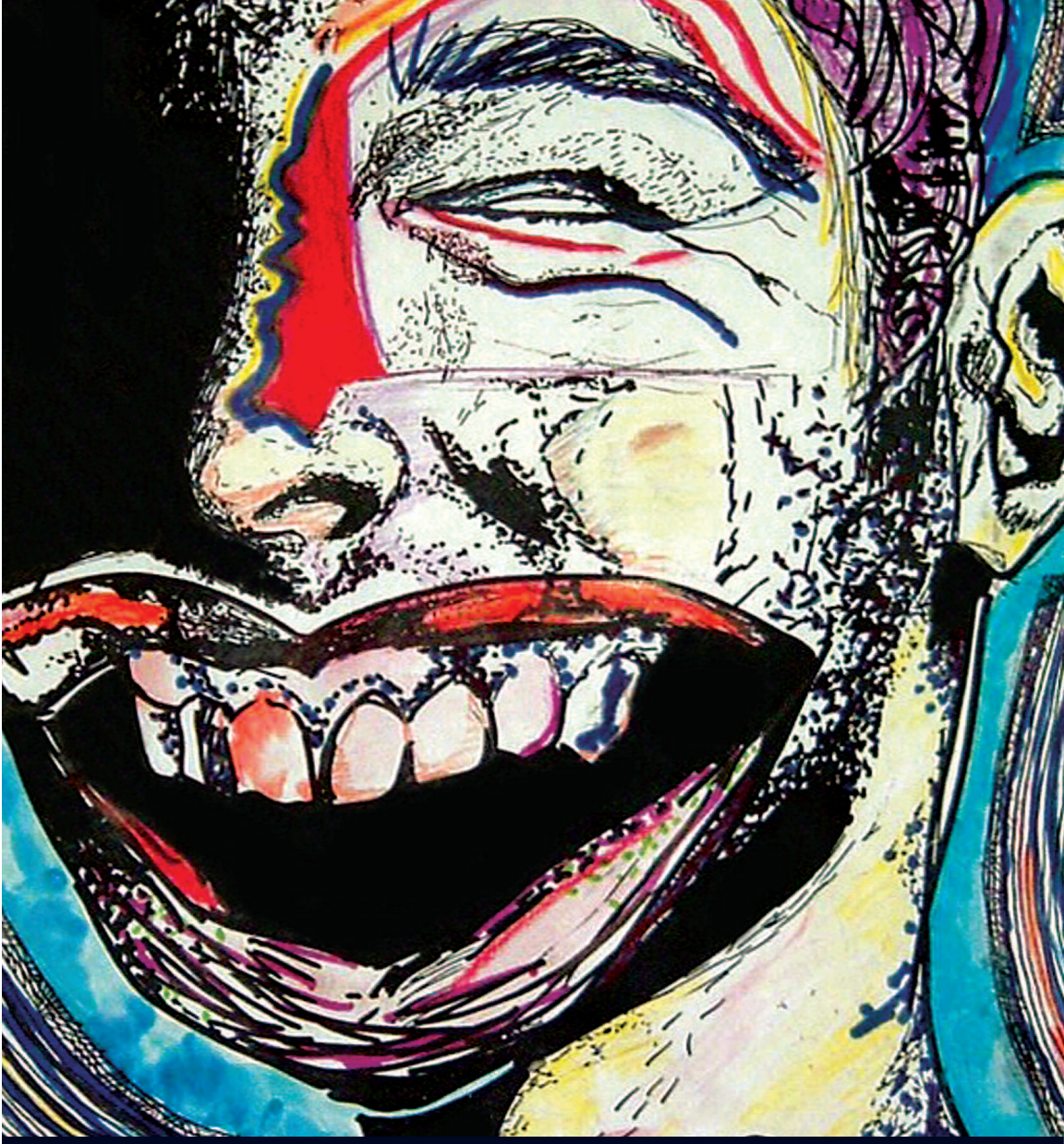
arché

journal of philosophy
volume 3



arché

Vol. III No. 1
Spring 2009

Journal of Undergraduate Philosophy
published at Boston University

Special thanks to:

The Student Union Allocations Board

Professor Daniel Dahlstrom, Chair of the Boston University Department of Philosophy

Provost David Campbell

NEH Distinguished Teaching Professor Diana Wylie

ἀρχή

ARCHÉ: A JOURNAL OF UNDERGRADUATE PHILOSOPHY AT BOSTON UNIVERSITY

ORDERING

Back issues are available for $10.00 per copy. Additional charges may apply for shipping overseas. To place an order and arrange payment, please contact the Managing Editor at buarche@gmail.com.

SUBMISSIONS

Undergraduate authors are invited to submit all types of philosophical discourse: research papers, critical analyses, conference proceedings, interviews, translations, letters and commentary on texts previously published in *Arché*, and original philosophical works. Submissions and correspondence should be sent as Word or .rtf attachment, with "Submission" in the subject heading, to buarche@gmail.com. Submissions may also be sent by mail to:

Undergraduate Philosophy Association
c/o Boston University Department of Philosophy
745 Commonwealth Avenue
Boston, Massachusetts 02215

Please include a 100-word abstract. The name of the author(s) should appear on the first page only. Citations should follow Chicago style, with all references occurring as endnotes. Authors are encouraged to query if they would like assurance that their submission is of a kind usually considered by the editors, or to follow-up if their submission has not been acknowledged in a timely manner.

Cover art: Katherine Otlowski

ISSN: 1946-1801

ἀρχή

CONTENTS

VOLUME 3, ISSUE 1: SPRING 2009

Boston University Department of Philosophy
Faculty 2008-2009

FULL TIME

Hugh Baxter
Alisa Bokulich
Peter Bokulich
Klaus Brinkmann
Tian Yu Cao
Daniel O. Dahlstrom
James Devlin
Juliet Floyd
Aaron Garrett
Charles L. Griswold
Jaakko Hintikka
Walter Hopp
Victor Kestenbaum
Manfred Kuehn
David Lyons
Krzysztof Michalski
Robert C. Neville
David Roochnik
John Silber
C. Allen Speight
Susanne Sreedhar
Daniel Star
Alfred I. Tauber
Judson C. Webb
Elie Wiesel

VISITING PROFESSORS

Lydie Fialova
Mihai Ganea
Susan James
Amelie Rorty

Introduction

Arché sprang from humble beginnings; initially, we merely aspired to provide a forum where deserving undergraduates can display their scholarly work. But in this third volume, we have grown bolder, and within these 64 slim pages we twice track through more than two thousand years of philosophy.

As any philosophical attempt has to begin with our faculties of reason, Rachel Bayefsky begins our traipse through history with an analysis of how those faculties can occasionally lead us in circles, and provides a diagnosis to help us avoid deductive vertigo. Our analytical skills now firmly rehearsed, Douglas Kremm, with a detailed analysis of two early Socratic dialogues, shows us how Socratic virtue can be unified without being impossible to achieve. Two millennia and 14 pages later, William J. Brady takes on contemporary epistemology with an intriguing and empirical refutation of Michael Huemer's dualistic theory of memory-beliefs. Finally, Peter Moore returns us to ancient Greece, enjoining us to seek out the enlightening harmonies hidden within the dissonant speeches of the Symposium.

The journey in these pages is varied and arduous, but I hope the reader will not find it too presuming. Philosophy today is as much a study of history as it is of any other discipline, and while any good scholar must know what has come before her, she must also be careful not to become bogged down in the past. I think that the quick flight through philosophy presented in the following papers will allow the reader to alight on the interesting ideas of our philosophical forefathers, without denying her room to pursue novel thoughts as well.

Ted Stinson
B.A. Philosophy, 2009
Boston University

A Diagnosis for the Circularity of Anti-Skeptical Arguments

RACHEL BAYEFSKY
Yale University

Arguments against skeptical hypotheses often raise the concern that their premises arise through a method that would be unreliable if the skeptical hypothesis turned out to be true. Such arguments may be considered guilty of circularity if their premises rely on the falsity of the skeptical hypothesis in an inappropriate manner—a manner I intend to clarify. In this paper, I plan to develop a diagnosis for the circularity of anti-skeptical arguments. The diagnosis, which I will call D, will be partially based on Robert Nozick's "subjunctive conditionals" account of knowledge. I will apply D to various anti-skeptical arguments. In particular, I intend to use D to contend that Descartes' proof of the truth of his clear and distinct perceptions is not necessarily circular.

I will argue that D provides an explanation for the way in which the effectiveness of an argument is crippled by circularity. The effectiveness of an argument is its ability to take people who are committed to evaluating arguments based on rational standards and bring them closer to the conclusion than they are when they first encounter the argument. An effective argument in this sense is not one that persuades by rhetorical force, but one whose logical steps make progress towards the conclusion. Circularity impedes an argument's effectiveness because the conclusion must be accepted before the premises, which means that the argument itself is not responsible for advancing the conclusion. Diagnosis D, by clarifying the reasons why the conclusion of a circular argument against a skeptical hypothesis must be accepted before at least one of the premises is accepted, provides us with a way to explain the decrease in the effectiveness of an argument due to circularity.

Here are some examples of arguments that might be considered circular. First, let us say that I am suddenly seized with the worry that the floor of my building, instead of being connected solidly to the ground, is floating on air. I try to reassure myself with the thought that, if I wanted, I could walk out of my room right now and go down the elevator in the ordinary way. Here is my "argument:"

 Arché, Volume 3, Issue 1: Spring 2009. pp. 3-13

(1-e): If the floor of my building has suddenly begun to float on air, I cannot walk out of my room and go down the elevator in the ordinary way.

(2-e): I can walk out of my room right now and go down the elevator in the ordinary way.

(C-e): Therefore, the floor of my building has not suddenly begun to float on air.

Alternatively, suppose a congenitally blind person receives surgery and slowly begins to recover her ability to experience visual imagery.[1] She sees the outline of a car in front of her, considers the situation, and reaches the fortunate conclusion: "My blindness has been at least partially cured!" Her "argument" runs as follows:

(1-b): If my congenital blindness has not been partially cured, I cannot see the outline of a car in front of me.

(2-b): I am able to see the outline of a car in front of me.

(C-b): Therefore, my blindness has been partially cured.

Are the "elevator" and "blindness" arguments circular? Not in an obvious way: neither conclusion appears explicitly among their respective premises. Clarifying the question of whether and how such arguments are circular is not a simple task. One response to the "elevator" argument, for instance, might be to object that if the conclusion is false, premise (2-e), at least, is also false: if the floor of my building has suddenly begun to float on air, I am unable to walk out of my room and take the elevator down. But for all valid deductive arguments, if the conclusion is false, one or more of the premises must be false. This does not imply that the truth of a premise "relies on the truth of the conclusion" in an inappropriate manner. Take, for instance, the classic argument:

(1-s): All men are mortal.

(2-s): Socrates is a man.

(C-s): Therefore, Socrates is mortal.

Here, the falsity of the conclusion would necessarily undermine the truth of at least one of the two premises. If the "Socrates" argument is circular, however, it is difficult to see how many valid deductive inferences could still stand.

I. A Diagnosis of Circularity

My diagnosis for the circularity of an argument will focus on a person's reasons for accepting the premises of the argument and whether or not they would be replicated in the case that the conclusion is false. In formulating my diagnosis, I will adapt one of the conditions of Robert Nozick's account of knowledge in Philosophical Explanations. Nozick presents a "subjunctive conditionals" account of knowledge, according to which S knows that p if and only if:

(1) p is true:

(2) S believes that p:

(3) If p were not true, S wouldn't believe that p; and,

(4) If p were true, S would believe that p.[2]

The idea behind conditions (3) and (4) is that one's beliefs should only count as knowledge if they are sensitive to the truth. That is, if one would believe p whether or not p is actually true, one's beliefs do not have the epistemic status of knowledge. Here, Nozick is presenting an account of knowledge, not putting forward criteria for evaluating circularity. His third criterion for knowledge, however, can be adapted to form part of my diagnosis D for circular arguments against skeptical hypotheses. According to D, an argument against a skeptical hypothesis is circular with respect to a specific premise p[3] if and only if the following conditions hold:

D1: If the conclusion were false, S would still believe that p for the same reasons as if the conclusion were true.

D2: If the conclusion were false, p would be false.

Together, these conditions entail the condition "If p were false, S would still believe that p for the same reasons as if the conclusion were true." While we might simply use this condition as an indication of circularity, it is useful to break it down into D1 and D2 in order to see why different types of arguments beg the question.

The basic idea behind D is that a circular argument, in the most general form, is one whose conclusion is contained in the premises of an argument. But it is not necessary for a conclusion to be explicitly present among the premises in order for an argument to be circular. Rather, an argument against a skeptical hypothesis is circular if the truth of the conclusion must be accepted before the reasons for at least one of the premises can count as justification for the truth of that premise. If I have the same reasons for believing the premise regardless of whether the conclusion is true, and the truth of the premise depends on the truth of the conclusion, then my reasons for

believing the premise do not point me towards the truth of the premise unless I am already assured of the truth of the conclusion. So the purpose of D is to evaluate the strength of the reasoning behind the acceptance of the premise. D explains that an argument against a skeptical hypothesis is ineffective because the conclusion must already have been accepted in order for the premises to be evaluated. Such an argument does not succeed in using the premises to prove the conclusion.

