical concepts and constructs—especially those of Occidental culture—are rarely discussed and scarcely understood outside the realms of ivy-covered university philosophy departments, austere law firms and courts of law, or some political think tanks. It does humanity as a whole little good for concepts so essential to its well-being, and perhaps very survival, to be the exclusive province of only a few.

However, if we accept the above definition of ethics, and take on the mantle of teaching and disseminating the great Western ethical constructs to the general populace, it does not take long before we run into a problem. These great ethical theories quite often conflict with or even directly contradict each other. So when counseling those who face moral dilemmas about how they are to live with and behave toward others, society, and themselves, or when we face such situations in our own lives, which ethical construct should we employ to help us make the best decisions?

The answer, I believe, lies in a concept called ethical contextualism. This essay briefly examines how contextualism may be developed as a method of teaching ethical theories to individuals unfamiliar with philosophy. I will postulate how ethical contextualism may be used to overcome the contradictions and conflicts that arise when ethicists consider and compare normative theories—such as Kant's categorical imperative and Mill's principle

of utility—about good/bad, and right/wrong actions. I will also show how ethical contextualism differs from moral relativism and briefly argue against claims made by philosophers such as Mark Timmons that ethical contextualism need not be metaphysically foundational. I will argue that ethical contextualism is grounded in a concept that is metaphysically necessary to all humans, and from which we derive our ethical constructs, namely, the concept of belief.

To begin, some definitions are in order. For the purposes of this paper, let us use A. W. Price's definition of contextualism as the view that "emphasizes the context in which an action, utterance, or expression occurs, and argues that that action, utterance, or expression can only be understood relative to that context."[1] But recall that we have defined ethics as how we live with and behave toward others, society and ourselves (in other words: how we *act*). So while contextualism can and does cover a wide range of topics, including epistemological and linguistic/semantic contextualism, we are here concerned with contextualism as it affects actions and first order moral questions and normative ethics, which is where most beginning philosophers or non-philosophers start their consideration of morality.

But just what are first order moral questions and normative ethics? In his book, *Morality Without Foundations*, Mark Timmons describes the former as being "questions about the moral status of persons, actions, institutions, practices, and the like."[2] For instance, we may ask, "What qualities about an individual or the actions of an individual make that person or those actions morally good or morally bad?" or "What features of a societal institution or social practice makes it morally good or morally bad?" When we attempt to answer these questions, we employ normative ethics, which is the formulation of "principles of right conduct and principles of value or goodness that state the most general conditions" of the rightness or wrongness of actions, the intrinsic goodness or badness of states of affairs, and the moral worth of persons.[3] Kant's categorical imperative and Mill's principle of utility are two of the most well-known normative ethical

theories; others include Aristotelian virtue ethics, the principles of hedonism, and the Judeo/Christian ethic.

These ethical constructs are designed to help us make correct decisions when confronted with difficult and confusing moral dilemmas, whether in broad social concerns such as abortion, euthanasia and capital punishment; in business and academic concerns such as preferential hiring practices and affirmative action; and in our inter and intrapersonal conflicts. But when embarking on a course of ethical studies many people soon find that while these normative ethical theories are indeed powerful decision-making tools, many of them conflict with each other, several of them set up restrictions that seem to be impossible to adhere to, and a few of them outright contradict each other. Additionally, the arguments for or against each construct may seem incredibly strong, leaving us to wonder which, if any, theory we should employ.

The classic examples of this involve Kantian duty ethics and act utilitarianism. One of the first things one learns about Kant, for instance, is the importance he places on duty from a good will, as he makes clear in the opening paragraphs of his *Groundwork of the Metaphysics of Morals*. Simply put, this devotion to duty leads Kant to form the categorical imperative, the first formulation of which is "I ought never to act except in such a way that I could also will my maxim should become a universal law" [Kant 4:402][4], the first conclusion of which is that lying under any circumstance is always wrong. Were lying to be universalized—more specifically, were all people to enter into contracts or agreements they have no intention of keeping—society would collapse. Banks would no longer lend money, nor could they be counted on to house and protect assets.[5] Credit would disappear. The very concept of money would be meaningless, since money is a symbol of wealth and not wealth itself. People could no longer trust each other to keep even the simplest and mildest of promises. We would by necessity return to the basest state of nature.

Contrast this with a basic principle of act utilitarianism, which may be stated as always acting in a way that maximizes

pleasure over pain for the greatest number of people involved in the action. We see examples of this quite frequently, particularly whenever a vote is taken to decide a specific course of action for a group of people. Indeed, our whole electoral procedure is almost purely utilitarian in nature: the most popular candidate (that is, the person who brings the greatest amount of pleasure over pain for the greatest number of people involved) is the one elected to office.

Even a cursory examination of these two theories quickly reveals the contradiction: Kant's prohibition of lying certainly seems to fly in the face of the utilitarian for whom lying may be permissible so long as it maximizes pleasure over pain for the majority. Taking it a step further, Kantian duty does not seem to be concerned with consequences; while consequences for the utilitarian seem to be the *only* thing that matters.

The problem grows more complex when we go out of the realm of theory and into the realm of the practical. For Kant, the classic example is the Nazi at the door. Suppose you lived in Nazi-occupied Amsterdam and were hiding ten Jews in your basement. One day, a Gestapo agent knocks at your door and asks you directly, "Are you hiding Jews in your basement?" As a Kantian, you are prohibited from lying no matter what the circumstances may be. But is adhering to duty and allowing the Nazis to haul innocent people off to almost certain death in a concentration camp really the correct moral choice? Surely the utility principle of maximizing the pleasure (not to mention the very lives) of the ten Jews over the pain of the one Gestapo agent is the preferable choice.

There is a similar classic example of a practical dilemma for the act utilitarian, which we will call "the fat person in the cave". Suppose you and ten of your friends go hiking along the coast one afternoon, and discover a cave. You all decide to explore it and spend several hours doing so, not realizing that the tide has blocked the coastal exit, and the water is rising rapidly. You all go further up and in the cave to escape the water and find a small opening to the outside. One of your friends, a very fat man, panics

and shoves his way ahead of everyone else to get out. But his girth is too great, and he gets stuck in the opening, unable to go forward or back. The water is still rising, and you and you friends are trapped thanks to your portly companion. But just when you resign yourself to the fact that you and eight of your friends will soon drown, let us further suppose that someone in the cave with you has in her possession a small stick of dynamite, which you can use to dislodge the fat man. Unfortunately, it will also most likely kill him. As a utilitarian, your goal is to maximize pleasure over pain for the greatest number of people involved, and not drowning is certainly more pleasurable than the alternative. But does that really include obliterating a person whose only offenses are an excess of body weight and a tendency to panic? Surely a sense of Kantian duty that precludes the murder of one person to save nine people should prevail.

As we can see, both Kantian deontology and act utilitarianism have very attractive and preferable goals. And yet, strict adherence to either theory can also lead to undesirable moral outcomes. So when faced with ethical dilemmas, which should we choose?[6] Ethical contextualism obliges us to choose the course based on the context of each situation. In fact, that is exactly what we have done by asking questions at the end of each of the above examples. The questions are designed to alter the actions normally compelled by both Kantian duty ethics and utilitarianism to fit the contexts of each individual situation. The people in the examples start out by understanding and adhering to strong ethical constructs—Kantian duty ethics in the first case and act utilitarianism in the second case. When the dilemmas come, each try to apply their ethical construct to that dilemma in order to arrive at a right course of action, and each find their desired ethical construct wanting *because of the context of the dilemma*. Each then modifies their construct, or even turns to a different construct, in order to ascertain the best course of action. It is this analytical approach that keeps ethical contextualism from becoming a mere reliance on intuition or feelings when making moral choices.

Many may see similarities between ethical contextualism

and Aristotle's doctrine of "the mean" or W. D. Ross' *prima facie* duties, and this is certainly understandable. J. O. Urmson defines the doctrine of the mean as feeling and manifesting "each emotion at such times, on such matters, toward such people, for such reasons, and in such ways, as are proper."[7] Or as Aristotle puts it, it is to feel and manifest each emotion "to the right person, to the right extent, at the right time, with the right motive, and in the right way" (NEII.9 45).[8] In other words, we need to act based on the context of the situation.