Before testing D against examples of potentially circular arguments, I would like to explain why D1 includes the clause "for the same reasons," as in, "if the conclusion were false, S would believe that p *for the same reasons* as if the conclusion were true." The addition of "for the same reasons" rules out a situation in which S would believe that p for different reasons than if the conclusion were false. In such a scenario, S's reasons could still serve as independent evidence for p even though S would believe that p if the conclusion were false, because S might be able to distinguish the types of reasons available if the conclusion is true from the types of reasons available if the conclusion is false. For instance, let us say the conclusion of an argument is that aliens have landed on earth. One of the premises of the argument is that three aliens have been spotted in a field in Northern Ontario. If the conclusion is true, S's reasons for believing the premise might include a news report; if the conclusion is false, S's reasons for believing the premise might instead be limited to a dramatic dream-vision. If the conclusion were false, therefore, S would believe that three aliens had been spotted for a different reason than if the conclusion were true. S might well be able to use the differences in his reasons for believing the premise to evaluate the truth of the conclusion. The fact that S would fail to meet the criterion laid out in D1 reflects the fact that his argument is more effective than an argument judged to be circular under D. This is because S would not necessarily have to accept the conclusion in order to evaluate the evidence for the premise.

II. Evaluating the Conditional

How are we to understand the conditional that is D1? One way of answering this question is to appeal to the possible-worlds account of subjunctive conditionals. We can follow the example of Nozick, who explains the application of his third criterion for knowledge as follows:

> "What the subjunctive 3 [if p were false, S wouldn't believe that p] speaks of is the situation that would hold if p were false. Not every possible situation in which p is false is the situation that would hold if p were false. To fall into possible worlds talk, the subjunctive 3 speaks of the not-p world that is closest to the actual world, or of those not-p worlds that are closest to the actual world, or more strongly (according to my suggestion) of the not-p neighborhood of the actual world. And it is of this or these not-p worlds that it says (in them) S does not believe

> that p. What happens in yet other more distant not-p worlds is no concern of the subjunctive 3."[4]

In other words, in order to evaluate whether the subjunctive conditional 3 holds, we identify the possible world or worlds that are closest to the actual world but that differ from the actual world in the sense that p does not hold. Then we determine whether S believes that p in the closest not-p world or worlds. Applied to D1, in order to evaluate whether S would believe that p if the conclusion were false, we must identify the possible world or worlds closest to the actual world in which the conclusion is false and then determine whether S believes that p in these worlds.

The concept of a "closest possible world or worlds" is, of course, problematic. How do we determine which possible world is closest? Can there be two possible worlds which are different from each other, but equally "close" to this world? The coherence of the "closest possible world" concept has been discussed extensively in the literature on Nozick and others, and I cannot provide a compete solution of this problem here. In general, however, the "closest possible world" is one in which we see the fewest changes from the actual world in terms of physical laws, objects and people that are present, and historical trajectory. Since it is possible for some of these conditions to change while others remain the same, it may be hard to determine, especially in close cases, which worlds count as "closer" than others. The difficulty of choosing the closest possible world when two worlds are rather similar, however, should not preclude us from acknowledging that the distinction is meaningful when applied to possible worlds that are radically different. An example may be instructive. Let us say I am sitting in a park on an ordinary day. Now let us think of possible worlds in which it is false that I am sitting in the park. In possible world W1, I am not sitting in the park because I decided to go to the grocery store instead of the park that morning. In possible world W2, I am not sitting in the park because I have been pulled up into the sky by a reverse gravitational force, which took effect only for one instant and only in one place on earth—the very park bench where I was sitting.

It seems plausible to suggest that W1 is closer to the actual world than W2. This is because although W1 differs from the actual world in that I am no longer sitting in the park, it differs in a way that is predictable and routine from the perspective of someone sitting in the actual world. W1 involves, for instance, no reversals of the laws of nature of the kind that take place in W2.

The notion of possible worlds on which I will be relying is comparative. I will be determining which of different possible worlds is closest to the actual world, not attempting to delineate the exact features of the possible world that is absolutely closest to the actual world. Even without an exhaustive list of characteristics that would make a possible world closest to the actual world, it is possible to determine which of a selection of worlds is most likely to be closest. But the notion of the closest possible world, and the diagnosis in general, must be clarified through examples.

III. Testing the Diagnosis

Let us return to the "elevator" example, the argument against the skeptical hypothesis that my floor of the building is floating on air:

> (1-e): If my floor of the building has suddenly begun to float on air, I cannot walk out of my room and go down the elevator in the ordinary way.
>
> (2-e): I can walk out of my room right now and go down the elevator in the ordinary way.
>
> (C-e): Therefore, my floor of the building has not suddenly begun to float on air.

We must evaluate both premises of the argument to determine whether or not D1 and D2 hold. Let us consider what would happen if (C-e) were false. If my floor of the building had suddenly begun to float on air, I would be likely to fall to the ground as soon as I walked out of my room. It seems almost impossible for my floor of the building both to be floating on air and to be connected solidly to the ground by an elevator shaft. In the closest possible world in which (C-e) were false, then, (1-e) would be true. I could not simply ride down the elevator. So D2 does not hold with respect to (1-e). Now we diagnose the argument's circularity with respect to premise (2-e). If my floor of the building has suddenly begun to float on air, I have the same reasons for believing that I can walk out of the room and go down the elevator as I would have had if my floor were as solidly connected to the rest of the building as ever. I have gone down the elevator many times before, I have no particular reason to suppose a great anomaly in the earth's physical laws has occurred, and so on. So D1 holds with respect to (2-e). Moreover, if (C-e) is false, (2-e) would also be false; if my floor of the building is floating on air, I cannot go down the elevator in the ordinary way. Both D1 and D2, then, hold for the elevator argument with respect to premise (2-e), and so the elevator argument is circular according to diagnosis D.

How does the diagnosis of circularity point out the "elevator" argument's ineffectiveness? All of my evidence for believing the premise "I can walk out of my room and go down the elevator" is compatible with the premise's falsehood. This premise will only be plausible if I already believe that my floor is not floating on air. But if I already believe that my floor is not floating on air, the anti-skeptical argument has not been responsible for bringing me any closer to the conclusion that my floor is not floating on air. If I am only soothed by rational arguments, then the "elevator" argument will not be particularly soothing.

Secondly, let us test D on the "blindness" argument:

> (1-b): If my congenital blindness has not been partially cured, I cannot

see the outline of a car in front of me.

(2-b): I am able to see the outline of a car in front of me.

(C-b): Therefore, my blindness has been partially cured.

The truth of (1-b) follows from the nature of congenital blindness at this stage of scientific research. Congenitally blind people are widely thought to be unable to experience visual imagery at any time, including in their dreams.[5] So D2 would not hold of (1-b), because (1-b) is never false. The question is therefore whether or not the "blindness" argument is circular with respect to premise (2-b). There are multiple scenarios under which a women's blindness has not been partially cured. One such scenario is a situation in which this women seems to be able to see but has not actually been cured of blindness. Say the women is suddenly abducted by aliens and made into a bodiless brain in a vat (BIV) with no eyes at all. It would be possible for the aliens running the BIV system to stimulate the neurons responsible for sight in the women's brain. The aliens could give her an experience of seeing the outlines of a car without her blindness actually having been cured. As soon as the aliens withdraw their tentacles, affix the women's brain to her body once more, and let her go, the women will be as blind as ever.

If we turn to the BIV world when we consider what happens when the conclusion of the woman's argument ("My blindness has been partially cured), then the woman's argument would be circular. D2 would apply if the woman is not actually seeing the outlines of the car. This would be because "seeing" requires some causal connection between an external source of visual images and the eye. Though the BIV experience provides images of the outline of a car to the blind woman, these images correspond to nothing in the outside world, and so the woman is not actually seeing the outlines of a car. D1 would hold because the woman would have the same reasons to believe she is seeing the outlines of a car whether or not her blindness is partially cured: her sensations and ideas of the car would be the same. So the woman's argument that her blindness has been at least partially cured is circular under D if the closest possible world is the BIV world.

But why should we base our judgment of whether D1 holds in the possible world in which the woman is a BIV? As Nozick points out, the subjunctive conditional "if p were false then S wouldn't believe that p" can be true "even though there is a possible situation where not-p and S believes that p," are both true, because "not every possible situation in which p is false is the situation that would hold if p were false."[6] Regardless of whether D1 would hold in all of the situations in which (C-b) is false—and it would hold in the BIV world—a *closer* possible world in which (C-b) is false is a world where the surgery was simply unsuccessful. In this world, if (C-b) were not true, the woman would not believe (2-b). If the woman's blindness were not partially cured, she would not think she could see the outlines of a car. D1, therefore, does not hold, and the "blindness" argument is not circular according to D. This is a consequence of the fact

that the scenario in which the woman is a BIV being fed images by aliens is farther away from the actual world than the scenario in which the surgery simply did not work.

The fact that the "blindness" argument is not circular under D provides us with a reason why it is more effective than the "elevator" argument. The congenitally blind woman receives sensory impulses that she recognizes, based perhaps on the testimony of others, as sight. She plausibly supposes that the most likely scenario under which these impulses appear to her is the scenario in which she can actually see. She does not have to believe that her blindness has been cured in order to realize that she is actually seeing; rather, she uses the fact that she is actually seeing to deduce that her blindness has been cured. The argument is therefore more effective than the "elevator" argument because it has brought the woman closer to its conclusion than she was when she began considering the premises.

IV. Application to the Cartesian Circle

Now that I have provided a diagnosis of circular arguments and applied it to the "elevator" and "blindness" arguments, I would like to apply my diagnosis to the Cartesian Circle. I intend to show that the Cartesian Circle is not necessarily circular under my diagnosis. In the Meditations, Descartes seeks to argue for the conclusion that as a rule, his clear and distinct perceptions are true.[7] One potential analysis of Descartes' argument is Keith DeRose's interpretation, according to which Descartes argues as follows:[8]

> (1): For certain propositions, such as the cogito, 2+3=5, and the causal principle of ideas (i.e. that there must be "at least as much reality in the cause as in its effect")[9], if I clearly and distinctly perceive these propositions, then they are true.
>
> (2): **God exists and is no deceiver.
>
> (3): I clearly and distinctly perceive the Rule of Truth: for all p, if I clearly and distinctly perceive that p, then p is true.
>
> (C): All of my clear and distinct propositions have the status of *scientia*, that is, certain knowledge that cannot be doubted.

The ** represents Descartes' argument for the existence of a non-deceiving God. According to this presentation of Descartes' argument, before Descartes recognizes the existence of a non-deceiving God, he can attain a fairly high level of certainty regarding his clear and distinct perceptions. But even a small doubt of the general rule of truth can undermine Descartes' acceptance of (1) by casting doubt on the reliability

of the belief-forming mechanism that results in (1): clear and distinct perception. Descartes must therefore prove the existence of a non-deceiving God in order to clearly and distinctly perceive the general rule of truth, which tells him that as a rule his clear and distinct perceptions are true. When Descartes reaches clear and distinct perception of the general rule of truth, all of his clear and distinct perceptions are raised to a higher level of certainty called *scientia*. The advantage of *scientia*, as opposed to mere clear and distinct perception, is that it is impervious to skeptical attack arising due to doubt of the general rule of truth.

We can now examine whether this two-level version of Descartes' argument exhibits circularity according to D. I will concede Descartes' argument for the existence of a non-deceiving God based on the clear and distinct perceptions in (1).[10] I will also accept the link between (3) and "C", because the reasons for why (3) leads to "C" are contained in the idea of *scientia*. In order to compare this argument to the others I have considered, I will reformulate it as follows:

> (1-d): If I clearly and distinctly perceive certain propositions, then they are true. These propositions include 2+3=5, the cogito, and the causal principle of ideas.
>
> (2-d): There exists a non-deceiving God who can guarantee the truth of my clear and distinct perceptions.
>
> (C-d): I clearly and distinctly perceive the Rule of Truth: for all p, if I clearly and distinctly perceive that p, then p is true.

Since I have conceded Descartes' argument for a non-deceiving God, D2 does not apply with respect to (2-d). Whether or not (C-d) is false, I will have to concede that (2-d) is true. Since the argument is not circular under D with respect to (2-d), I will ask whether the argument is circular under D with respect to (1-d). To test D1, we ask: if (C-d) is false, would I still believe (1-d) for the same reasons as if (C-d) were true? To test D2, we ask: if (C-d) is false, would (1-d) be false?