So, too, with Ross' concept of the *prima facie* duties, which was developed in part because Kant's ethical absolutism seems unsustainable since the moral situations we may encounter on a daily basis are not static but fluid; that is, those situations are too varied and diverse for a one-size-fits-all moral approach to be effective. Instead, Ross came up with six categories of *prima facie* duties, and counseled that when we are faced with moral decisions, we should seriously consider all of these duties and decide which is the most important to follow for that particular circumstance. As Ross puts it:

> When I am in a situation, as perhaps I always am, in which more than one of these prima facie duties is incumbent upon me, what I have to do is study the situation as fully as I can until I form the considered opinion … that in the circumstances one of them is more incumbent than any other.[9]

While Ross' theory and Aristotle's doctrine are certainly in line with ethical contextualism, the latter is much broader in scope than the two former, in that ethical contextualism asks us to consider not just a certain doctrine or six categories of duties when making moral decisions, but to also consider the broad spectrum of ethical thought—from strict Kantian absolutist duty on the one end, to ethical egoism and even hedonism on the other— and from that reflection to determine which is the best course for the situation at hand.[10]

If ethical contextualism compels us to make moral choices based on the situation, then how does it differ from moral rela-

tivism? Moral relativism claims that morality only exists in relation to culture or society and is neither absolute nor foundational, nor can any moral decision be considered right or wrong, or good or bad. On the surface, this seems to agree with some of the assertions about ethical contextualism made in this paper; for instance, ethical contextualism as defined here does not oppose the relativist stance that all moral theories are equally valid. However, there is an important difference between the two. While ethical contextualism holds that moral decisions are (and should be) contextual in nature, I also argue that ethical contextualism claims that moral choice *itself* is foundational in that, since the choice of which ethical theory to employ depends on the context of the moral dilemma, there is, therefore, a right choice for each moral dilemma.

There is another foundational consideration in ethics that should be addressed. In *Morality Without Foundations*, Mark Timmons uses ethical contextualism to basically affirm the relativist view that morality is cultural and not foundational. In the book's introduction, Timmons writes:

> My version of ethical contextualism represents a plausible metaethical view without needing to look for a metaphysical foundation in some realm of moral facts and without needing to find some unshakable epistemological foundation upon which to rest our justified moral beliefs. In short, we get a philosophical account of morality without foundations (original emphasis).[11]

However, Timmons has placed a foundational element within his denial that we need foundational elements. That element is *belief*. Belief is necessary to our very existence. Every scientific discovery, every medical advancement, every engineering feat, every proven mathematical or logical theorem started out as, and still remains, a belief. Belief is essential not only to our concepts of ethics, but also to our humanity; that is, we cannot be rational, sentient beings without belief.

In *The Nicomachean Ethics*, Aristotle tells us, "It is no easy

task to be good." (NEII.9 45)[12] This is especially true when, as we have seen, ethical theories conflict with or even contradict one another. Ethical contextualism is a powerful analytical tool, grounded in the metaphysically-necessary concept of belief, that helps us surmount the contradictions and conflicts that arise between competing normative theories about good/bad, and right/ wrong actions, and provides us a way through the conflicts to help us make the best moral choices we can.

Notes

1. Price, A. W. *Contextuality in Practical Reason*. Oxford: University, 2008.

2. Timmons, Mark. *Morality Without Foundations: A Defense of Ethical Contextualism*. Oxford: University, 1999. 9.

3. Ibid. 10

4. As reprinted in *Practical Philosophy: The Cambridge Edition of the Works of Immanuel Kant*. Ed. Mary. J. Gregor. Cambridge: University, 1996.

5. We are getting a first hand example of this in our society right now with the sub-prime loan debacle.

6. Some have pointed out that the conflicts between act utilitarianism and Kantian duty ethics are solved by adjustments within the theories themselves. For instance, proponents of utilitarianism would cite rule utilitarianism as an answer for the problems found in act utilitarianism. In rule utilitarianism, rather than considering the consequences of a particular act, we instead deliberate on rules of conduct; whichever rule has the best overall consequences is the best rule, and, therefore, the one we should follow. So in the case of "the fat man in the cave" example we might conclude that the rule "Murder is always forbidden" will bring about the best consequences overall, thus bringing our actions in line with Kantian duty ethics. Similarly, many Kantian scholars will point out that the apparent rigidity of the Categorical Imperative is greatly mitigated by how we form the maxim, or rule, to which the CI is to be applied. In the "Nazi at the door" example, if we make it our maxim to never lie unless by not doing so we hurt innocents or cause them to be hurt by others, this would bring our actions more in line with the consequentialists. But both of these examples only serve to strengthen the contextualist claim; in fact, they are both the very essence of how contextualism actually works, for in each example, actions depend entirely on context.

7. Urmson, J.O. "Aristotle's Doctrine of the Mean." Rpt in *Essays on Aristotle's Ethics*. Eds. Amelie, Oksenberg, Rorty. U.S.: University of California Press, 1980. 161.

8. Aristotle. *The Nicomachean Ethics*. Translated by David Ross. New York: Oxford, 1998. 45.

9. Ross, W.D. *The Right and the Good*. Ed. Philip Stratton Lake. Oxford: University. 2002. 19.

10. Aristotle believed that we are not virtuous by nature and so this "determining of the best course" can only come about by constant practice. I do not disagree with this assertion; however, I would modify it to say that though we may not be virtuous by nature, we still do, nonetheless, carry out ethical contextualism by nature. For instance, we all seem to know, or at least have some sense of, when an ethical decision, say by a judge in a court of law, seems unjust. We may not know why, we just know that it goes against our core sensibilities of justice. Knowing the why is where Aristotle's practice comes in; we may never fully know why we feel the way we do about an ethical decision (even ones we make), but learning all we can about the broad spectrum of ethical thought may bring us a step closer to figuring it out.

11. Timmons, Mark. *Morality Without Foundations: A Defense of Ethical Contextualism*. Oxford: University, 1999. 7.

12. Aristotle. *The Nicomachean Ethics*. Ed. David Ross. NewYork: Oxford, 1998. 45.

Bibliography

Aristotle (1998) *The Nicomachean Ethics* Ed. David Ross. (New York: Oxford University Press)

Kant, Immanuel. (1996) *Practical Philosophy: The Cambridge Edition of the Works of Immanuel Kant*. (Cambridge University Press)

Price, A. W. (2008) *Contextuality in Practical Reason*. (Oxford University Press)

Ross, W.D. (2002) *The Right and the Good*. Ed. Philip Stratton Lake. (Oxford University Press)

Timmons, Mark (1999) *Morality Without Foundations: A Defense of Ethical Contextualism* (Oxford University Press)

Urmson, J.O. (1980) "Aristotle's Doctrine of the Mean." Rpt. in *Essays on Aristotle's Ethics* Eds. Amelie, Oksenberg, Rorty (University of California Press

DECONSTRUCTING UTILITARIANISM: WHY HAPPINESS IS NOT THE ONLY BASIS FOR MORALITY

David Nagy

I. INTRODUCTION

Ethics tells us what we should do and how we should live our lives. Though ethics is important, no one knows exactly what the correct actions are. When we say an action is wrong, what do we mean? Apparently that the action should not be done—but what does that mean? *Why* should that action not be done?

These questions seek a basis for morality, a reason for us to judge some actions moral and others immoral. The debate over the correct basis for morality has subsumed the majority of philosophical thought in ethics for thousands of years. Philosophers, being the type of people they are, have thought of a wide variety of bases. A theory that states a basis of morality, some sort of set of rules that we can judge actions by to determine if they are right or wrong, can be called an ethical theory.

Obviously, there is not enough space in this article to analyze every ethical theory that has ever been invented—there is probably not enough space in all the books of the world. Therefore, this article will focus on one popular ethical theory; as you might guess from the title, this would be Utilitarianism. While there are many forms of Utilitarianism, I will be presenting a critique that can, I hope, be applied to all (or almost all) varieties of it.

In particular, I will consider two very different (yet quite related) types of Utilitarianism, which can be called Psychological Utilitarianism and Ethical Utilitarianism. Psychological Utilitarianism states that the only goal of human action valued *in itself*— that is, the only good that humans value as good in itself, and not

because it achieves other goods—is happiness. As a result of this, Psychological Utilitarians argue that the only end goal of human action is happiness. Ethical Utilitarianism states that the moral thing to do is, in some fashion, to maximize everyone's happiness and to minimize everyone's suffering. While the way to do this varies from Ethical Utilitarian to Ethical Utilitarian, all of them agree on this basic point.

II. Summary

The purpose of this article is to argue that both forms of Utilitarianism, Psychological and Ethical, are false. First, in section III, I will critique Psychological Utilitarianism, by providing four examples of things that *some humans* value in themselves apart from happiness; i.e., I will argue that sometimes people *do not act* for the sake of happiness. Then, in section IV, I will critique Ethical Utilitarianism, by providing four arguments as to why, sometimes, it is *not* moral to simply maximize happiness and minimize suffering.

In section III, I will start out by examining the link between Psychological and Ethical Utilitarianism. Specifically, many people use Psychological Utilitarianism to try to prove Ethical Utilitarianism. The argument boils down to this: Happiness *is* the goal of all human action, so it *should be* the goal for all human action. However, I will demonstrate that there are at least four goals of human action that do not include happiness: virtue, freedom, power, and remembrance. This argument fails because, sometimes, we desire these things even if we have to suffer to achieve them. I will also examine the Psychological Utilitarian's response, which is to define happiness such that it includes the desires I mention. The problem is, the concept "happiness," however you define it, must include *pleasure* somehow—if the Psychological Utilitarian defines the term too broadly, it ceases to be happiness. When that happens, the Psychological Utilitarian has undermined her own position.

Next, in section IV, I will critique Ethical Utilitarianism

by providing four arguments against the idea of happiness as the basis for morality. First, I will argue that it is impossible to measure happiness. Second, I will try to show that it is impossible to predict happiness. Third, I will argue that Utilitarianism implies that we should seek only bodily pleasures, which is an undesirable outcome. Fourth, I will argue that Utilitarianism implies that we should kill people who will have more suffering than happiness in their lives, which is also an undesirable outcome.

Finally, in section V, I will give a brief synopsis of my argument, and deliver some closing observations.

A final note before we begin: this article is entirely concerned with Psychological Utilitarianism and Ethical Utilitarianism. Other controversies in ethics (e.g., Deontology vs. Consequentialism vs. Virtue Theory, the right vs. the good), while interesting and valuable, are beyond its scope. I am only trying to disprove Psychological and Ethical Utilitarianism.

III. PART ONE: HAPPINESS IS NOT THE ONLY GOAL OF HUMAN ACTION

According to Psychological Utilitarianism, the only thing humans value as a good in itself is happiness; in other words, the end goal of all human action is happiness. At this point, you might wonder why this theory is relevant to the overall aim of my article. After all, the main idea of this article revolves around ethics—what the correct basis for morality is. What role does a seemingly psychological theory play in morality?

The answer is that Psychological Utilitarianism has often been used to argue in favor of Ethical Utilitarianism. The thrust of that argument goes as follows: Happiness *is* the only goal of human behavior, so it *should be* the only goal of human behavior. John Stuart Mill, perhaps the most famous Utilitarian, gave this defense of his theory; I shall be using Mill's arguments as a foil in this essay.

Before moving forward, however, it will be necessary to actually define "happiness." In *An Introduction to the Principles*

of Morals and Legislation, Bentham ("the father of Utilitarianism") seems to define happiness as "pleasure," and the opposite of happiness as "pain" (Bentham 1996, pp. 585-586). In other words, to Bentham, happiness is a kind of experience—this interpretation is supported by the fact that he thinks we can measure happiness through things like its intensity, duration, and certainty (Bentham 1996, p. 589). Mill is more vague on this point, but in *Utilitarianism*, at least, he seems to accept this definition. For the purpose of this essay, the following definition will suffice:

X is happiness if and only if:

(i) X is an experience; i.e., X is some sort of experiential state; and

(ii) X is pleasurable.

Obviously, what it means to be "pleasurable" is a key part of this definition. For now, the commonsense view of pleasure should suffice. As we will see later, Utilitarians can redefine pleasure to rebut some of my arguments—but we will get to that when the time comes.

In *Utilitarianism*, Mill argues that, to prove that something is a desired end of human action, one must merely prove that it is desired. Since happiness is a desired goal for human action, it is an end. And, since happiness is the *only* end of human action—since happiness is the only good valued as an end in itself, and not merely as a means to an end—it should be the *only* goal of morality. In other words, humans *do* desire only happiness, so humans *should* desire only happiness (Mill 2001, pp. 35-39). In the terms I've invented, Mill's argument amounts to the following: *Psychological* Utilitarianism is true, therefore *Ethical* Utilitarianism is true.

At this point, one objection becomes clear. Just because humans desire happiness does not mean happiness is moral. For example, some humans desire inflicting pain, but it does not follow from that that those humans *should* inflict pain. While I think this counter-argument is powerful in a certain way, in another way it is lacking. As Mill himself argues, it is essentially impossible

to argue in favor of first principles in any discipline, not just morality.[1] As happiness (or, as Mill puts it, utility) is the first principle of the ethical theory of (Ethical) Utilitarianism, it cannot be justified in ethical terms—there is no argument *in principle* that could prove why humans *should* desire happiness (Mill 2001, p. 35). This is true not just of Ethical Utilitarianism, but also of all other moral theories that have basic principles. But that does not mean we should abandon basic moral principles; as stated above, it is extremely difficult to see how one could achieve a coherent moral framework without them.[2]

With this in mind, Mill's argument becomes significantly more attractive—not on a logical or analytical level, perhaps, but on an intuitive level. Or, to put it another way, an Ethical Utilitarian could argue that she is not providing a set of ethical rules *as such*. Instead, she could merely be offering a framework, a way of better achieving what we already want to achieve, which, per Psychological Utilitarianism, is happiness. We already want to maximize happiness and minimize suffering; all Ethical Utilitarianism does is bring this out into the open, and give us a method to achieving our goal more effectively. While this might not be an ethical theory *per se*, the prescriptions it gives us *are* ethical.

For these reasons, I do not believe it enough to attack the logical structure of Mill's argument. If I want to show why his argument is wrong, I must show why his premise is wrong; I must show that humans desire things other than happiness, on a fundamental level. And I believe that humans desire at least four things other than happiness: virtue, freedom, power, and remembrance.

By *virtue*, I essentially mean morality itself; sometimes, people want to do the right thing even if it results in less happiness. This is not a circular argument, in that I am not saying that these people are necessarily correct in what they see as virtue. It may be possible that they are in fact incorrect, and the morally good action is always the one that increases happiness and decreases suffering, as some Utilitarians argue. Nevertheless, some people believe that a morally good action sometimes results in greater suffering than a morally bad action might; and sometimes, those

people wish to do the good action *merely because* it is good.

Mill himself admits that this is the case. In *Utilitarianism*, he states that "[people] desire, for example, virtue, and the absence of vice, no less really than pleasure and the absence of pain" (Mill 2001, p. 36). Mill's response to this point is to say that, while it is true that virtue is sometimes the end of human behavior (and not happiness), this is only because "in those who love [virtue] disinterestedly it has become so, and is desired and cherished, not as a means to happiness, but as a part of their happiness" (Mill 2001, p. 37). Mill is arguing that, yes, some people do value virtue as an end in itself, but only because they have made virtuousness *a part of happiness for them*. Mill is perfectly fine with people finding their own path to happiness (a point that will come up again when I talk about freedom), and if that means being virtuous, all to the good. Utilitarianism still stands.

At this point, an observation must be made: Mill could respond this way to *any* example demonstrated to him. Mill could say that a monk, whose life revolves around denying himself pleasure, is really just doing what makes him the happiest. You could drag a miserable drug addict in front of Mill, sacrificing so much happiness for the briefest flitter of intoxication, and Mill could say that he is pursuing what he *thinks* will make him happy, even if the addict knows his addiction makes him miserable. Indeed, Mill could posit happiness as the ultimate, true goal of all human behavior without truly giving a coherent argument in favor of it.[3]

However, just because Mill can posit happiness as the end does not mean it actually *is* the end. I could posit suffering as the goal of all human action, and come up with some explanation for every counter-example presented to me. Indeed, one can postulate practically anything in this manner. For Mill to be right, he needs to do more than show that he *could* be right, or that it is *possible* that happiness is the only thing valued as an end in itself. He needs to show that he *is* right, or at least that one would be justified in believing that he is right.[4]

The inherent problem in Mill's theory—which will come up again in the other three ends of human desire I will talk about—is

162

that his stance is intuitively wrong. Thinkers as old as Confucius have believed that the most moral person does good *because it is good*, and for no other reason. In fact, they would have argued that the truly good person always does good, *even if it brings them suffering*, and that idea has lots of intuitive appeal (consider how much we value martyrs). Certainly, one could argue that such a person takes "doing good" as part of her identity; but that the person takes virtue as a *part of her happiness* is much less obvious. At the very least, the process through which this integration of virtue and happiness is supposed to take place is very opaque, and Mill does not go into detail about what it actually is.

To give a concrete example to this admittedly vague point, consider a jury deciding a case about a vigilante. The vigilante kills a murderer, who lacked any family, friends, or loved ones. One can easily imagine a jury convicting the vigilante for murder solely on the basis of justice, which is part of virtue.[5] However, such a ruling clearly *contradicts* happiness: the murderer had no defenders, while the vigilante had loved ones who will be put into anguish by the ruling (not even counting the anguish of the vigilante). We could even imagine the jury members themselves being put into despair by sentencing an otherwise exemplary person to prison, whereas even they realize that they would be *happier* by letting the vigilante go. Certainly, one could argue that letting vigilantism run rampant would cause great suffering, but it is not rational to expect this will be a natural consequence of a not-guilty ruling. This action, made on the basis of virtue, seems quite clearly to contradict happiness, so it is hard to see how the jury has made virtue a part of a desire for happiness in this case.

I use words like "seems" and "hard to see" because, again, it is certainly *possible* that happiness lies at the foundation of their desires. Maybe the jurors take such pleasure in justice that it outweighs the suffering they feel when they see the vigilante's family (though this seems like, at least sometimes, it would be false). But it is possible that *anything* lies at the foundation of our desires. I am not arguing that Mill cannot be right; I am arguing that he is probably wrong. Examples like this show that Mill needs

to provide a much stronger argument than the one he offers; he needs to provide more evidence. Until he does, we are only justified in believing that *sometimes* happiness is the ultimate goal of human desires—and sometimes, virtue is.

At this point, I have accomplished the goal I have set out to do in this section—disprove Psychological Utilitarianism. However, all I have done for this purpose is cite one counter-example. My argument would be much stronger if I could provide additional ones, for the simple reason that it would make it *more likely* that Mill is wrong. In other words, maybe Mill is right that a desire to be virtuous is ultimately a desire to be happy; therefore, an argument that includes other things desired as ends in themselves is much stronger than an argument that includes only virtue.