The question about D1 can be answered in the affirmative. Even if it is not the case that as a rule, every single one of my clear and distinct perceptions are true, I would still believe that I clearly and distinctly perceive certain propositions to be true. The closest possible world in which the general rule of truth is false is not the world ruled by an utterly vindictive evil genius who ensures that all of my clear and distinct perceptions are false. A closer possible world in which (C-d) is false is one in which some, but not all, of my clear and distinct perceptions are true, including those I seem to perceive most clearly and distinctly, such as "I exist." In such a world, I would indeed have the same reasons for thinking "I exist" is true; it would seem impossible for me to doubt this proposition, it would conform with everything I have experienced so far, and so on. D1, in other words, would be satisfied: if (C-d) were false, I would still have the same reasons for believing (1-d) as if (C-d) were true.

D2, however, would not be satisfied. Even if the general rule of truth were false, it does not have to be the case that all of my clear and distinct perceptions are false. In the closest possible world in which the general rule of truth, (C-d), is false, the proposition "I exist," and likely some other clear and distinct perceptions, could still be true. It is then not the case that the falsity of (C-d) implies the falsity of (1-d). My reasons for believing in the truth of certain clear and distinct perceptions such as "I exist," though these reasons would be present even if the rule of truth were false, are still able to support (1-d) independently of the falsity of the general Rule of Truth. Descartes' argument on DeRose's interpretation, therefore, is not circular according to D.

If we concede Descartes' argument for the existence of a non-deceiving God, diagnosis D can explain the effectiveness of Descartes' argument for the truth of his clear and distinct perceptions. Descartes can accept the truth of some of his clear and distinct perceptions without having previously accepted that all of them, as a rule, are true. So he can use the initial clear and distinct perceptions as springboards to prove the existence of God and thereby the truth of all his clear and distinct perceptions. The argumentative steps in Descartes' proof can bring someone closer to the general Rule of Truth than he or she was at the beginning of the argument. Descartes does not have to assume a world of widespread clear and distinct knowledge in order to work his way up to such a world. His argument, diagnosis D shows us, is thereby rendered more effective.

In this paper, I have sought to provide a diagnosis for the circularity of arguments against skeptical hypotheses and to argue that Descartes' argument for the truth of his clear and distinct perceptions is not necessarily circular. I suggested that the "elevator" argument was circular under D because my reasons for supposing I could walk outside and take the elevator only counted as reasons for this premise provided the conclusion of the argument—that my floor was not floating on air—was true. I also argued that the "blindness" argument was not circular under D because the woman can believe she is able to see for reasons that would not be available in the closest possible world in which she is blind: a world in which her continuing blindness is due to failed surgery instead of her being a BIV. In the "blindness" but not the "elevator" argument, the premises can be accepted on the basis of the evidence for them without previous commitment to the conclusions. The premises of Descartes' argument for the truth of his clear and distinct perceptions, in its two-level variety, can also be accepted independently of the conclusion, for even if the conclusion does not hold, the premises—individual clear and distinct perceptions—could still stand.

I have also suggested that D provides a way to explain why circularity detracts from an argument's effectiveness. In a circular argument, one cannot accept at least one of the premises without also accepting the conclusion, but according to the logic of the argument, one should not be accepting the conclusion without accepting the premises. One therefore has no reason to accept either the premises or the conclusion unless one has already accepted both—but if one has already accepted both, the argument has not been responsible for one's acceptance of either. Diagnosis D, by point-

ing out the conditions under which the acceptance of the premises is contingent on the prior acceptance of the conclusion, shows us why circular arguments are less effective.

To conclude, when calling an argument against a skeptical hypothesis "circular"—a claim that is intended to be devastating—it is useful to provide an analysis of why this is so. The advantage of providing a diagnosis of circularity is that it helps us to distinguish between circular and non-circular arguments in cases where our intuitions do not give us much insight into the argument's logical structure. Exploring the sources of circularity can encourage us to avoid circular anti-skeptical arguments and can show us the importance of appealing to those who can be properly convinced: those who are open to rational persuasion and who are not already persuaded.

Received February 2009
*Revised March 2009**

* *My thanks to Geoffrey Pynn, of Northern Illinois University, for his very helpful comments and suggestions.*

ENDNOTES

1: It has recently been shown that a congenitally blind person is capable of recovering vision even after twelve years of blindness. See: Ostrovsky, Andalman, and Sinha, "Vision Following Extended Congenital Blindness," *Association for Psychological Science Research Report* 12 (17), 1009-1014, 2006.

2: Robert Nozick, *Philosophical Explanations.* (Cambridge MA: Harvard University Press, 1981), 172-176.

3: Diagnosis D must be evaluated with respect to all the premises of an argument, because the purpose of D is to show that an argument is circular if S would still believe at least one of the premises even when this premise is false. If it can be shown that at least one of the premises fits the criteria outlined in D1 and D2, then the argument is circular under D. In order to show that D does not apply to an argument, one must show that D does not apply to any of its premises.

4: Nozick, 199.

5: C. Hurovitz, S. Dunn, G.W. Domhoff and H. Fiss, "The dreams of blind men and women: A replication and extension of previous findings," *Dreaming* 9 (1999): 183-193.

6: Nozick, 156.

7: René Descartes, "Meditations on First Philosophy," from *Selected Philosophical Writings*, trans. Cottingham, Stoothoff, and Murdoch, (New York: Cambridge University Press, 1988), 106.

8: Keith DeRose, "Descartes, Epistemic Principles, Epistemic Circularity, and Scientia," *Pacific Philosophical Quarterly* 73 (1992): 224.

9: Descartes, 96.

10: In particular, Descartes uses the causal principle of ideas to assure himself that the cause of the idea of a perfect God—that is, God—must be at least as real as the idea itself.

The Unity of Virtue: Toward a Middle Ground Between Identity and Inseparability in Socratic Virtue

DOUGLAS KREMM
University of Pittsburgh

At the center of much of the discussion about the virtues in the earlier Socratic dialogues lies an issue concerning the relationship among those virtues—specifically about whether or not they are necessarily linked in some way. In the *Protagoras* and the *Laches*[1] (the dialogues to which I will confine most of my discussion here for reasons mentioned below) it is clear that Socrates[2] subscribes to a view in which the virtues are somehow *united*. It is far from clear, however, what specific form that unification is supposed to take. Although he does at times appear to be arguing for his own stance on whether the virtues are unified, the dialogues are generally structured in such a way that it is not Socrates, but his interlocutors, who are in the position to defend a given point of view. For this reason, much of what Socrates says about the exact relationship among the virtues is left unexplained.

In most of the scholarship concerning Socrates' doctrine of the unity of Virtue, there has been a tendency to endorse one of two interpretive stances, neither of which, unfortunately, seems sufficiently able to make sense of seemingly contradictory parts of the dialogues. Daniel Devereux assesses the situation quite well in his article "The Unity of the Virtues,"[3] but I will argue that the conclusion he draws there is not entirely satisfactory. In the first two sections of this paper, I will discuss the two alternative interpretations of the doctrine of the unity of Virtue, the Inseparability View and the Identity View, in order to bring to light the problems of each. But whereas Devereux is content with referencing external accounts (e.g. those of Xenophon and Aristotle) of the historical Socrates and attributing the problematic areas to an inconsistency in Plato's representation of that historical figure[4], I will suggest an interpretation that makes Socrates' claims self-consistent. Central to this effort will be a general redefining of what Socrates is arguing about, and on that front I will share an important starting point with Terry Penner's "Identity View." I will argue, however,

 Arché, Volume 3, Issue 1: Spring 2009. pp. 15-30

that Penner's strict Identity View gives rise to its own complications, and that there is a better way to understand Socrates' doctrine of the unity of Virtue. My larger aim is to suggest a way of thinking about inseparability and identity which will avoid the pitfalls of the Inseparability View as they are outlined in Section I, as well as the puzzling ontological implications I find in the Identity View.

I. The Inseparability View

The two most thoroughly argued interpretations of the doctrine of the unity of Virtue differ primarily in the strength of the respective claims that they understand Socrates to be making. The weaker Inseparability View takes Socrates to be asserting an equivalence of all the virtues—what Gregory Vlastos calls a "Biconditionality Thesis."[5] According to this view, the separate virtues remain distinct parts of a whole (Virtue), but they are necessarily coinstantiated. That is, having one of the virtues is a necessary and sufficient condition for having all the others, which is to say that X is courageous if and only if X is also just, temperate, pious, etc. It is important for this view, though, that the individual virtues remain distinct: courage is not *identical* to justice, justice is not *identical* to piety, etc. Whatever these virtues turn out to be, they are all distinct parts of a whole. The central claim is just that an agent cannot be courageous unless she is also just, temperate, pious, etc. Any argument for the Inseparability View relies heavily on Socrates' remarks at the end of the *Laches* (esp. 198d-200) where he seems to equate the whole of Virtue with knowledge of all goods and evils. At 199e5-6, for example, Socrates tells Nicias that "the thing you are now talking about [knowledge of all goods and evils, from 199d1], Nicias, would not be a part of virtue *but rather virtue entire* [emphasis added]." The salient idea there is that knowledge of all goods and evils just is Virtue, and it is necessary that anyone who has *any one* of the individual virtues necessarily has that knowledge which guarantees possession of all the other virtues.

The textual evidence for the Inseparability View comes from two premises drawn from the early dialogues. The first,

> (1) X is virtuous if and only if X is characterized in some way by moral knowledge or wisdom,

is a generalized inference from the many instances where Socrates and his interlocutors take it for granted that virtue is a "fine and noble thing," and that something cannot be fine and noble without being in some way attentive to or characterized by wisdom (see in particular *Laches* 192c6 and *Protagoras* 349e4-350c5). This can apply to either an action or an agent. A reckless or shortsighted action performed in dangerous circumstances, for example, would not really be virtuous (and thus not courageous), because it may also have been foolish. A foolish action is hardly commendable, and so, the argument goes, it is not virtuous. So there must be wisdom or knowledge

at play somewhere in the picture, and if we are talking about some individual virtue, the relevant kind of wisdom is of course going to be moral wisdom, wisdom in some way pertaining in some way to action.[6] The second premise, established in the argument at the end of the *Laches* (199b7-8), is

> (2) The same knowledge is of the same things whether future, present, or past,

which, together with (1), establishes a strong case for the Inseparability View. To speak of agents, if X has any one of the virtues, then she must have moral wisdom or knowledge (from (1)), and if X has moral wisdom or knowledge, then she must have knowledge of *all things pertaining to moral action*, (from (2)). And of course knowledge of all things pertaining to moral action will ensure that the agent is virtuous in all possible ways—that is, that she is courageous, just, temperate, pious, etc.

There is an underlying assumption fueling this interpretation, and that is that the "What is X" questions Socrates asks (what is courage, justice, piety, temperance, etc.) are questions after something like the "meaning" or "essence" of X.[7] A proper response to these questions, then, would be some sort of logical analysis about what we have in our conceptual understanding of X. The question we're asking is, "What do we mean when we say that F is X?" A satisfactory answer to these conceptual questions will consist in a definition of, say, courage, that uncovers the essential features at play when we say something like "F is courageous." To do this, we might begin with a definition that seems in some respects inadequate—at 192c1 in the *Laches*, for example, courage is "endurance of the soul"—and then work toward a more refined definition. This requires some logical analysis to yield what is really essential to courage; that is, what we *mean* when we say that F is courageous. A somewhat hackneyed example will be helpful here: take the sentence

> (3) F is a bachelor.

If the term "bachelor" were in question in the Socratic dialogues, proponents of the Inseparability View would find the solution in an analysis that yields from (3) both

> (3a) F is male, and

> (3b) F is unmarried.

This is to say that it is contained within what one means when she says "F is a bachelor" both that "F is male" and "F is unmarried." So to answer the *conceptual* question of what a bachelor is, we merely have to give a definition that uncovers these underlying, defining features of "bachelor" in a way that makes one better understand the term.

Socrates' argumentative moves, on this view, are seen as a way of showing why it must be the case that the meaning of the term "courage" is such that it carries with it

other claims (about wisdom, necessary coinstantiation, etc.). The analysis would run something like this:

(4) F is courageous.

From this, Socrates' various arguments are taken to yield

(4a) F is wise. (See the discussion of (1) above for why this is so.)

and as a result he concludes

(4b) F is just, temperate, pious, etc.

This is what constitutes the Biconditionality Thesis, and it can be schematically represented as follows:

$\forall f\,(Cf \leftrightarrow Wf \leftrightarrow Jf \leftrightarrow Tf \leftrightarrow Pf)$

Within this construction, any of the virtues can be freely substituted for any of the others without producing a false statement. This amounts to the coinstantiation of the virtues which the Inseparability View takes as Socrates' main point. If Socrates' arguments succeed, then, he will have shown that the coinstantiation represented above is always operative in any talk of courage, justice, piety, etc.