In addition to happiness and virtue, we desire *freedom* as an end in itself. By freedom, I mean something quite specific: the desire to make one's own decisions, as opposed to following the decisions of others. In other words, some people are not satisfied with merely doing what others say—they want to make their own decisions and follow their own beliefs, for better or for worse. And that "or worse" part is very damaging to Psychological Utilitarianism. It means that some people want to make their own decisions, *even if that results in more unhappiness for everyone*.

Of course, an easy rejoinder that Mill could make is that this is yet another example of people making a certain ideal a part of their own happiness. Certainly, following your own beliefs results in greater suffering sometimes—but maybe those people who desire freedom as an end in itself do so because that is the only way they can be truly happy. Can mindlessly doing what others tell you to do really be called happiness?

This counterargument runs into the same problems the one against the virtue point. Sometimes, following freedom clearly and explicitly *contradicts* following happiness. This is especially obvious when we consider what happens when we reflect on our previous actions. I can think of several times when, because I made my own decision, suffering resulted for everyone—and yet, I still prefer making that decision to simply following the opinions

of others. *Even if* I would be perfectly happy simply doing what somebody else said (and at the very least, this would absolve me of the suffering of guilt), I would *not* want to. The Psychological Utilitarian has an uphill road to climb if they want to claim that this is really a desire for happiness.[6]

The third thing that humans desire as an end in itself, apart from happiness, virtue, and freedom, is *power*. By power, I am referring to something akin to Friedrich Nietzsche's idea of the Will to Power. Briefly, Nietzsche had the idea that the only thing that humans desire is power (Nietzsche 1968, p.366). On my interpretation, Nietzsche meant power not necessarily over fellow humans, but more generally over ourselves and our environment. We have power over ourselves by being able to control our actions, and we have power over our environment by being able to affect changes in it.[7]

I think Nietzsche is right in that we have this desire—at the very least, *I* have this desire. But again, humans sometimes have this desire even when it contradicts happiness. Perhaps the clearest example of this is the case of insanity. Some insane individuals are quite happy people, and yet, even if being insane meant that I would be happier than I am now, I would rather be in full control of my mental faculties.

This point is actually quite similar to the freedom argument I made above. The major difference, to me, appears in the focus. My freedom argument dealt with the desire to follow one's own beliefs, as opposed to others'. The point about power involves the desire to avoid having your own beliefs directly affected by natural forces. In any case, both of them include examples that seem to directly contradict the theory that all desires can ultimately be reduced to a want for happiness.

At this point, I have provided three non-happiness goals of human action that are desired as ends in themselves. While I do believe all of them can be separated from happiness, there are also rather obvious ways for a Psychological Utilitarian to respond to them. While I believe those ways are *ad hoc*, I also recognize that none of my arguments would likely influence someone who

already believes in Psychological Utilitarianism, or at least in the premise of Mill's argument. If I want to persuade them, I need something more visceral, something with a larger emotional impact. I will begin with the following observation: all humans will die.

Fear is a very strong emotion, and fear of death is one of the most intense types of fear (perhaps only trumped by the fear of one's loved ones dying). All rational beings, on one level or another, are aware of their eventual demise. It is the fear of this demise that gives rise to, in my view, one of the strongest human desires there is: the desire for remembrance, i.e., the desire humans have to have an *impact* on the world, for there to be some lasting *proof* of one's life after one's death.

This proof can take many forms. At its most universal, it can be memory in one's loved ones. At its most specific, it can be the founding of a nation, or the overhaul of the world's thought. It ranges from a name carved into a monument to the continuation of one's genes in one's children. In all its forms, it shares the same characteristic: even though one is dead, one is still *remembered*, and so in some sense, one is still alive.

This desire, a desire that I suspect everyone has but I *know* that I have, eviscerates Psychological Utilitarianism. This is because, by its very nature, the desire for remembrance *cannot* be a desire for happiness. It is a desire for a state of affairs that occurs after death, after the ability to feel happiness and suffering vanishes. It is a desire for something to happen when one is physically incapable of feeling happiness.

One may argue that, while the object of desire may take place after death, the action itself takes place while one is still alive. Then, one could move on to argue that the actions people take when they desire happiness they *only* take because it makes them happy. In other words, it gives people pleasure to *think* that they will be remembered. The problem with this argument is that the desire for remembrance is not for something so superficial as to *think* that one will be remembered, or to *believe that* one will be remembered. The desire is to *actually* be remembered. Admit-

tedly, this is a fine distinction—but moving from a desire to be remembered to a desire to be happy is even more of a jump than moving from virtue, freedom, or power to happiness. There is just no reason to believe that the counterargument is accurate.

In the end, I think the major error Psychological Utilitarians make when they respond to these types of arguments is that they define happiness too broadly. Considering happiness to include such a wide variety of things as a desire to be virtuous, a desire to be free, a desire to have power, and a desire to be remembered, may very well remove all explanatory power from the word "happiness". In other words, the major observation Psychological Utilitarianism is supposed to make is that *happiness* (defined as having something to do with pleasure) is the only goal of human action. Redefining happiness to encompass all that humans really *do* want (e.g. literally *defining happiness* as "everything humans desire") would make Psychological Utilitarianism *trivially* true—it would be reduced to saying "The only goal of human action is whatever humans desire." However you define "happiness," it must include pleasure somehow, because that is the sense that the word actually has in English. Once we move outside the realm of pleasure—once we start replacing the word "happiness" with the word "desire"—we admit the falsity of Psychological Utilitarianism.

IV. PART TWO: WHY HAPPINESS IS NOT A PROPER GOAL OF HUMAN BEHAVIOR

In Part One, I argued that Psychological Utilitarianism is incorrect. In Part Two, I will attempt to show that Ethical Utilitarianism is also incorrect. Ethical Utilitarianism, recall, is the theory that the basis of morality is happiness. In a systematic form, it is the theory that the moral thing to do is to maximize pleasure and minimize pain.

So, why should you believe that happiness is not the basis of morality? First, nobody can actually measure happiness. Utilitarianism requires that we act based on some notion of increasing happiness and decreasing suffering. But what, exactly, does it

mean to increase happiness? Would I be happier eating a steak, or eating chicken? Or maybe I should eat a salad—while that may give me less pleasure than the steak, it might also be healthier, meaning it would give me greater opportunity for pleasure in the future.

Is there a time value for pleasure? It seems that certain pleasure now is more valuable than uncertain pleasure in the future—but how much more? Is there a formula for how to find the current value of future pleasure? And if there is, how can we determine the numerical value of each pleasure?

A huge part of Utilitarianism is increasing the happiness of others. How can we know if someone else is happy? Certainly, they can tell us; however, they might be lying. It seems egoistic to project our own pleasures and suffering onto others, and assume that others find the same things pleasurable that we do.

Is pleasure worth more when spread out, or concentrated? Certainly, someone who tortures people for pleasure is creating more suffering than pleasure if he tortures, say, ten people. But what if he tortures one person? Would our judgment of the morality of that action then hinge on whether or not the pleasure he gains from torturing is greater than the suffering caused in the tortured? But we already had trouble measuring two types of pleasure against each other—how can we possibly measure pleasure against suffering?

Mill's attempts to defend his arguments in Part One only make the situation worse. If happiness covers pleasure, virtue, power, and remembrance, then happiness becomes *even more* impossible to measure. Does virtue or power give more happiness? What about remembrance—is it worthwhile to cause suffering if it would let me be remembered for 1,000 years? How about 100,000? There are so many different types of pleasures, and so many different experiences that can be *called* pleasurable, that it is impossible to compare them to see which is greater.

Note that I am not just saying that happiness is *difficult to measure*, or that we have to spend a lot of time doing it. I am saying that, *in principle, happiness cannot be measured*, because

happiness is not the type of thing that can be measured—you can't assign a number to happiness. At least, *humans* cannot do such a thing. Certainly, we can sometimes measure one type of pleasure against another, e.g., I get greater pleasure from eating steak than from eating lettuce. But once the examples start becoming at all complex, the concept of "levels" of happiness spirals off into a vague realm of confusion. Since happiness cannot be measured, it cannot be maximized, which means Ethical Utilitarianism cannot actually be practiced.

This vagueness of happiness creates another problem for Ethical Utilitarianism. Not only is happiness impossible to measure, but it is also impossible to predict.[8] If happiness really is the basis of morality, then we need to act with increasing happiness (and decreasing suffering) in mind. This necessarily means that we must have some idea when an action will increase happiness, and when it will increase suffering. However, it is impossible to do this—we can never even have a justified belief about which action will increase happiness and decrease suffering the most.

There are three main reasons for this. First is the problem that all such Consequentialist theories have: A Consequentialist theory is one where morality is judged on the basis of consequences, but when you project the effects of actions out far enough into the future, the actual consequences become unknowable. However, the Utilitarian can easily respond that we can still know the consequences to some extent, and we should act on those consequences that we do know, or could reasonably predict.

The second reason is that happiness cannot be measured, so it cannot be predicted. But this makes my second argument contingent on the success of my first one, meaning that I have not truly made my overall stance any stronger.

There is a third reason happiness is impossible to predict, however. It is a reason that is fairly unique to happiness, and that is because of the *nature* of happiness. Everyone is made happy by different things. I may like (and derive pleasure from) vanilla ice cream, while you may like chocolate ice cream. I may like philos-

ophy, while you may like physics. In other words, most things do not make *everyone* happy.[9] So, if we want to increase happiness for everyone, we need to know which *specific* things make various people happy.

But we cannot know this. Have you ever thought doing something would make you happy, only afterwards to learn that it did not? It could be something as simple as seeing a bad movie you thought would be good, or as life-altering as taking a job that ends up being boring. But the fact that it happens means that, a lot of the time, we do not even know what makes *ourselves* happy. So we know even less about what makes *others* happy, which means that, at least some of the time, it is impossible for us to predict what makes *anyone* happy.

Too many times, our actions have an impact on people we do not even know. Too many times, we are forced to act when we do not know what will make others happy. To have a justified belief about something, we need to at least be fairly certain of that belief. If we are not—if we have no idea if a certain belief is true or not—it is not justified. But too often, we are forced to act when we have no idea, or only a very vague idea, of what will result in the most happiness. Ethical theories are supposed to guide our actions in our everyday lives. If Utilitarianism cannot do this, it is not a proper ethical theory.

So, at this point, I have shown why Ethical Utilitarianism cannot be used as a guide for behavior. While this shows why we *cannot* follow Ethical Utilitarianism, it does not necessarily follow that we *should not* follow it. Maybe Ethical Utilitarianism is an ideal that we cannot reach, but we should try our best anyway, even if we must necessarily be imperfect. I do not think a morality that cannot be successfully followed is a good one—but I have not yet critiqued the core of Ethical Utilitarianism, the *ideal* of happiness. That is the purpose of my last two arguments.

My third argument is one that John Stuart Mill himself considered, as it seems to have been a popular argument against Utilitarianism. According to Ethical Utilitarianism, happiness, or pleasure, is the highest good, the thing we ought to maximize.

However, this seems to imply that we should seek only the base, physical pleasures of life—indulging in food, drink, sex, and the like. Why be an intellectual when one could have so much more fun—and be so much more certain of pleasure—if one enjoys, to put it strongly, the pleasures of swine? The burden is on the Utilitarian to provide reasons why merely indulging oneself in the pleasures of swine is not a good way to live.

Mill's response to this argument is to say that there are classes of happiness. He says that some kinds of pleasure are qualitatively better than others, and so presumably, we should seek those better types of pleasure (Mill 2001, pp. 7-8). How do we find out which types of pleasures are better than others? By asking people who have experienced both which they would rather experience (Mill 2001, pp. 8-9).

This response has several major problems. The three obvious ones are: First, it seems to be completely *ad hoc*, with no real justification inside a Utilitarian framework to make this distinction other than to respond to the objection. Second, it makes measuring and predicting happiness even more difficult—does one intellectual experience more or less pleasure than two bodily-pleasure-seeking dolts? And third, why should we expect *every person* who experiences two types of pleasure to *always* prefer one type to the other?

There is another, even more central problem to Mill's response: it essentially admits the complete falsity and uselessness of Ethical Utilitarianism. By setting the quality distinction of happiness to be "what people would rather experience," Ethical Utilitarianism's basis of morality changes from "happiness" to "what people like to experience." In other words, Mill ends up arguing merely that, to be moral, we should have people experience what they want to experience (which, if I had added even more arguments in Part One, Mill might have ended up with anyway). This new ethical theory may be a good one; however, it is a *new* ethical theory. The supposed central observation of Ethical Utilitarianism was that morality was based on *happiness*, or *pleasure*. The theory Mill ends up with makes no mention of

happiness—it removes happiness as its basis for morality—and therefore, it is *not* Ethical Utilitarianism.

These three arguments are enough to show that, not only is Ethical Utilitarianism useless to morality, but that in the process of defending it, one must give up the supposed insight that the basis of morality is happiness. But there is one more argument which shows that *any* type of Ethical Utilitarianism is inherently incoherent. I have termed this argument the "problem of suffering."

Consider a person in a hospital, ravaged by a painful disease. There is no doubt that the rest of this person's life will be defined by pain. She will surely experience more suffering than happiness, *if she lives*. On the other hand, *if she dies*, she will, by definition, experience neither happiness nor suffering.[10] In other words, if she lives, she will have more suffering than happiness—but if she dies, she will not. If we value happiness as the basis for morality, this means that we have the *moral imperative* to kill this person. This is all well and good if the person wants to die. But what if she does not? In Utilitarianism, that does not matter—the desire for autonomy plays no role in the Utilitarian's moral calculus. If happiness is the basis of morality, then no matter what, we must kill any person we reasonably believe will experience more suffering than happiness.

This idea is, intuitively, morally repugnant. But earlier, I claimed that this argument shows why Utilitarianism is incoherent, not that it is contrary to our intuitions in one specialized case. It does not seem I have done what I promised. However, the case I have mentioned is not specialized—it applies to the majority of human beings.

Having lived a comfortably middle-class life myself, it is hard for me to imagine what the majority of this world's lower-class lives are like. But I have heard descriptions: starvation, disease, poor education, degrading jobs, and miserable living conditions. I could go on and on. Many people in the world really do live in a wretched state of existence; it would be extremely difficult to argue that they experience more happiness than suffering.

Ethical Utilitarianism is supposed to be a moral theory. Not

only that, it is supposed to be an *altruistic* theory. Its proponents were the major driving force behind many good movements in many countries, including prison reform. But, if we follow its logic strictly, it implies that we should kill a large portion of the people on the planet because there is more suffering than happiness in their lives.

There are very few ways an Ethical Utilitarian could respond to this. Certainly, there are at least *some* people who freely admit that they experience more suffering than pleasure, yet still want to live. Ethical Utilitarianism cannot just give them a pass because of their desire—otherwise, as I argued before, they would give up happiness as the basis of morality and replace it with human desires. Ethical Utilitarianism cannot inherently value pleasure over suffering, because then the sadist might be justified in torturing people for the pleasure it brings her. The only remaining route I can see is to argue that existence *in itself* is a pleasure; that the pleasure we get merely from existing so outweighs the lack of pleasure we get in death that the Ethical Utilitarian scales are heavily weighted on the "life" side.

In my opinion, this response makes the same mistake Mill did back in Part One. You can define happiness to include whatever you want, but eventually it stops being happiness (e.g., if you define happiness to include suffering, you are not truly talking about happiness anymore). Many people commit suicide every year because of the suffering in their lives—indeed, those people often suffer less than others who desire to continue living. It is clear from this that the desire to continue living has more to it than a simple pleasure over pain calculation; some people want to live even if their suffering far eclipses their pleasure. The desire to live is outside the pleasure-pain continuum, and thus it must be outside of happiness.

This means that the logical conclusion of Ethical Utilitarianism is genocide. Any ethical theory that leads to this is not an ethical theory—it is the opposite of an ethical theory. Ethical Utilitarianism's focus on a crass calculation of happiness over suffering as the supreme moral force renders it utterly incoherent.

V. CONCLUSION

Morality—what we should and should not do—has a huge impact on how we lead our lives. We need some sort of standard, or basis, to determine the moral thing to do. Throughout history, many such bases have been given, but one of the most popular has been Ethical Utilitarianism, which posits that *happiness* (or pleasure) is the basis of morality—we *should* (morally speaking) maximize pleasure and minimize pain (or suffering). Some people have used the related theory of *Psychological* Utilitarianism, that happiness is, in fact, the only goal of human action, to argue for Ethical Utilitarianism: Psychological Utilitarianism is true, therefore Ethical Utilitarianism is true.

However, Psychological Utilitarianism is wrong. Humans desire, as ends in themselves, things other than happiness. These include *virtue* (being a good person), *freedom* (acting on one's own beliefs; making one's own choices), *power* (having an effect on one's mental and physical world), and *remembrance* (being remembered after one's death). The only way a Psychological Utilitarian can respond is to define "happiness" broadly enough to include all of these, among others.

But, by doing this, the Psychological Utilitarian is effectively admitting the falsity of their own position. For Psychological Utilitarianism to be true, it is not enough to merely define "happiness" to mean "what humans desire"—that would make Psychological Utilitarianism *trivially* true. We need to gain some additional insight by using the word "happiness." The only insight we can gain is that all humans desire *only* some form of pleasure; if the Psychological Utilitarian admits that pleasure is not enough to explain the totality of human action, she has admitted that Psychological Utilitarianism is false.

Therefore, a major argument in favor of Ethical Utilitarianism (that is, the truth of Psychological Utilitarianism) has a faulty premise. But Ethical Utilitarianism, even considered on its own merits, has serious faults. Happiness is impossible to measure and impossible to predict. This means that Ethical Utilitarianism,

as a practical matter, is impossible to follow with anything approaching reliability.