This interpretation works quite well for reading the *Laches*, but it encounters some serious difficulties when confronted with certain parts of the *Protagoras*. Indeed, by 333b in that dialogue, Socrates thinks he has shown that the pair of propositions

(i) For one thing there is only one opposite; and

(ii) Wisdom is different from temperance yet both are parts of Virtue,

are mutually untenable. If *folly* is the opposite of both wisdom and temperance (332-332e14), then (i) and (ii) cannot be held simultaneously. If (i) is right, then wisdom and temperance cannot be different from each other, because they have the same opposite. Wisdom and temperance, then, are either *one thing*, or there was some misunderstanding in the talk about opposites. It seems clear, though, that both Socrates and Protagoras are less willing to abandon (i) than (ii), since the conversation after that point moves toward a reluctant agreement on Protagoras' part that "wisdom and temperance are one thing" (333b6).[8]

The real problem here is that Socrates seems to begin to abandon altogether the view that there are "parts of Virtue" at all. We can see this if we consider that the *Protagoras* is structured in general as a debate over two possible ways in which Virtue might be "one." Either "virtue is a single thing, with justice and temperance and piety

its parts" or these things are "all names for a single entity" (329d1-3). And when Protagoras endorses the former (329d4-5), much of the proceeding dialogue is largely an attempt on Socrates' part to refute that claim. This is made even clearer in 349b, where Socrates summarizes the main point of contention:

> Wisdom temperance, courage, justice, and piety—are these five names for the same thing, or *is there underlying each of these names a unique thing, a thing with its own power or function, each one unlike any of the others?* "[emphasis added]"

At this point, it seems clear that the primary issue at stake is not one of similarity, or even of inseparability, but rather one of identity or non-identity. The dilemma as Socrates formulates it above concerns whether or not the so-called "separate virtues" are actually just *one thing* which admits of different names. And, since most of Socrates' arguments in this dialogue (concerning piety and justice in 330c-331b7; wisdom and temperance in 332-333b8; wisdom and courage in 349e-350c5) are aimed at refuting the claim that the virtues are non-identical, it would seem that he is ready to endorse the other horn of the dilemma—namely, that all the "parts of Virtue" are just five names for the same thing.

The problem the Inseparability View faces here is that it must ignore or distort the instances in the *Protagoras* where Socrates seems to be situating himself against the view that there are actually parts of Virtue. There are two ways to deal with this troubling inconsistency. The first option is to take Socrates' argument in the *Protagoras* as merely an attempt to undermine Protagoras' claims. But to read the *Protagoras* in this way is likely to preclude any possibility of understanding Socrates' stance on the issue, since it reduces many of his claims to mere argumentative tools. This reading would discredit not only the Inseparability View, but all of Socrates' interpretations.[9] Furthermore, there are passages which provide definite obstacles to the reading that Socrates is not arguing for anything in particular, such as the one at the end of the *Protagoras* (361b1-361b4): "but now you [Socrates] are arguing the very opposite *and have attempted to show that everything is knowledge—justice, temperance, courage*" (emphasis added). The speaker here is Socrates, and he is giving voice to what he thinks the discussion would say "if it had a voice of its own" (361a5-6). This passage, among others,[10] shows that it is likely that Socrates is arguing for a particular stance; what is not immediately clear is exactly what his stance is.

The second option, then, is to take Socrates' skirmishes with Protagoras as arguments in which Socrates himself does have something to establish. Once we agree to do this, we can examine in more detail what his position that the virtues are in fact "one" implies. This line of thought leads to the development of the Identity View, a very convincing account of which can be found in Terry Penner's "The Unity of Virtue."[11] I think that Penner succeeds there in revealing how the Inseparability View not only faces textual contradictions, but also reduces many of Socrates' argumentative moves to sheer nonsense. In the next section I will briefly discuss how Penner's ac-

count paved the way for the necessity of an interpretation that asserts something stronger than inseparability among the virtues. I will then take issue with and alter some aspects of Penner's Identity View in an attempt to avoid the problems that arise from that view.

II. The Identity View

The Identity View meets the above complications by attributing to Socrates a stronger claim than inseparability. The thought here is that when Socrates makes his claims about Virtue being "one," he is not simply saying that there are many virtues which are linked in such a way that they cannot be separated. He is saying, rather, that there is only Virtue, and that Virtue admits of a number of different names, depending on the circumstances in which it is active. To quote Penner, this is a stronger claim than the one attributed to Socrates in the Inseparability View, "since it carries ontological implications not carried by [inseparability]."[12] The ontological implications Penner identifies are:

> (1) 'Bravery,' 'wisdom,' 'temperance,' 'justice,' and 'piety' are five different names of the same thing, and
>
> (2) In addition to brave men there is such a thing as bravery.[13]

This stronger claim does indeed seem more consistent with the areas in the *Protagoras* mentioned above, where Socrates seems to argue against Protagoras' formulation of Virtue as a whole comprised of distinct parts. Furthermore, when speaking of justice at 330c1-3, Socrates poses the following question: "Is justice a thing or is it not a thing? I think it is. What about you?" There are other instances where he speaks similarly about piety as a "thing" (330d1), and in the whole argument from opposites (332d-333b10), he is speaking of temperance, wisdom, justice and piety as *things*. The ontological implication of (2) sits well with these areas of the dialogue, and so the claim which gives rise to this implication naturally seems more likely to be what Socrates meant.

But Penner also identifies a more fundamental problem with the Inseparability View. As I formulated it above, the Inseparability View understands Socrates' questions as *conceptual* questions about what is carried in our concept of a virtue. Penner argues[14] that what Socrates is after is in fact *not* some sort of logical analysis, but something altogether different—some sort of substantive theory saying *what courage is*—and that the Inseparability View is misguided from the start. The key difference here is that in conceptual questions the reference of "courage" is the *meaning* of courage, while in substantial questions the reference of courage is, to quote Penner, "*that psychological state which explains the fact that certain men do brave acts.*"[15] The reasons given above for preferring a claim which implies (2) are also strong reasons for reading Socrates'

"What is X" questions as substantial rather than conceptual. And there are still further reasons: it is clear as early as 330a in the *Protagoras* that Socrates is speaking of the virtues as having "powers" or "functions" analogous to the powers of the eyes or the ears.[16] In the following discussion on 332a-c, he argues that it is "by temperance" that people "act temperately" (332a11), just as it is "with quickness" that an action is "done quickly" (332b13). A similar discussion can be found in the *Laches* at 192b-c, where Socrates is trying to explain to Laches in what way he wants to discuss courage. Again, the example he gives is a *power*: "[…]what I call swiftness is the power of accomplishing a great deal in a short time" (192b2-3), and he then asks Laches

> …to speak in the same way about courage. What power is it which, because it is the same in pleasure and in pain and in all the other cases in which we were just saying it occurred, is therefore called courage? (192b5-8)

It is also important to note here that the whole discussion with Nicias and Laches is an attempt to find out "the manner in which virtue might be added to the souls of [Lysimachus' and Melesias'] sons to make them better" (190b5-6). It seems clear by this point—without even delving into the contextual confusion that arises from taking Socrates' "What is X" questions as conceptual ones[17]—that what Socrates is after is certainly something more like the identification of a motivational state of soul which gives rise to a specific kind of action. "What is courage," then, can translate roughly into "What is that state of soul, the power of which is to produce courageous actions," and *mutatis mutandis* for justice, temperance, piety, and all the rest. And it seems clear that the Inseparability View doesn't quite address these questions in a satisfactory (or at least in a consistent) way.[18]

If we agree to take Socrates' questions as substantial and not conceptual, then the central issue is this: does the same state of soul gives rise to each of the different virtuous actions, or are there are actually different states of soul, each of which has its own distinct power to produce a certain type of virtuous action? Is courage the state of soul whose power it is to produce courageous actions, justice the state of soul whose power it is to produce just action, and so on? Or is it, rather, one state of soul—Virtue—which has the power to produce courageous actions as well as just, temperate, and pious ones? Penner answers the latter question in the affirmative, and it follows that all the virtue words collapse into names for that single state of soul which is the explanatory entity for all virtuous actions. Penner is content with identifying Virtue with the knowledge of goods and evils[19] and positing that as the single entity to which each virtue-term refers; he does not carry his conception of the unity of Virtue any further. There remains room for distinction, then, only among virtuous actions; the loss of distinction occurs at the deeper level of the "virtues themselves"—that is, in the soul. The difference between a courageous action and a pious one, on this view, is determined entirely by the circumstances in which the action occurs. A virtuous action performed in conditions of danger or affliction, for example, can be

characterized as courageous, while a virtuous action performed in conditions which demand attention to the gods' demands is a pious one, and so forth for all the other virtue-adjectives. The real difference between this view and the Inseparability View is the lack of distinction among the virtues. It is a central feature of the latter view that there is some way to distinguish among the virtues *in themselves* (and not only in the light of external circumstances: see note 18). On Penner's Identity View, there need not be any such distinction at this level. While we are willing to label certain actions as instances of this or that virtue, these actions will all have as their explanation *the same state of soul*.

There is a common-sense objection here that, though worth mentioning, will not by itself refute Penner's strict Identity View. The problem, in brief, is that it just goes against everyday talk about virtues to say that all the virtues are simply names for the same thing. It hardly seems plausible that when we call a person *courageous*, we are making the same claim as when we call her *just*. If this concern is not entirely eliminated by the end of Penner's article, he at least defeats it implicitly by posing Socrates' argument as substantial rather than conceptual. Socrates is simply not talking about virtue in the same way as we do in our everyday conversations. His specific idea of Virtue is not of something that can be separated into distinct parts. Our confused intuition, if Penner's interpretation is right, arises from thinking misguidedly about distinctive virtuous actions as being caused by different states of soul, instead of thinking about them as products of the *same* state of soul occurring under different external circumstances. But there is a more pressing problem here which is at least tangentially related to the previous one, and which is enough to warrant some hesitation toward the strict Identity View outlined above.

III. Problems With the Identity View and a Middle Ground Between Inseparability and Identity

It will be helpful here to assess some of the ontological implications of Penner's Identity View. His interpretation allows us to claim that there exists such a thing as virtue which is by definition identical to both:

(5a) Knowledge of all goods and evils,[20] and

(5b) The state of soul which gives rise to all virtuous actions.

It is also central to the Identity View that

(6) 'Courage,' 'justice,' 'temperance,' and 'piety' are all names for the same thing—namely, the thing whose existence is claimed in (5).

About a virtuous action—I will use courage as an example—we can say

> (7a) A courageous action is produced by Virtue, and
>
> (7b) A courageous action occurs in circumstances of danger or affliction.[21]

For any action G rightly to be called *courageous*, it is necessary both that it

> (a) be produced by Virtue, and
>
> (b) occur under circumstances of danger or affliction.

If (a) is not the case, then G cannot be courageous, just, temperate, or pious; if (b) is not the case, the action might be just, temperate, or pious, but it cannot be courageous. The trouble here is that, given (a), what can be an eligible candidate for a virtuous action is *extremely* limited. We can make this complication clearer if we substitute "knowledge of all goods and evils" for "Virtue" and then rephrase (a). What we get is

> (a′) An action is courageous if and only if it is produced by (or with) knowledge of all goods and evils.

We can see, then, that even to speculate about G being a courageous act is to presuppose either that the speaker herself has knowledge of all goods and evils, or that someone else (namely the person who committed G) has knowledge of all goods and evils. Certainly Socrates does not seem very inclined toward the idea that *anyone* has knowledge of all goods and evils, committed as he is to the idea that acknowledgment of his ignorance is what makes him the wisest man in Athens.[22] The problem here is not *per se* that there are no virtuous individuals, but, rather, that there is no room in the Identity View for a virtuous action to come into being *at all* if no one has knowledge of all goods and evils. This is because Penner's Identity View, by reducing courage to just another name for Virtue, has eliminated the possibility of separating *courage* from *Virtue itself* in order to identify courageous actions on some other grounds which do not presuppose knowledge of all goods and evils. Our labeling an action as *courageous* is in some sense mere happenstance. Besides the fact that it was produced by Virtue, there is nothing special—nothing *outside* or *independent* of Virtue itself—about a particular action which makes it courageous; we just happen to talk that way about actions which *seem* "right" or "good" and which occur in circumstances of danger or affliction. We can of course still speak of "right" actions, but the point is that without Virtue, there is no way of saying that an action is exemplary of Virtue (that it is virtuous). Since there is no virtue in the way Vlastos understood it—as a standard by which to judge whether an action is just *or* pious *or* courageous, etc. [23]—we have noth-

ing outside of Virtue itself by which to make such judgments.

This is especially problematic because Socrates does not hesitate to discuss examples of virtuous actions. Indeed, many of his arguments take for granted that certain actions *are* examples of virtuous behavior.[24] If we accept that one cannot speak meaningfully of individual actions as virtuous (which I take to be a consequence of the Identity View), then many of Socrates' arguments would be begging the question, and much of what he claims would be outright contradictory to this formulation of his doctrine of the unity of Virtue. My suggestion here is to establish some sort of textual basis for distinguishing the virtues themselves at the level of the soul—without construing Socrates' arguments as merely conceptual—in a way that allows us to talk sensibly about virtuous actions. Such an account will, I think, be able to avoid the complications mentioned above, and it will also address our intuitional worry about collapsing all the virtue-terms into one entity. This requires a conception of the virtues such that they remain *quantitatively* undifferentiated—Virtue is still knowledge of all good and evils, and the individual virtue-terms still collectively refer to that state of soul—but also such that the virtues are somehow *qualitatively* distinct—there is some unique quality about each of the virtues by which it has the power to produce a specific type of virtuous action.