There are also two problems with the very concept of happiness as the basis of morality. If we are merely seeking to maximize pleasures, we should just indulge ourselves in food, drink, and sex. The only response amounts, essentially, to redefining "happiness" as "what humans want to experience," which puts the Ethical Utilitarian in the same boat as the Psychological Utilitarian; defining happiness that broadly removes all explanatory power the word "happiness" was supposed to have.

Finally, there is the problem of suffering: most peoples' lives contain more suffering than pleasure. Under Ethical Utilitarianism, this means that we should kill the majority of the human race, *even though* most humans want to live. The possible counterargument that existence in itself is happiness is confusion— mere existence does not bring pleasure. If it is possible for an ethical theory to be subject to a *reductio ad absurdum*, logically leading to genocide would be it.

However, do not read more into this conclusion than you should. I am not saying that happiness is not *a* basis for morality. Quite the contrary—pleasure is an extremely important goal, and there are much worse things to do than maximizing the human race's happiness. While happiness is, in my opinion, *one* basis for morality, *one* standard, *one* rule, it is not the *only* one. That is all that I have tried to show.

Notes

1. By saying this, I think Mill may have in mind something like the Foundationalist theory of knowledge—some things we know without reference to other things. Just like those foundational pieces of knowledge cannot be justified by argumentation, neither can the first principle of morality. Foundationalism has been challenged as the correct theory of belief justification; nevertheless, a similar format does seem to apply to ethical theories like Utilitarianism.

2. This issue, which one might call an issue of moral nihilism, is much more complex than I imply here. This is not the article to address it in, however. For now, let us assume that it is possible to argue for a moral first principle in non-moral terms.

3. Note that this is much like Psychological Egoism, which claim that all human action is done to benefit the self. Even the most altruistic of people could be said to merely be trying to get a good reputation, or to avoid guilt, or acting for some other selfish motive. It is interesting that the altruistic Utilitarianism could have this similarity to egoism; however, this observation is not truly relevant to the discussion at hand. I mention it merely to elucidate my point.

4. I shall not talk about the issue of when one is justified in believing something in this article. For now, I think it is safe to assume that a standard of justification exists, and Mill's argument falls somewhere short of that standard. I doubt most would disagree with me on those points (though philosophy being what it is, I am sure a few would).

5. Some might argue that justice is synonymous with virtue. I do not wish to argue this here; I think most would agree that justice and virtue are inextricably linked, in any event.

6. This line of argumentation is very much related to one I will give in Part Two, when I talk about how following Utilitarianism would lead us to seek only the sensual pleasures of life. For now, it is enough to point out that "happiness" is an inherently vague term, and defining it to include all the cases I point out makes it even more vague, to the point of meaninglessness. Again, I will go into more detail on this point in Part Two.

7. Keep in mind that this is only *my interpretation* of Nietzsche. In any event, this is not a Nietzsche exegesis, so whether or not Nietzsche actually believed this is irrelevant. I only want to give credit to this idea where credit is due—I did not come up with this power concept myself.

8. These two points are, of course, fundamentally related, in that they both depend upon the fact that nobody really knows what "happiness" is. I make both points because they show different aspects of this observation.

9. It could be argued that there are a few things that *do* make everyone happy. I have doubts about that—but let us assume that it is true. However, it should be undeniable that the *vast majority* of things do not make everyone happy; the set of things that make all humans happy is much smaller than the set of things that do not. This means that any ethical theory that is constructed solely out of the things that make everyone happy would be woefully incomplete.

10. This is not the place to get into an argument regarding the afterlife. In this article, let us just assume that, if the afterlife exists, any experiences there will be so different from experiences in this life that they cannot be compared.

Bibliography

Bentham, Jeremy. (1996) "An Introduction to the Principles of Morals and Legislation," in: D. Wootton (Ed), *Modern Political Thought: Readings*

from Machiavelli to Nietzsche, pp. 585-589 (Indianapolis, IN: Hackett Publishing Company)

Mill, John Stuart. (2001) G. Sher (Ed), *Utilitarianism* (Indianapolis, IN: Hackett Publishing Company)

Nietzsche, Friedrich. (1968) W. Kaufmann (Ed), W. Kaufmann & R. Hollingdale (Trans), *The   Will to Power* (New York: Random House)

Kant, Korsgaard, and Civil Disobedience

Roberto Lewis

The Careful Moralist

The requirements of morality exert their influence on our lives constantly, even as we struggle to understand and define them. We feel compelled by some element in our nature to make decisions even when the relevant moral facts are uncertain. At times we are faced with moral choices that entail significant consequences. We struggle to satisfy the demands of morality, demands that are unclear in their implications for action and, as Christine Korsgaard puts it, "unconditional and overriding" (Korsgaard 2008, p. 233). The requirements of morality often lead us to take positive action in response to what we perceive as injustice. The moral actor will also be concerned about the justice or injustice of his own action, and the consequences that result from it. His decision will be easier to make if he has a moral system to evaluate his options. When attempting to negotiate moral dilemmas, the careful moralist ("careful" in the sense that he maintains, as best he can, an awareness of the full moral implications of his actions and beliefs) will have to balance his own needs and moral urges with those of his neighbors.

The Kantian system of morality provides useful standards for individual action, but problems arise when we attempt to apply the Kantian system to acts of injustice committed by the state. Kant's views on life lived under the restrictions of political society are strict, as are his views on the requirements of living a moral life. There are few openings where a citizen who takes it upon himself to promote his conceptions of justice and freedom for his fellow citizens could be said to be acting morally, or even in a fashion that Kant would not consider criminal. For Kant, the law is right, and he who breaks it is wrong, and the circumstances do

not seem to matter. In "On the Common Saying: That may be correct in theory, but it is of no use in practice," Kant writes:

> But if a public law is in conformity with this, and so beyond reproach, with regard to right, then there is also joined with it authorization to coerce and, on the other's part, a prohibition against actively resisting the will of the legislator; that is, the power within a state that gives effect to the law is also unopposable, and there exists no rightful commonwealth that can hold its own without a force of this kind that puts down all internal resistance, since each resistance would take place in conformity with a maxim that, made universal, would annihilate any civil constitution and eradicate the condition in which alone people can be in possession of rights generally.

> From this it follows that any resistance to the supreme legislative power, any incitement to have the subject's dissatisfaction become active, any insurrection that breaks out in rebellion, is the highest and most punishable crime in the commonwealth, because it destroys its foundation. And this prohibition is unconditional, so that even if that power or its agent, the head of state, has gone so far as to violate the original contract and has thereby, according to the subject's concept, forfeited the right to be legislator inasmuch as he has empowered the government to proceed quite violently (tyrannically), a subject is still not permitted any resistance by way of counteracting force (Kant 1999, "On the common saying…" 8:299-300).

Kant is saying that, if we are behaving correctly as citizens of a society, we are not only forbidden to question the rightness of the laws that govern our lives, we must not act against the wishes of those who create and administer those laws. Even if our rulers are tyrants, rebelling against them is absolutely prohibited. No state could operate successfully if this were not the case, because if it were morally permissible to defy our legislators, then, conceivably, every citizen could defy them, which would create a

condition of lawlessness in which no one's rights were respected. Rights, he argues, can only be claimed within the jurisdiction of government.

The view of government as the sole source of justice leaves little, if any, room for us to create positive change outside of the channels provided by our rulers, who are often concerned primarily with their own interests. Something within us resists this restriction. As we have only limited control over our emotional states, we similarly have limited control over our perceptions of and reactions to instances of justice or injustice. When a regime claims to be acting on our behalf and yet commits acts that are substantively unjust, and especially when they violate the standards of justice the regime itself claims to promote, we feel moved to act. We feel moved to protest such activity, even as we think of ourselves as law-abiding people. Even if we do not act, out of fear of our rulers, we feel as though we should. Our understanding of the rationale behind the rule of law fails to resolve this quandary.

Civil disobedience, defined here as the conscious non-violent breaking of a law, or the refusal to obey a command, in order to demonstrate opposition to the policies of a government or organization, is, I will argue, a technique of resistance to authority that satisfies the Kantian conditions for right moral action. Kant's views on the role of the individual in civil society seem, at first glance, to discourage the use of civil disobedience as a tactic to create social change, because it involves unlawful activity. I will argue in favor of Korsgaard's position, that civil disobedience occupies a middle ground created by the conflict between the moral citizen's duties of virtue and of justice, an area in which he is forced to act unjustly in the interest of justice (Korsgaard 2008, pp. 259-260). I will argue that the individual unjust actions involved in a program of civil disobedience, taken together, compose a just action.