My suggestion is to begin by keeping separate our conception of a virtue itself, on the one hand, and a virtue's power or function on the other. Socrates hints at such a distinction at 330b1-2 in the *Protagoras*, where he asks whether the virtues "are unlike each other, both in themselves and in their powers or functions,"[25] and this distinction allows us to begin to question the idea that virtue is just the power to act virtuously. I think that if there is a qualitative distinction to be made at all (and it seems there must be), it should be located somewhere in the area of the distinction hinted at above. In what follows I will elaborate on the thought in 330b1-2 in hope that such an elaboration will lead naturally to a genuine distinction among the individual virtues.

Both the Inseparability View and the Identity View agree that Virtue itself is understood as knowledge of all goods and evils, and I think this account is consistent with the claims Socrates makes at the end of the *Laches* (199b-199e6). There is a problem, however, in speaking of Virtue as just the power to produce virtuous actions. Penner airs some hesitation at the end of his article[26] about how "knowledge, by itself, could be a motive-force in any way," but unfortunately he doesn't proceed to address the issue. I think the key to the problem lies in Penner's own conflation of Virtue and a virtue's power. By reducing Virtue itself simply to the state of soul characterized by knowledge of all goods and evils, he has left out an important aspect necessary to call that knowledge a "state of soul" at all. What I am suggesting is this: we should not think of knowledge *itself* as the *only* thing relevant to Virtue. The fact that Socrates is talking about this knowledge as already relevant somehow to behavior is putting it in a certain context: the context of an agent whose possession of that knowledge is bound up in some way with her actions. We are talking about two separate things when we speak of knowledge in the abstract—as a collection of facts or truths which might exist independently of its possessors—and when we speak of knowledge as something

which "might be added to the souls" of young men "to make them better" (*Laches* 190b5). And if we are going to talk about Virtue as knowledge of all goods and evils, we must get clear about what it means for that knowledge to exist *in the soul* in such a way that it influences action.

Once we realize that there is some necessary and presupposed context here (namely the soul), it becomes natural to make the distinction I urged above about keeping Virtue itself separate from a virtue's power or function. We have to keep these separate because we are talking both about Virtue as a kind of knowledge *and* about the kinds of *actions* that the virtuous *state of soul* produces. Without going into the epistemological questions concerning Socrates' conception of knowledge, there will always already be some distinction here between that knowledge and the power which acts in tandem it. Naively we might say, for example, "Surely X's knowledge won't produce virtuous actions if her muscles have turned into jelly!" This is of course a bit ridiculous, but it shows that there must be more to the picture. There must be, in addition to knowledge, some power which acts together with that knowledge to produce a virtuous action. To dwell at length on the particulars of what that "power" is or what it must be would take me too far astray from the Socratic dialogues. I will say briefly that the distinction I am urging must be effected in some way "in the soul," if we are talking about "states of soul" as the surrounding context for Virtue. We do speak intelligibly of dispositions and tendencies, and it seems dangerous to identify either as something purely and outwardly physical.[27] With this in mind, I think it is plausible not to equate "knowledge of all goods and evils" with "the power to produce virtuous action," which means that we must understand Virtue in a different way.

Here, I am proposing that Socrates holds something like a bi-leveled conception of the soul.[28] Suppose, for example, that an agent is truly virtuous in Socrates' sense. At the first level of this person's soul lies knowledge of all goods and evils. This knowledge can be seen as informing and influencing what I will call the second level of the soul—the level at which I posit the powers to produce certain virtuous actions under certain circumstances. All of these powers or functions are equally influenced by the knowledge which exists at the first level of the soul; in the truly virtuous soul, this is knowledge of all goods and evils. For this reason, all of the *powers* which give rise to specific actions are always *virtuous* in the truly virtuous agent—that is, they will always produce an action which is informed by knowledge of all goods and evils. But they will sometimes be active and sometimes not. In a situation which demands that a person endure affliction or danger, for example, her power to act in an enduring way, *influenced by her knowledge of all goods and evils*, will produce what we might call a courageous action. The "courage" which gives rise to that action, then, is not identical *only* to her knowledge of all goods and evils; it was rather a combination of that knowledge and its influence on the part of her soul which is the power to produce courageous actions. And the state of soul in operation here is not the same as the one that would be in operation if she were to act temperately in a situation whose circumstances demanded that she resist temptation. In that case, her temperate action is a result of her knowledge of all goods and evils and the influence of that knowledge on

her power to act virtuously under conditions which demand self-control. Her temperance, then, is distinct from her courage, and we could make the same case for her justice, her piety, and so on.

One obvious implication of this view is that, so long as a person has knowledge of all goods and evils, there cannot be instances where one virtue conflicts with another. A person's courage, for example, cannot conflict with her justice (as Protagoras suggests at 349d4-8), because both of these virtues are anchored by knowledge of *all* goods and evils. This knowledge, presumably, will inform and influence the right power for acting virtuously under whatever circumstances. It is only in the event that a person does *not* have knowledge of all goods and evils that the "second level" of her soul could possibly produce an action that conflicts with the demands of Virtue. In that case she might act in a way that seems virtuous but is really shortsighted or unconsciously self-advancing, or something else not wholly commendable. Her lack of knowledge, for example, might act in tandem with a power in the second level of her soul and produce a "courageous" action where it was in fact foolish to endure in the given circumstances. But the fact that she has something operating which "recognizes" the circumstances as *the kind in which to endure* (even if that "recognition" is misled because it is not fully informed) shows that the situation was such that a virtuous action might have been possible. What I mean here is this: the "powers" I am positing at the second level of the soul recognize or pick out certain circumstances as the kind in which to act a certain way. It is of course *knowledge* which, for Socrates, will have the final say in the overall process, but that knowledge will have to "fuel" the correct *power* to act in accordance with what is right or wrong, good or evil. And since we all, as agents, have the same "powers" (to act courageously or justly or temperately, etc.) we are able to recognize "virtuous-like" behavior on some grounds independently of Virtue itself. We all have the ability to recognize that such-and-such behavior in such-and-such circumstances seems virtuous *because it was produced by a virtuous power*.

If we are allowed to talk in this way about Virtue, then we can also talk about specific examples of virtuous action. In the Identity View, we were unable to distinguish among individual virtues because all we had to work with was knowledge of all goods and evils, the products of which were impossible to speculate about without presupposing that very knowledge (either in the speaker or in the one who is acting). My distinction between knowledge and power allows for the existence of the individual virtues in a way that is at least partly intuitive, not only to Socrates but to his interlocutors as well. On my account, there is more to Socrates' conception of virtuous action than knowledge of all goods and evils, and this is what we needed to allow for virtuous actions to occur which are not *necessarily* the product of that knowledge. In this conception, it is possible for a "virtuous output" to result from a person who is not virtuous in the sense that she has knowledge of *all* goods and evils. What this would amount to, briefly, is one's power to produce virtuous actions under such-and-such circumstances (the "second level" of her soul) being influenced by what knowledge of goods and evils she *does* have (the "first level" of her soul) in such a way that she performs what is commonly recognized as a "virtuous" action. The reason such

actions can be "commonly recognized" is that they are recognizably produced by the second-level powers of the soul, which are always producing eligible candidates for "virtuous" actions. To be sure, there is much more that would need to be said about this process and about what I have termed a bi-leveled soul. It should be sufficient, however, that I have created room for a separation between our talk of an action being virtuous (e.g. courageous) and a person being virtuous, and have thereby eliminated the necessary condition that a courageous action be produced by knowledge of *all* goods and evils. This conception succeeds in understanding Socrates' questions as substantial and not conceptual (we are still talking about explanatory states of soul), and it also creates the possibility for talking about virtue as something manifested not only in a virtuous agent (who has knowledge of all goods and evils), but in a virtuous action as well (an action recognizably produced by the powers of the second-level of the soul). The problem with the Inseparability View was mainly that it provided a conceptual link among the virtues instead of a substantial one. The substantial link I have provided here is perhaps more akin to inseparability, but it meets the demands of Penner's Identity View without all the puzzling ontological implications. This expands the ontological situation in order to allow for the existence of individual virtuous actions, whether or not there exist persons who are virtuous in the sense that they have knowledge of all goods and evils. Thus, it is possible to understand Socrates and his interlocutors as speaking intelligibly about examples of virtuous actions in a way that is consistent with Socrates' own doctrine of the unity of Virtue.

Received January 2009
Revised March 2009

ENDNOTES

1: All citations to Plato herein come from the versions of the dialogues printed in: John M. Cooper, *Plato: Complete Works*. (Hackett, 1997) The translations are by Rosamond Kent Sprague (*Laches*) and Stanley Lombardo/Karen Bell (*Protagoras*).

2: I am referring here (and will be throughout) to the Socrates as he is represented by Plato in the early Socratic dialogues. In an attempt to get a better understanding of the claims being argued there, I will confine myself in this paper to a discussion of the earlier dialogues (primarily the *Protagoras* and the *Laches*) where the influence of Platonic innovation is not yet very evident. For discussion of how Plato's own views might have developed and departed from those of the actual historical Socrates (or what is known of them), see: Daniel Devereux, "The Unity of the Virtues" *A Companion to Plato* (Grand Rapids, 2006), 325-339 (esp. 336-338), and John M. Cooper, "The Unity of Virtue" in *Reason and Emotion—Essays on Ancient Moral Psychology and Ethical Theory* (New York, 1999), 76-117 (esp. 107-117).

3: Devereux, 325-339.

4: It should be noted here that Devereux advances this view as only one possible way of making sense of the contradictions. In other areas, he does mention that it might in fact be worthwhile to reconcile

the differences in a way that is more judicious to Plato's representation of Socrates, and I am, to some extent, just following him up on that suggestion here.

5: Gregory Vlastos, "The Unity of the Virtues in the Protagoras," in *Platonic Studies* (Princeton: Princeton University 1973), 221-269. Vlastos' Biconditionality Thesis is, for all intents and purposes, terminologically interchangeable with what I am calling the Inseparability View: I have simply adopted the latter term throughout this paper because I will be speaking more widely about the general method of interpretation and not just the thesis Socrates is understood to be asserting.

6: It is not clear in these dialogues whether Socrates himself actually distinguishes among different types of knowledge or wisdom in a way that would allow us to speak in detail about what might make "moral wisdom" different from, say, "theoretical wisdom." The important point here, though, is that for a person to perform *good* or *morally commendable* actions, she must have wisdom of some sort (this is the general thought behind the discussion in the *Protagoras* 349E4-351B), and that wisdom is concerned with action. It is this wisdom—the wisdom which guides an agent to perform morally commendable actions—that I am referring to as *moral wisdom*.

7: Indeed, Vlastos finds the strict Identity View preposterous mainly because he is worried about the "essence" of some individual virtue being the same as the "essence" of a different virtue. He seems especially committed to understanding Socrates' questions as a request for meanings on pp. 227-228, where he expresses concern over whether or not piety, "as a 'standard' by looking to which we can tell whether a given act is or is not pious," could ever be the same thing as Courage. That is, if the virtues are *identical*, we could use Courage as a standard by which to judge whether an action is pious. Just after establishing his Biconditionality Thesis, he claims that he has established a "conceptual" or "definitional connection [...] between virtues" (233). And again in his section on Pauline predications (252-259, esp. 258): "he [Socrates] did not mean to assert that Justice is a just *eidos* and Piety a pious one, but the analytic truth that the *eidos*, Justice, is such that all of its instances are just, and the *eidos*, Piety, is such that all of its instances are pious." These remarks make it clear that the Inseparability View as Vlastos would have it (and as I have outlined it in sec. I) is committed to understanding Socrates' questions as conceptual—that is, as requests for meaning or analytic truths, and not as substantial questions, as Penner understands them (see section II below).

8: It might be objected here that what I am taking to be Socrates' claims are in fact just rhetorical questions meant to further the discussion. There might be something to this objection, but I think there is reason to believe that Socrates does hold at least some of the claims he makes in these dialogues. I elaborate more on the problems of reading Socrates' questions as wholly non-committal below.

9: Thus Vlastos accounts for certain seemingly contradictory parts of the dialogues through an assessment of the goals of the *elenchus* (Vlastos, 268-270). I will not argue here for or against the merits of his interpretation—whatever results such an interpretation might yield, it seems beneficial to develop an account of the doctrine of the unity of virtue that can stand for the most part independent of normative claims about the *elenchus*.

10: There are similar instances in the *Laches* (194b1-4, 198b2-3, 198c10-21, to name a few) where Socrates seems clearly to see himself as having some personal footing in the arguments.

11: Terry Penner, "The Unity of Virtue," *The Philosophical Review* 82 (1973): 35-68.

12: *Ibid.*, 36.

13: *Ibid.*

14: *Ibid.*, 38-42.

15: *Ibid.,* 41.

16: It should be kept in mind, however, that Socrates does distinguish here between the virtues "in themselves" and "their powers and functions" (330b1-2). This will be important in the account of virtue I will develop in the next section.

17: Such an understanding, as Penner points out (91-92), reduces Socrates' argument in 332a3-333b6 (on wisdom and temperance) to a rhetorical move that makes little sense and is blatantly fallacious. He does not mention the skirmish at 330c-331b7 (on piety and justice), but I think the same could be said of that argument as well.

18: Vlastos does talk of "recognizably different moral dispositions" as a way to distinguish among virtues at least some of the time (231). But the whole discussion on dispositions there seems somewhat confused, and he slides in and out of talk about virtuous actions and talk about virtuous agents in a way that makes it hard to understand what he means by moral dispositions. There is, at any rate, a host of problems inherent in thinking of virtue as an "essence" which manifests itself in things we call virtuous, on the one hand, and a moral disposition on the other. Penner's discussion (pp.44-49) is quite helpful in revealing some of the complications of thinking about virtue as a disposition or "tendency," but in the end I think that bringing something akin to Vlastos' dispositions back into the picture is essential to keeping Socrates' doctrine from being self-refuting (see section III below).

19: By page 60 Penner has made it quite clear that he's talking about a "single entity which makes men brave, wise, temperate, just, pious, virtuous, knowledgeable."

20: From the *Laches* 199d-199e6, and from Penner pp. 61-62.

21: I feel obliged to note here, again, that it is not of primary importance what we decide to posit as the circumstances in which a courageous action can be labeled as such. I am following Vlastos here in using "affliction or danger," but the circumstances could just as well be something quite different. The important point is just that there do exist some special circumstances under which it is right to say of an action that it is a courageous action.

22: This view is made evident in many Socratic dialogues, but particularly in the *Apology.*

23: Vlastos, 227.

24: In the *Laches*, Socrates agrees with Laches that a man who fights the enemy while remaining at his post is acting courageously (191a1-5), and he then proceeds to give additional examples of courageous behavior in 191a5-191e3. In the *Protagoras*, the whole argument for the identity of wisdom and courage (349e-350c5) proceeds from the assumption that men who dive into wells and men who fight on horseback are engaging (at least sometimes) in acts that are courageous.

25: It is admittedly not clear whether Socrates is engaged at this point in a mere argumentative (that is, a non-committal) maneuver. I do think, though, that Socrates must be taking for granted some distinction like this in order for his doctrine to be intelligible. This will become clearer when I elaborate on my interpretation of that doctrine.

26: Penner, 67.

27: That is, physical in the sense of "muscles" or "bones" or something along those lines. Even if we take dispositions as being identical to some physical make-up of the brain, we are shifting from the outer and blatantly physical to something inner, something "in the soul" as it were. It seems at any rate unreasonable to speak of neurons or brain construction in an attempt to understand Socrates' conception of the soul.

28: I think it is worth mentioning here that my calling the soul "bi-leveled" is metaphorical; the imagery

invoked by this term should not be taken too literally. I am not suggesting here that Socrates had an elaborate theory of the soul and that part of this theory was something analogous to Plato's idea of the "parts" of the soul. I have simply found the imagery to be helpful in understanding what I take to be a coherent way of making sense of Socrates' claims about the unity of virtue and the soul.

Unjustified Veridical Memory-Belief

WILLIAM J. BRADY
University of North Carolina at Chapel Hill

That memory plays a fundamental role in our knowledge is apparent. Almost all of our knowledge involves learning which took place prior to the present. For example, I might claim to know that Mars has two moons, even though I learned it over a decade ago. Involved in my knowing that Mars has two moons is my belief and my justification—both of which are provided by memory.[1] I know that Mars has two moons in virtue of my remembering the past belief, and I am justified either by some psychological relation to my memory of the belief (e.g. I seem to remember that Mars has two moons) or the accuracy of my memory. While it seems clear that knowledge depends on memory, the relationship between a belief that P and a memory of the belief that P is not obvious. If we are to understand knowledge subsequent to the present, a theory of justification for beliefs provided by memory (henceforth, memory-belief) is essential.

Developments of such a theory has been attempted in the literature recently and many have failed on intuitive grounds.[2] In "The Problem of Memory Knowledge" (1999), Michael Huemer develops a justification theory called the dualistic theory which he supposes accounts for the intuitions that past theories have not. I begin the paper by examining how the dualistic theory succeeds in accounting for certain intuitions about the justification of memory-belief. I then propose three cases, all of which fall under a sort of case I call unjustified veridical memory-belief. I argue that these cases serve as counter-examples to the dualistic theory, and then develop motivation for such cases by considering results from psychological research on false memory.

Before we begin there is a distinction to be made among different sorts of memory—the most notable for our purposes is between event memory and propositional memory. Event memory involves a recollection of personal experience, e.g. I have the memory *of* being at the party last night. Propositional memory involves recollection of factual information, e.g. I have the memory *that* Mars has two moons. Both Huemer (1999) and Senor (2005) focus on propositional memory in the analysis of memory-belief, and so shall we in the present paper. Therefore, we can define a mem-

 Arché, Volume 3, Issue 1: Spring 2009. pp. 31-41

ory-belief as a proposition of the sort, "I remember that P" which refers to a recollection of some proposition previously believed (which is currently believed as a result of the recollection).

I. The Dualistic Theory

The dualistic theory holds that a memory-belief is justified if and only if

(a) one was justified in adopting the belief that P, and

(b) one was justified in retaining the belief that P.

Huemer assumes an internalist deontological view of justification for both conditions, in this case to be understood in terms of "epistemic responsibility."[3] On this view he posits that the condition of justified adoption will be satisfied as long as a person forms her belief in an epistemically responsible manner; namely, by acting as rationally as she possibly can in forming her belief. In contrast, if one forms a belief that P for example merely due to wishful thinking, then her belief that P will not be justified since she is not acting in a rational manner.

Similarly, the justified retention condition can be satisfied as long as one's memory is retained in an epistemically responsible manner. Huemer explains this to mean that, "the normal functioning of memory, in the absence of specific reasons for revising a belief, constitutes an epistemically responsible manner of retaining beliefs."[4]

What Huemer means by the "normal functioning of memory" is made clear in his explanation of the dualistic theory's advantages. According to Huemer, the dualistic theory captures the "intuition that I am rational in believing something I seem to remember even if on this particular occasion, unbeknownst to me, my memory is deceiving me—even if, that is to say, I never really had that belief before."[5] Thus the justified retention condition is satisfied by the normal functioning of memory, where "normal functioning" will include *seeming to remember* even though one's memory is actually deceiving them—just so long as they are unaware that their memory is erring.

Huemer claims that the dualistic theory accounts for two intuitive conclusions about justification for memory-belief that two opponent theories, the preservation theory and the foundational theory, cannot. The preservation theory holds that the justificatory status of a memory-belief is transferred from the original formation of the belief. For example, if I am justified in forming the belief that P, my memory-belief is automatically justified upon remembering that P. The preservation theory, however, fails to account for the intuition that the justification of memory-belief should depend on the internal state of the believer.[6] Huemer considers a rendition of Russell's five minute hypothesis to argue that not only do we hold the latter intuition, but that the dualistic theory accounts for it while the preservation theory fails to do so.

Suppose that, five minutes ago, an evil deceiver created someone who had the exact memories and was in the exact situation that Jones was in five minutes ago. This person, henceforth Jones2, is identical with Jones except for the fact that Jones2's memories are false in that he never *actually* experienced them (since he was created five minutes ago). It would therefore make sense for Jones2 to believe the same things that Jones believes, such as eating pizza for breakfast. Indeed both Jones and Jones2 have a memory of picking up pizza and eating it in the same way. The fact of the matter, however, is that Jones actually ate pizza while Jones2 did not.

Huemer thinks that in the above case Jones2 is justified in his belief that he ate pizza for breakfast. In terms of phenomenological experience, Jones and Jones2 are completely indistinguishable. Jones2 has no reason to suspect that his memories are false, so he will be justified based on the internalist "epistemic responsibility" view. Surely Jones2's belief that he ate pizza for breakfast is perfectly rational from his perspective. But the preservation theory, which supposes that a memory-belief retains the justificatory status of the original belief, is committed to concluding that Jones2 has no justification at all for his belief. Since Jones2 actually did not eat pizza for breakfast, according to the preservation theory there is no initially justified belief to transfer to the memory.

The intuition that Jones2 is justified is not present for one who holds a reliabilist account of justification (and perhaps other versions of externalism). According to the reliabilist, the fact that Jones2 has a faulty memory bars his belief that he ate pizza from being justified.[7] However, for our present purposes we will continue to assume an internalistic deontological view of justification.

While the preservation theory fails in the case of Jones2, the dualistic theory can account for the internalist intuition. Jones2 satisfies the dualistic theory's condition of justified retention because he seems to remember eating pizza and has no defeaters he is aware of (e.g. Jones2 is aware that his memory usually leads him to hold false beliefs about his past). Since the memory of his eating pizza also happens to be the source from which he formed his belief, he meets the conditions for justified adoption as well. Since both criteria are satisfied, according to the dualistic theory, Jones2 is justified in his memory-belief as we have concluded on intuitive grounds.

The foundational theory, on the other hand, holds that seeming to remember that P is self-justified, and therefore a memory-belief that P is prima facie justified merely in virtue of seeming to remember that P.[8] But a consequence of this theory is that a previously unjustified belief could gain justification merely by passing into memory, and this is in opposition to a second intuition that the justification of a belief cannot be increased by passing into memory; rather, justification of a belief should only be lowered as it passes into memory.

Considering an example elucidates this point: while traveling in Europe, Mary read in the *National Enquirer*, a supermarket tabloid news source, that aliens were seen strolling suspiciously around Williamsburg, Brooklyn. Mary reads the *National Enquirer* regularly and trusts it because she wishes its stories were real (she thinks life would be boring otherwise), and thus forms the belief that aliens were seen walking in

Williamsburg. It happens to be that Mary lives in Williamsburg herself, and upon returning from Europe three weeks later she remembers that aliens were spotted in her neighborhood. Mary has since forgotten from where she read about the aliens, but nonetheless she still remembers reading and believes that they were spotted. Mary decides to take preventative measures. She nails her all windows shut with wood.

Two points seem clear in the case of Mary: first, when Mary originally forms her belief that aliens were spotted in Williamsburg, she is unjustified (she formed the belief by reading a non-credible source and wishful thinking). Moreover, when she is back at home and has the memory-belief that aliens were spotted she is still unjustified. But the foundational theory would consider Mary justified, since, in virtue of seeming to remember that aliens were spotted in Brooklyn her memory-belief is justified. This cannot be satisfactory however, because it allows Mary's previously unjustified belief to become justified merely by passing into memory. Where the foundational theory fails in the case of Mary, the dualistic theory accounts for our intuitions: since Mary originally formed her belief in an unjustified manner, she did not meet the justified adoption condition and therefore the dualistic theory concludes that Mary is unjustified in her memory-belief.

Thus, according to Huemer, the dualistic theory is at least *prima facie* motivated as it accounts for intuitions that other justification theories such as the foundational and preservation theories allegedly cannot.

II. Cases of Unjustified Veridical Memory-Beliefs

There is, however, a sort of case in which the dualistic theory fails to account for our intuitions: those in which one's memory stores the original veridical belief that P, but memory distorts the details of the belief, leading us to call the belief-holder unjustified. For example, one may recall that P, where P is a true belief she held at one point, but her memory may confuse the source of her belief. We now turn to three cases in which the latter sort of situation leads to our calling the belief-holder unjustified.

Three plausible types of memory distortion come to mind:

> (1) those of source misattribution, in which one forgets or misremembers the source of her belief;
>
> (2) those of inflation, in which one falsely remembers that an event has occurred; and
>
> (3) those of misinformation, in which one receives false information that biases her memory of the original information surrounding the belief.

We begin by considering the case of Jean which involves source misattribution.

Jean is a normally functioning individual. She is a regular reader of the tabloid news source the *National Enquirer* because by wishful thinking she thinks it is credible (she wishes for example that "Batboy" is real). One Monday, however, Jean happens to be reading *The NewYork Times* and sees the headline that the stock market has hit an all-time low. Jean forms the belief that the stock market has hit an all-time low on Monday. A week passes and Jean remembers that the stock market was at an all-time low last Monday, but has misremembered the source of her information—she now believes that last Monday she read it in the *National Enquirer*. However, as a regular reader of the *National Enquirer* she nonetheless maintains her belief.

Surely Jean was justified in holding her belief when she originally formed it since she read it from a credible source. But when the belief passes into memory, and unbeknownst to her she misattributes the source, her memory-belief becomes unjustified. She now believes that she learned the information from a non-credible source, one which she personally thinks is credible only due to wishful thinking. Because she believes that P in the latter irrational way, she is barred from being justified especially according to an internalist deontological view.