KANT AND KORSGAARD

Korsgaard's subject is the "conscientious revolutionary," a "good person" who sometimes "finds she must rebel," even while

accepting that "morality is unconditional and overriding," and that "revolution is always wrong" (Korsgaard 2008, p. 234). She begins by analyzing Kant's division of the duties of the individual into duties of virtue (or ethical duties) and duties of right (or duties of justice). Duties of right are derived from Kant's Universal Principle of Right, which states that "any action is right if it can coexist with everyone's freedom in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone's freedom in accordance with a universal law" (Kant 1999, "The Metaphysics of Morals" 6:231). In other words, in order to act justly, individuals must act in a manner that does not restrict the freedom of other individuals. It follows from that, Kant says, that any action that is a "hindrance" to a just action (one that can coexist, if made into a universal law, with the freedom of everyone) is an unjust action (Kant 1999, "The Metaphysics of Morals" 6:231). If an action that hinders a just action is wrong, the argument goes, then an action that hinders an action *that hinders* a just action is *right*. In short, Kant has provided a rationale for coercion. In order to protect freedom *for* everyone, we are permitted to restrict the freedoms *of* everyone (Kant 1999, "The Metaphysics of Morals" 6:231-232).

In the prescription of an action as either a duty of virtue or a duty of right, there must be a *law* that makes the action into an objective duty, and there must be an *incentive*, or a subjective reason for a person to perform the action. A duty of virtue differs from a duty of right in that the incentive for a person to perform a duty of virtue is *the fact that it is their duty*. Kant calls this an *ethical* incentive. Duties of right, on the other hand, have *juridical* incentives. Juridical incentives are external and have to do with things other than duty. According to Kant, "the mere conformity or nonconformity of an action with law, irrespective of the incentive to it, is called its *legality* (lawfulness); but that conformity in which the idea of duty arising from the law is also the incentive to the action is called its *morality*" (Kant 1999, "The Metaphysics of Morals" 6:219). The relationship between law and incentive tells us how to think about actions, how to prioritize them, and to what

degree we can coerce others to perform them. I make sure to call my mother on Sundays because I know that I *ought* to, because I feel that it is the *right* thing for me to do, and not because I am worried about being formally punished if I *do not* call her. If I forget to put the correct number of coins into a parking meter when I park in front of it, the fine I will receive acts as an incentive for me to remember in the future (Kant 1999, "The Metaphysics of Morals" 6:218-219).

Korsgaard points out that under Kant's conception of political society, to have a "right" to something means to have the authority to force others to respect that right. If I assert a right, for example a right to property, I am in a sense making a law, because my right, and my authority to enforce it, extends to everyone (Korsgaard 2008, pp. 238, 239). If I buy a book, no one is permitted to steal it from me, even if they did not witness my buying it.

Kant argues that human beings have a "natural right" to freedom and to private property that can be derived a priori. People will those rights to be respected and enforced, which requires joining together in mutual coercion, so Kant argues that political society is the expression of the combined will of all of its members:

> Now, a unilateral will cannot serve as a coercive law for everyone with regard to possession that is external and therefore contingent, since that would infringe upon freedom in accordance with universal laws. So it is only a will putting everyone under obligation, hence only a collective general (common) and powerful will, that can provide everyone this assurance. - But the condition of being under a general external (i.e., public) lawgiving accompanied with power is the civil condition. So only in a civil condition can something external be mine or yours (Kant 1999, "The Metaphysics of Morals" 6:256).

When people establish procedures for enforcing individual rights, they create a *juridical condition*, a state of mutual and universal coercion under which human rights can be upheld and enforced.

This, according to Kant, is political society (Korsgaard 2008, p. 240). Without this system in place, rights cannot be enforced and there is no justice. Living in political society is the only way to ensure the rights and freedom of everyone, so living in political society is a duty of justice. An action that threatens political society or its institutions is in violation of that duty.

Kant calls the ruler who represents the will and coercive authority of the people the "sovereign." The sovereign authority may take a variety of forms, ranging from a single dictator to a network of governing institutions (Kant 1999, "The Metaphysics of Morals" 6:339). In challenging the legitimacy or authority of the sovereign, an individual who revolts or disobeys opposes the General Will of the society as a whole and threatens the web of coercion that enables each member of the society to enjoy freedom and the protection of their rights. It follows from this that individual actions that do not follow the laws laid down by the sovereign, the "legislative authority" mentioned above, are in violation of the Universal Principle of Justice, and are therefore unjust.

Korsgaard characterizes Kant as holding the view "that all governments should be taken to be legitimate," that "any regime's decisions are the voice of the General Will of its people; and its procedures for making those decisions must be taken to be the ones the people have agreed to" (Korsgaard 2008, p. 241). At the same time, she argues that "Kant of course does not mean that all governments and all of their decisions are perfectly just," because Kant's vision of government is an ideal to be worked towards (Korsgaard 2008, p. 241). If this is true, and a legitimate government can make unjust decisions, it follows that acting in accordance with duties of virtue, and following the Categorical Imperative (which requires that it's agent act only from a maxim that that she can at the same time will to be a universal law of nature) will at some point bring a citizen into conflict with the rule of societal law.

When these conflicts occur, Korsgaard argues that duties of virtue, by Kant's own reasoning, should take precedence:

Kant thinks that there are duties, and so ends, that belong specifically to the territory of virtue: the pursuit of the happiness of others, and the cultivation of our own talents. powers, and character. But ethics encompasses all of our duties. It is a duty of virtue to do the duties of justice from the motive of duty. In other words, justice itself is a virtue. And Kant says that the virtue of justice is possessed by one who makes the rights of humanity his end (Korsgaard 2008, p. 255).

Korsgaard's view is that duties of virtue apply to the arena of justice, and that all actions, in order to be moral, must be in line with our ethical duties. She argues that the potential for conflict between our duties to the Universal Principle of justice and our duties to the Categorical Imperative will sometimes lead a virtuous person to rebel against the sovereign, and the General Will, in order to create room for a fuller expression of morality and freedom.

AN UNKNOWN WILL

Korsgaard illustrates the difficulty in separating the concept of the collective will of the members of a society from the logically derived authority of a sovereign:

Kant's argument, as I've suggested, depends on a deeply procedural conception of the general will. Our general will, according to this argument, just is whatever follows from the procedures that make collective action possible, and so, in Rousseau's extravagant language, it can do no wrong. Suppose we allow, instead, that there is such a thing as the general will, independently of our procedures, and that our procedures should be viewed as a fallible device for ascertaining it. Then we can allow, contrary to Kant, that the extant regime may not represent the will of the people and so may fail to be legitimate. Even so, we get the problem. It is still true that the people cannot speak as a people until they have a voice. A revolutionary who claims to be the

representative of the people merely because of the spirit he senses among them or even because he has taken a favorable vote is misdescribing the situation. The people can only give their mandate through some duly constituted voice, through someone who has the right to represent them. If we admit the possibility that the extant regime does not represent the general will then there is *no way* to tell what the general will is (Korsgaard 2008, p. 251).

The uncertainty and fear caused by the difficulty of discerning the nature of the General Will can lead a Kantian thinker to minimize the implications of admitting that there is room for injustice in the decision process of a legitimate sovereign authority. But if we are going to discuss the General Will in terms of something that confers legitimacy and authority, we are implicitly assuming that the General Will exists. If we allow that the General Will exists, then it follows that its nature can be ascertained, given the correct procedures, whatever they might be. If we allow that, then we must conclude that the procedures currently available to us are fallible, and that it is possible for a sovereign authority to misinterpret to some degree the General Will of its society. I hold that civil disobedience is a technique of resistance that functions as an effective, though not perfect, tool for uncovering moral facts about the General Will of a given society.

CIVIL DISOBEDIENCE IS NOT WRONG

To restate Kant's Universal Principle of Right, "Any action is right if it can coexist with everyone's freedom in accordance with a universal law, or if it on its maxim the freedom of choice of each can coexist with everyone's freedom in accordance with a universal law" (Kant 1999, "The Metaphysics of Morals" 6:231). I argue that civil disobedience takes as its end the freedom of all humanity. Civil disobedience, if successful, opposes a hindrance to freedom by obstructing actions made by the sovereign authority that are not in line with the General Will. If unsuccessful, civil disobedience does no harm, and in fact strengthens the sovereign

authority by confirming its legitimacy in regard to the General Will. While the individual actions comprised by a program of civil disobedience may be "wrong," according to the Universal Principle of Right, the program taken as a whole is not wrong.

I agree with Korsgaard that, if we are to accept Kant's concept of political society as the only source of justice and also accept the idea of government being the expression of the General Will (no matter what form that government takes), no violation of the law can be justified in itself. When a person practices civil disobedience, however, their individual violations of the law are part of a larger project in service to the General Will. This project is necessarily in service to the General Will because it is dependent upon the will to achieve its aims. This project plays a role in modifying the policies and institutions of government, and helps them to become part of a more focused and accurate expression of the General Will.