However, the dualistic theory is committed to concluding that Jean is justified in holding her memory-belief. Since Jean was justified in forming her belief that the stock market hit an all-time low, she satisfies the justified adoption condition. Moreover, since she is unaware of the fact that her memory is faulty, she satisfies the justified retention condition since she has no specific reason to doubt her memory. But if Jean satisfies both conditions, then according to the dualistic theory she is justified, while we have concluded that Jean should *not* be justified. The dualistic theory therefore fails to account for the case of Jean, which involves a veridical memory of a belief, but also involves source misattribution that renders her unjustified.

One may object that upon remembering that the stock market hit an all-time low, Jean is forming a new belief. In other words, when Jean believes the stock market hit an all-time low via *The NewYork Times*, she has a different belief than when she believes the stock market hit an all-time low via the *National Enquirer*. After all, subsequent to recall, if asked where she learned her information about the stock market, Jean would report that she read it from the *National Enquirer* (while she originally would have reported reading it from *The NewYork Times*). If a new belief is actually formed, then this new belief would be considered unjustified by the dualistic theory since the adoption condition is violated: she will have adopted the belief from a non-credible source (the National Enquirer). Therefore, the dualistic theory does not conflict with our intuitions.

While the objection above proposes a correct analysis of a newly formed belief, it is not true that Jean forms a new belief upon remembering that P. It is important to distinguish between a change in details surrounding a belief that P, and a change in the belief that P itself. Jean might believe that the stock market hit an all-time low and that she read it from *The New York Times*. However, if the latter is viewed as one belief, it is a separate belief from the belief in question in the case of Jean. We are in-

terested in her belief that the stock market hit an all-time low, and while details may have changed surrounding that belief upon her remembering (namely, the source of her information), the belief in question remains that the stock market hit an all-time low when she remembers it. The fact that she misremembers the source of her information leads us to call her unjustified, but the dualistic theory is insensitive to this change because it focuses on the initial adoption of the belief, and then shifts sole focus to the retention. Since Jean passes both the adoption and retention condition, the dualistic theory will call Jean justified, while we have concluded the opposite.

A second type of memory distortion is that of inflation, in which one remembers facts about something without actually experiencing it (e.g. due to merely imagining an event). Consider the case of Chase: as a young child, Chase was abused sexually by his uncle. Being confused and hurt by the abuse, Chase chose not to tell anyone and repressed all memories of the abuse. Twenty years later, Chase began to feel high stress and crippling depression. He went to a therapy session in which his psychologist, Dr. Sanborn, asked Chase if there could be any past events underlying his current depression. Chase reported that there were no such events, but then Dr. Sanborn asked Chase to engage in mental imagery so he might uncover hidden memories. He asked Chase to imagine someone close to him touching him inappropriately and to fill in the details as vividly as he liked. Chase began to imagine a scene of child abuse but reported no uncovered memory. The next morning, however, when Chase woke up he suddenly remembered a frightening feeling upon thinking of his childhood. He immediately phoned Dr. Sanborn. When Dr. Sanborn answered Chase exclaimed, "Oh my goodness! I was sexually abused by my uncle as a child!"

The case of Chase is in a sense parallel to the case of Jones and Jones2: Chase, like Jones2, forms his belief that P as a result of remembering that P. It is true that Chase was sexually abused as a child, but as a result of repression he does not believe that he was sexually abused until the morning after his therapy session.[9] As a result of the mental imagery exercise in which he vividly imagined abuse occurring, Chase suddenly remembered that he was abused and subsequently formed the belief that he was abused. Chase's memory has conflated his imagination with a true occurrence, although as it turns out the abuse that Chase remembers as a result of imagination did actually occur to him.

In the case of Jones2, he is thrown into the world by an evil deceiver and only believes that P as a result of remembering that P. As we have seen above, Huemer argues that the dualistic theory concludes that one in the situation of Jones2 is justified: "Since [Jones2] acquired his belief that he ate a bagel this morning by seeming to remember it, he is rational in accepting it."[10]

Since Chase similarly forms his belief that he was abused as a result of remembering that he was abused (after the mental imagery), the dualistic theory will conclude that Chase is justified in his memory-belief that he was abused as well. Chase's belief that he was abused satisfies the justified retention condition because like the case of Jones2 he has no "specific defeaters" he is aware of in his remembering that P. Also like Jones2, he satisfies the justified adoption condition because the dualistic theory

allows that seeming to remember P is a rational way of adopting a belief.[11] The dualistic theory therefore concludes that Chase is justified in his memory-belief that he was abused.

In the case of Chase, however, it seems that our intuitions call Chase unjustified in his memory-belief that he was sexually abused. While it is true that he was abused as a child, he only comes to remember that he was abused due to his imagination of an abusive event. If we removed the fact that Chase was actually abused from the case, then surely we would conclude that he is unjustified in his memory-belief that he was abused—he would only believe that he was abused as a result of the imagination of the event. As the case goes, Chase is in a sense in the latter situation: while the abuse actually did occur, his repression prevents him from remembering that he was abused *because* he was abused. He only remembers he was abused because he imagined it, and he is completely aware of his voluntary engagement in imagining the abusive event. Therefore, on the present internalist account of justification, we should conclude that Chase is unjustified in his memory-belief that he was sexually abused. The dualistic theory, however, yields the opposite conclusion.

Admittedly, the analysis of the case of Chase is contingent upon whether or not one thinks Chase is acting in an epistemically responsible manner when he comes to believe he was abused as a result of remembering. According to our analysis, Chase is called unjustified in his retention because he voluntary engaged in imagining the abusive event, and this lead to a false memory. Huemer may object, however, that on the internalist picture we should think that Chase is acting in a perfectly responsible manner, since he is completely unaware of his memory erring due to inflation. Therefore, our intuitions should coincide with the dualistic theory's conclusion that Chase is justified in his memory-belief.

I reply that the case is at best borderline in terms of epistemic responsibility, and in the least Huemer has a case where the dualistic theory does not obviously account for our intuitions. Perhaps details of the case could be altered to make it certain that we call Chase unjustified, thus securing our original conclusion about the case. For example, we can alter the case so that Chase has a history of visiting psychologists with the hopes that imagery therapy will help him uncover repressed memories that he wishes he had (this way he could easily explain his depression).

A third case of unjustified veridical memory-belief that poses a problem for the dualistic theory is one which involves misinformation. Cases of misinformation involve an original memory being biased towards false information learned prior to forming the memory. Consider the case of Jeff. Jeff and his girlfriend Sarah are having problems—they don't see each other very often these days, and when they do see each other it usually results in fighting. One day while walking on campus, Jeff saw Sarah with another man, but he couldn't tell who the man is. Jeff became angry because the two looked like they were flirting, but then suddenly he saw them start to kiss each other. Jeff ran home upset, forming the belief that Sarah cheated on him with someone. The next day Jeff confided in his friend who is a huge gossiper on campus. Jeff asked the gossiper if he knew anything about Sarah and this new man. The gossiper

said that he knew who the man was, and said that Sarah and the man had been dating for months already. Jeff knew that the gossiper was a non-credible source—he often made up elaborate stories about people on campus because he liked to spread rumors. However, Jeff was so upset that he just had to know who Sarah was cheating on him with, and he chose to believe the gossiper. As it turns out, everything the gossiper said was a lie. However, later that day Jeff called Sarah and told her that he couldn't see her for at least a week. After a week, Jeff called Sarah ready to lecture her about everything she had been doing covertly over the past two months. He remembered that Sarah cheated on him with somebody, now recalling it from the gossiper source with the false details instead of his original perceptive source of information.

The case of Jeff is one of misinformation—Jeff forms his belief that P by witnessing an event, but then later is told incorrect information about the event which biases the details of his memory towards the false source. When Jeff forms the belief that Sarah cheated on him with someone as a result of his perception of the event, he is justified in his belief—on our present account of justification he passes the criteria of epistemic rationality. However, when he forms the memory-belief that Sarah cheated on him he is unjustified. Even though Jeff's belief that Sarah is cheating on him with someone is true, Jeff comes to remember the event with details from the gossiper source that contains false information. Moreover, Jeff only believes the gossiper due to his wishful desire to know who Sarah is cheating on him with. Therefore, Jeff is acting irrationally in epistemic terms and his belief is unjustified.

The dualistic theory, however, once again draws a contradictory conclusion compared to our intuitions on the case. According to the dualistic theory, since Jeff acquires his belief that P in a justified manner, he satisfies the criteria of justified adoption. Moreover, Jeff's trust of his memory satisfies the justified retention condition since Jeff is unaware of the fact that his original perceptual memory has been conflated with the gossiper story. Perhaps Jeff is aware that the gossiper is a non-credible source, but he has no reason to think that his memory should have conflated the gossiper's story with his perception of the cheating. Therefore, Jeff satisfies both conditions and the dualistic theory must conclude that he is justified in his memory-belief that Sarah cheated on him. On the other hand we have concluded that Jeff should be considered unjustified, and thus the dualistic theory fails to account for the misinformation case.

One may object that Huemer will actually call Jeff unjustified in his adoption, because it seems that he adopts the belief that someone cheated on him as a result of listening to the gossiper. Therefore, the dualistic theory will call Jeff unjustified by adopting the belief in an irresponsible manner (by using a non-credible source). However, this objection only succeeds if Jeff is forming his belief that someone cheated on him by way of the gossiper. In the case, Jeff adopts the belief rather by perceiving the event—talking to the gossiper skews his original belief but does not *form* it (this is the misinformation aspect of the case). The fact that his information is skewed leads us to call his memory-belief unjustified, but the dualistic theory is only worried about Jeff's retention of the belief from his perspective. Since Jeff is unaware that his memory has been biased, the dualistic theory will consider him justified in his retention. Thus,

the dualistic theory calls Jeff justified while we have concluded that he is unjustified in his memory-belief.

III. Empirical Support for Cases of Unjustified Veridical Memory-Belief

The dualistic theory fails to account for three sorts of cases dealing with unjustified veridical memory-belief. In the cases, while the memory-belief that P is veridical (the proposition is true), details of the belief are distorted in the memory which leads to our calling the person unjustified in her memory-belief. For example, in the case of Jean, while Jean's memory that the stock market hit an all-time low is true, her false memory that she read it from a non-credible source bars her from being justified.

It may be objected at this point that the presented cases of unjustified veridical memory-belief involve memory distortions which are rare exceptions. These odd cases, one may object, need not be accounted for so long as a theory works in usual circumstances. However, empirical evidence from psychological research on memory suggests that the sorts of memory distortions involved in the three cases seem to be highly prevalent in normal populations. If the latter is true, then the cases of unjustified veridical memory-belief serve as solid counter-examples to the dualistic theory rather than exceptions that can be glossed over.

The case of Jean is an example of a common memory error known in the psychological literature as source misattribution—the tendency of subjects to confuse or forget the source of their memory. Jean misremembers the source of her information that the stock market hit an all-time low as from the *National Enquirer* instead of *The New York Times* where she actually read it. Source misattribution has been shown to be a robust phenomenon through numerous studies.[12] Moreover, subjects may have a correct memory of information but nonetheless confuse where the information came from.[13] Source misattribution has also been shown to generalize to real world situations such as eyewitness cases[14] and memory of childhood events.[15]

The second case dealt with Chase, in which his imagining of an event caused him to falsely remember information from the event and that the event actually occurred. This sort of phenomena, known in the psychological literature as imagination inflation, has also been shown across multiple studies to be a robust effect. One study found that just by having subjects imagine a childhood event (e.g., getting a hand cut by glass) increased their confidence ratings of the event occurring whether or not it actually did.[16] Studies have extended research to find that subjects can be convinced they have recently performed an action (e.g. flipped a coin) merely by imagining they have done so.[17]

The last case presented was the case of Jeff which involved misinformation biasing the details present in his memory. This effect is brought out in experiments at a high rate through the post-event misinformation paradigm.[18] The paradigm is known for finding that subjects who read or hear misleading reports often answer biased towards false information they are presented with (as compared to the control groups).

Some studies suggest that the misinformation effect decreases when subjects hear a report from something known to be non-credible.[19] However, if a long enough delay occurs between actually witnessing the event and reporting what occurred, subjects will fall into the misinformation trap even if they previously knew the source was non-credible.[20]

IV. Conclusions

The psychological research on false memory provides an array of evidence that memory distortions can occur in normally functioning individuals. The precise conditions in which they occur may still be an open question, but we do learn something clear about the relationship between a belief and the memory of that belief: namely that there is no necessary connection between a belief being stored and its being likely to be true upon recall.[21] Since the psychological evidence generally supports the latter claim, then it seems that cases of unjustified veridical memory-belief that result from memory distortion cannot be ignored.