PROTEST AND THE LEGISLATIVE PROCESS

Criminal action that causes a change in legislation does not carry an additional penalty for sedition. This is because the criminal did not, through his action, directly force the change in legislation. The citizens initiated the change (through voting, or by petitioning their elected representative, both of which would be legitimate procedures in a representative democracy) based upon a realization they had upon examination of the criminal's case. Forcibly changing the government would be opposing the General Will, and, on Kant's view, would be wrong. Influencing the General Will itself through the opinions of the members of society would not be wrong (Korsgaard 2008, p. 250).[1] Any changes in government brought about through civil disobedience are initiated by the government itself. A program of civil disobedience is intended to influence the worldview of a society's population, in the hope that it will pass that influence on in some form to its sovereign. Whether the sovereign is a single ruler who modifies his policies to appear benevolent in the eyes of the international community or

an elected body of representatives who pass legislation that they hope will placate their angry constituencies, the sovereign enacts reform freely, without any sort of direct coercion. The judgment as to whether the initial legislation did a poor job of interpreting the General Will is still, ultimately, done by a representative of government. Civil disobedience preserves the authority of the sovereign. It increases a population's confidence in the responsiveness of its sovereign, and provides proof of the sovereign's skill at divining and expressing the General Will. In this sense, a successful program of civil disobedience can actually strengthen the sovereign's legitimacy.

Unlike the prototypical form of rebellion, involving armed struggle against the representatives of the state and the intent of ultimately overthrowing the controlling regime, the practice of civil disobedience is not rooted in a total rejection of the state's authority. Instead, it is a passive plea for a re-examination of the government's expression of the will of its subjects that stands in contrast to Kant's "prohibition against actively resisting the will of the legislator" (Kant 1999, "On the common saying…" 8:299). The citizen who practices civil disobedience defies the will of the state, but offers his body in payment, allowing the state to confine, damage, or destroy it. This act indicates his recognition of the state's authority.

The practice of civil disobedience voices popular dissent as a passive statement that allows the government to conduct its negotiations with its discontented citizens on its own terms. It requires a faith on the part of its practitioner in the ability of the established authority to eventually recognize the practitioner's conception of justice. The practitioner's faith indicates his belief in the potential for wisdom and justice in his rulers. I would offer as an example the demonstrators who practiced civil disobedience during the American Civil Rights movement of the 1950s, who placed their trust not, I imagine, in the *actual* wisdom or kindness of their tormentors and adversaries, but a the hope for the development of those qualities.

CIVIL DISOBEDIENCE AND JUSTICE

It is not obvious that a program of civil disobedience is by definition a just program, in the sense of being in the service of a just, or at least, not unjust, cause. It is a tactic of resistance that might be used to push back against the will of all sorts of governments or organizations. It seems inaccurate to say that civil disobedience as a concept takes as its end the freedom of humanity by virtue of its method, in the same way as Korsgaard's conscientious revolutionary. At the same time, it does seem as though social cultures that practice non-violence tend to have relatively tolerant attitudes toward the diversity of human life—causes that are opposed to the freedom of humanity in general (by this I mean that they promote agendas that seek to redistribute the benefits and burdens of society using essentially arbitrary criteria, like race, ethnicity, or gender) seem to prefer more aggressive political action. A group that is actively opposed to the freedom of humanity in general could theoretically undertake a program of civil disobedience. If such a program were successful, and resulted in changes in legislation that reflected the political philosophy of that group's members, it would mean that they had correctly interpreted the inclination of their community's General Will. In relation to their sovereign authority, however, they have done no harm.

The General Will does not necessarily produce just actions or outcomes, but it is one of the few tools we have with which to build conceptions of justice. Rather than viewing morality as a pre-existing set of moral facts that we can deduce purely through reason, I conceive human morality as a set of observed tendencies of natural human behavior. To use Philippa Foot's example, a human being that does not keep his promises is a "defective" human being, in the same sense that a wolf that hunts alone, rather than with a pack, would be considered a defective wolf (Foote 2001, pp. 25-37). I believe that an educated populace, allowed to function in a peaceful and prosperous society would eventually, if we could accurately discern its will and translate it into social policy, move toward "universal" justice. A society that prized indi-

vidual autonomy and freedom of expression and of the exchange of ideas would be more likely, I believe, to accurately discern the moral tendencies that guide human moral practices.

Notes

1. "The government contains agencies for both determining and interpreting what the general will is. Of course the people may decide that the government is not doing a good job of this. But this judgment can only be made by someone who has the right to speak for the people, and that right belongs to the government itself. Therefore, the government can reform itself, but the people *as subjects* cannot reform government" (Korsgaard 2008, p. 250).

Bibliography

Kant, Immanuel. (1999) "The Metaphysics of Morals," in: M. Gregor (Ed), *Practical Philosophy/Immanuel Kant* (New York: Cambridge University Press)

_____. (1999) "On the common saying: That may be correct in theory, but it is of no use in practice," in: M. Gregor (Ed), *Practical Philosophy/ Immanuel Kant* (New York: Cambridge University Press)

Korsgaard, Christine M. (2008) "Taking the Law into Our Own Hands: Kant on the Right to Revolution," in: A. Reath, B. Herman, & C. Korsgaard (Eds), *Reclaiming the History of Ethics: Essays for John Rawls* (New York: Cambridge University Press)

Foot, Philippa. (2001) *Natural Goodness* (New York: Oxford University Press)

CONTRIBUTORS

Joshua Allen. B.A. Philosophy, Portland State University. Kant, Nietzsche, Ricoeur, Bachelard, Lefebvre, Lugones, Yancy, Phenomenology, Post-Structuralism, Hermeneutics, Feminist Philosophy. Joshua is currently a master's student at California State University, Los Angeles.

Theresa Cheng. Ethics, Moral Psychology. Theresa is currently an undergraduate student of philosophy at California State University, Los Angeles.

Melissa Foote. B.A. Philosophy, California State University, Los Angeles. Philosophy of Mind, Feminist Philosophy, Philosophy of Religion. Melissa is currently completing her master's degree at California State University, Los Angeles and intends to teach philosophy.

Jessica Gonzalez. B.A. Philosophy and Political Science, University of California, Los Angeles. Metaphysics, Philosophy of Language. Jessica has received her M.A. from California State University, Los Angeles and plans to teach philosophy at the community college level.

Paige Harrison. B.A. Cognitive Science, Occidental College. Epistemology, Metaphysics. Paige received her M.A. from California State University, Los Angeles and plans to teach philosophy.

Michael Hatcher. B.A. Philosophy, Biola University. Epistemology, Philosophy of Mind, Philosophy of Language, Metaphysics. Michael is currently a master's student at California State University, Los Angeles and intends to apply to Ph.D. programs in philosophy.

Roberto Lewis. B.A Film and English, University of California, Berkeley. Ethics, Political Philosophy. Roberto is currently a master's student at California State University, Los Angeles, and plans to apply to Ph.D. programs in speech and rhetoric.

Phil Lollar. B.A. Philosophy and TV, Radio, and Film, California State University, Fullerton. Ethics, Philosophy of Mind. Phil is currently a master's student at California State University, Los Angeles and intends to continue his work performing cartoon vocalizations.

Peter Miller. B.A. Philosophy, Chapman University. Philosophy of Self, Philosophy of Mind, Social/Political philosophy, Moral Psychology, Normative Ethics. Peter is currently a master's student at California State University, Los Angeles and intends to apply to Ph.D. programs in philosophy.

David Nagy. Ethics, Political Philosophy. David is currently an undergraduate philosophy student at California State University, Los Angeles, and intends to study Japanese overseas after graduation.

Anthony Ristow. B.A. Film Studies, University of Michigan. Continental Philosophy, Existentialism, Postmodernism. Anthony is currently a master's student at California State University, Los Angeles and intends to apply to Ph.D. programs in philosophy.

MASTER OF THE ARTS IN PHILOSOPHY
CALIFORNIA STATE UNIVERSITY, LosAngeles

The Department of Philosophy at California State University, Los Angeles offers a program of study leading to the Master of Arts degree in Philosophy. The program aims at the acquisition of a broad background in philosophy. It is designed for those preparing for further graduate study or community college teaching, and for self-enrichment. Although the department is analytically oriented, it encourages work in other areas, for example Asian philosophy, feminist philosophy, and interaction between European and Anglo-American thought. The Department includes faculty members with diverse backgrounds and interests actively working in a wide range of philosophical specialties. Classes and seminars are small with a friendly, informal atmosphere that facilitates student-faculty interaction.

The academic programs in philosophy at California State University, Los Angeles are intended to engage students in philosophical inquiry. They aim to acquaint students with noteworthy contributions by philosophers to the tradition; to explore various philosophical issues, problems, and questions; to provide students with principles of inquiry and evaluation relevant to the many areas of human activity, such as science, law, religion, education, government, art, and the humanities; to develop in them skills of analysis, criticism, and synthesis needed for advanced work in various scholarly fields; to encourage the development of skills and attitudes leading to self-reflection and life-long learning.

PHILOSOPHY IN PRACTICE
SUBMISSION INFORMATION

Each of the eleven student contributors was specially selected to submit a paper for this issue of *Philosophy in Practice* by one or more faculty members in the Department of Philosophy at California State University, Los Angeles. All writers are currently students in the master's program of philosophy, undergraduate students of philosophy, or have obtained degrees from the Master of the Arts program. All philosophy students at California State University, Los Angeles are eligible for nomination, and those who were chosen to contribute have demonstrated a superior ability to develop and compose works of advanced philosophical writing.

For more information on *Philosophy in Practice*, please contact:
philosophyinpractice@gmail.com