Future psychological evidence may help elucidate conditions in which memory distortions occur, but presently I have argued that there is a solid set of counter-examples that can be raised against the dualistic theory, and these are the cases in which one's memory holds a veridical belief, but false details of that belief render the memory-belief holder unjustified. Perhaps the dualistic theory can be amended to account for these cases of unjustified veridical memory-belief, but how such an amendment can be made remains to be seen. By developing a sort of case that gives the dualistic theory problems, this paper has hopefully given future theories another set of criteria to account for when seeking to be as valid as possible. The aim of this paper has been at least to extend the dialectic in regards to epistemological problems of memory towards necessary conditions for the justification of memory-belief.[22]

Received January 2009
Revised March 2009

ENDNOTES

1: Assuming that Mars has two moons is a non-basic belief, and that justification is required for knowledge—some externalists like Dretske have argued that justification is unnecessary for knowledge e.g., Fred Dretske, *Knowledge and the Flow of Information*, (Cambridge: The MIT Press, 1981).

2: For a more complete summary of proposed theories and objections than presented here, see: Michael Huemer, "The Problem of Memory Knowledge," *Pacific Philosophical Quarterly* 80:4 (1999): 346-357; Thomas Senor, "Epistemological Problems of Memory." *Stanford Encyclopedia of Philosophy*, 2005.

3: Huemer, 355.

4: *Ibid.*, 351.

5: *Ibid.*, 357.

6: *Ibid.*, 350.

7: Alvin Goldman, *Epistemology and Cognition*, Cambridge: Harvard University Press, 1986.

8: Huemer, 348.

9: Let us assume here that Chase does not hold the belief that he was abused non-occurrently prior to the therapy session

10: Huemer, 351.

11: *Ibid.*, 351.

12: See, for example: Larry L. Jacoby et al, "Becoming famous overnight: limits on the ability to avoid unconscious influences of the past." *Journal of Personality and Social Psychology* 56:3 (1989): 326-338; Johnson, Hashtroudi, and Lindsay, "Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions." *Journal of Memory and Language* 32:1 (1993): 1-24; Robert Belli and Elizabeth Loftus, "Recovered memories of childhood abuse: a source monitoring perspective." *Dissociation: Clinical and theoretical perspectives*. New York: Guilford Press, 1994; Schacter, Harbluck, & McLachlan, "Retrieval without recollection: An experimental analysis of source amnesia." *Journal of Verbal Learning and Verbal Behavior* 23:5 (1984): 593-611.

13: Schacter et. al.

14: Kleider, Pezdek, Goldfinger and Kirk, "Schema-driven source misattribution errors: Remembering the expected from a witnessed event," *Applied Cognitive Psychology* 22:1 (2007): 1-20.

15: Ceci, Loftus, Leichtman, and Bruck, "The possible role of source misattribution in the creation of false beliefs among preschoolers," *International Journal of Clinical and Experimental Hypnosis* 42:4 (1994): 304-320.

16: Garry, Manning, Loftus and Sherman "Imagination inflation: imagining a childhood event inflates confidence that it occurred." *Psychonomic Bulletin & Review* 3:2 (1996): 208-214.

17: Goff & Roediger, "Imagination inflation for action events: Repeated imaginings lead to illusory recollections," *Memory and Cognition* 26:1 (1998): 20-33; Thomas & Loftus, "Creating bizarre false memories through imagination," *Memory and Cognition* 30:3 (2002): 423-421.

18: Loftus, Miller and Burns, "Semantic integration of verbal information into a visual memory," *Journal of Experimental Psychology: Human Learning and Memory* 4:1 (1978): 19-31.

19: Smith and Ellsworth, "The social psychology of eyewitness accuracy: Misleading questions and communicator expertise," *Journal of Applied Psychology* 72:2 (1987): 294-300.

20: Underwood and Pezdek, "Memory suggestibility as an example of the sleeper effect," *Psychonomic Bulletin and Review* 5:3 (1998): 449-453.

21: This point was argued by Senor (2005) independent of empirical evidence in order to critique internalistic justification theories.

22: Special thanks to Ram Neta and Jesse Prinz for testing their intuitions on the three cases of unjustified veridical memory-belief.

The Symposium: An Acoustic Illusion

PETER MOORE
Boston University

The relationship between *eros* (love) and beauty dominates a significant portion of the dialogue in Plato's *Symposium*, appearing most explicitly in the disagreement between Socrates and Agathon over whether *eros* itself is beautiful or whether it is *of* a beautiful object. Moreover, Socrates concludes his speech with a statement about the role this relationship plays in the life of the philosopher: he is enamored with the essence of beauty. Insofar as we are philosophically concerned, therefore, with the teaching that the dialogue offers about *eros*, we would do well to ask ourselves what makes it beautiful. For it is a beautiful dialogue. At times one is even overcome with madness and jealousy for this beauty; for the account that Apollodorus gives us, by which we are twice removed from the original symposium at which the dialogue purportedly takes place, reminds us that these words were not meant for us, but for the ears of another.[1] As readers we stand in an erotic relation with the dialogue: we expect that the dialogue holds answers for us, and we must find them.

So we must concern ourselves with the beauty of the dialogue, insofar as we perceive this beauty and wish to follow Socrates to its form, the supreme object in the philosopher's development. Thus it would be natural for us to ask, "What form does the dialogue take, and what does this "form" teach us about *eros*?"[2] One answer to this question takes its bearings from the visual metaphor that *eidos*[3] form suggests, positing geometrical representations of the dialogue's form. Accordingly, one figure typically used is a step-pyramid, and[4] another possible figure is the circle. While I will ultimately reject these representations, I will nonetheless present their outlines in order to identify features of the dialogue that are indispensable to understanding Socrates' teaching of *eros*.

A pyramid suggests a finite ascent; likewise, the speeches proceed in an apparently dialectical manner towards a higher and more refined *logos*[5] of *eros*. According to the pyramidal representation, each speech ascends by overcoming to the contradictions of its predecessor. At the base, Phaedrus praises *eros* for its usefulness: *eros* provides the motivation necessary for the accomplishment of great and virtuous deeds. The

 Arché, Volume 3, Issue 1: Spring 2009. pp. 43-51

lover wants to appear virtuous in the eyes of the beloved, while dreading the shame of appearing vicious (178d). Here, however, Phaedrus neglects a crucial distinction: the lover acts out of vanity, since the desire to appear virtuous motivates him to act virtuously. But since the appearance that the lover project depends upon the beloved's conception of the good, the lover might actually act wrongly, in an effort to impress his beloved, such as one in love with a thief might act when robbing a store. In other words, the lover wishes to appear virtuous, but whether he appears so depends upon the one to whom he appears. Thus ethical norms are wholly absent from Phaedrus' account.

Precisely this ethical component becomes Pausanias' addition. *Eros* is not one, but of two kinds (180c): a "heavenly," noble, and good kind, and a "popular," crude, and bad kind. According to Pausanias, the factor that determines the ethical character of *eros* is described by the following rule: "for every action it may be observed that as acted by itself it is neither noble nor base" (181a). But as this ethical principle threatens to devolve into relativism[6], Pausanias must later abandon it, favoring instead the object of *eros* for its ethical determinant: love of the *soul* characterizes the good and heavenly love, while love of the *body* characterizes the bad and popular love (183e). Eryximachos, perhaps perceiving the relativistic tendency in the ethical principle that Pausanias introduced, attempts to give *eros* a non-relativistic basis in his natural-scientific worldview: the heavenly *eros* and the popular *eros* comprise two opposing, natural forces that, moreover, can be known (objectively) and controlled.[7] In this way he introduces the possibility of gaining knowledge of *eros*, rather than simply offering eulogies. Medicine under this conception of *eros* is just knowledge of the methods by which the practitioner may balance and harmonize these two forces.

However, in extending the domain of *eros* to encompass all natural phenomena, Eryximachos diminishes the distinctly human aspect of love, namely, that it consists in a relationship between two persons. Hence Aristophanes begins his speech by limiting the domain of *eros*, recommending that the other speakers begin with the "nature of man and its development" (189d). Following his own recommendation, Aristophanes relates a *mythos* (myth) and *istoria* (history) of humankind that purports to explain why love is as we know it currently: each person longs for his other half, and love is just this desire to restore one's wholeness (193a). In this way Aristophanes includes Eryximachos' contribution (knowledge) to the discussion, since he offers a definition of *eros*, but also refines it by limiting it to human relationships. Yet Aristophanes' *mythos* seems to endorse a bleak view of *eros* and the human condition. *Eros* was contrived by the gods as a punishment for humanity's predecessors after they stormed the heavens (190c-d). Hence in contrast with Aristophanes, Agathon praises love for its *beauty*: love itself is beautiful, and as such only the talents of a poet can properly reveal its nature.[8]

According to the pyramidal representation of the dialogue, Socrates' speech would correspond to the peak in virtue of its references to every preceding speech: *eros* is useful because it provides the basic, vital energy for all of our pursuits (Phaedrus);[9] at the proper stage in his development, the lover will feel compelled to prefer the

soul to the body, and undertake the development of another soul (Pausanias);[10] at a still higher stage in his development, the lover will behold the beauty in "different branches of knowledge" (Eryximachos);[11] the telling of a story about the birth of *eros* from *Poros* and *Penia* suggests that *Mythos* is an appropriate means for relating the nature of *eros* (Aristophanes);[12] *eros* itself is not beautiful, but its object is (Agathon).[13]

Moreover, the ascent passage (210a—211d) seems to function as a miniature of the form of the whole dialogue. Just as each speech ascended to a higher stratum of the pyramid by overcoming the contradictions of the previous, so too does the lover continually push the object of *eros* to a higher plane by recognizing the temporal imperfection in the original object. The beauty of one particular body will fade, and so upon recognizing the universality of beauty in beautiful bodies, the lover will seek beauty at a higher temporal plane, namely, the soul (210b). But, as the soul only endures one lifetime, the lover will seek beauty in the *nomoi*, or laws and institutions of the *polis*, which endures beyond the life of any particular individual (210c). Knowledge likewise represents an advance to a higher temporal plane, since knowledge at best is true at all times, that is, eternally true; but, being forgetful creatures, we must constantly preserve knowledge in memory (208a). Thus at the peak of this ascent, the lover seeks an eternal object—the essence of beauty—which never perishes or changes (211a-b). So regarded, Socrates' speech is like the golden capstone of the pyramid: it requires the lower strata for its support, but, being enthroned in gold, we behold it as more precious and wondrous than the rest.

Nonetheless, there are difficulties in the dialogue which the pyramid schema cannot overcome. For one, the speeches do not proceed as continuously as the pyramid would have them. For example, Aristophanes incorporates almost nothing explicit from the previous speeches into his own, and in fact prefaces his own by dismissing of the seriousness of Eryximachos' speech, saying that he "unsays all that [Eryximachos] has said" (189b). In other words, the development of the speeches may only be apparent; they are primarily the contributions of individual personalities.[14] The second difficulty is Alcibiades' speech, which criticizes, and in many ways undermines, Socrates' speech. Nor can we dismiss the difficulty in Alcibiades' speech as a dramatic flourish meant to show Plato's fondness for his teacher, at least not if we intend to contemplate the teaching of *The Symposium* seriously. It is worth noting that, besides Socrates,[15] Alcibiades is the only other character in the dialogue who professes to speak truly (213a). This profession puts him in competition with Socrates for the truth—a competition in which Socrates later partakes when he agrees (upon Alcibiades' suggestion), to interrupt wherever Alcibiades speaks falsely about him.[16] More significantly, Socrates never once interrupts. Given these details, we must regard Alcibiades' words as true, and consider them in relation to whatever teaching *The Symposium* contains.

Alcibiades begins his praise of Socrates with "*eikones*," or "representations." Socrates is *hubristes* (215b); he is like the satyr Marsyas (215b), who, also in a hubristic manner, challenged Apollo's supremacy in the musical arts and incurred due punishment for it.[17] So there is a kind of arrogance in Socrates' pursuit of the universal: like Marsyas,

Socrates challenges the divine, perhaps even seeking to become divine himself. Later, however, Alcibiades abandons these *eikones*, claiming that Socrates is "unlike any other person, ancient or modern" (221c). Moreover, Socrates gives no value to beauty, wealth, honor—or even his fellow men.[18] In Alcibiades' praise, therefore, Socrates is a radical particular—he has no comparison. But precisely this peculiarity undermines Socrates' ascent. In his attempt to know universals, to abandon the procession of time for the eternal, Socrates forgets that the object of *eros* is not universals, but particular individuals. Thus Alcibiades presents us with the paradox of Socrates' personality for which the pyramid cannot account. Socrates seeks to live his life entirely as an expression of the universal, but in doing so, he makes himself a radical particular.

Even if the pyramid fails as a general representation of the dialogue, it nonetheless captures one crucial detail about the teaching of the dialogue: the philosopher thinks of dialogue as a process of dialectical ascent. He wants to stand atop the pyramid with the universal, and from those heights survey all of the various kinds of individual *beings*: bodies, souls, the *polis* (city), the different branches of knowledge, etc. *Eros* is the energy that propels the philosopher upward in his ascent. Hence whatever representation we adopt, it must allow us to pursue the objects of *eros* to their heights.

However, another geometrical figure is available that might account for Alcibiades' speech. We might represent the form of the dialogue as a circle with *eros* at the center. The details of the dialogue in fact suggest this figure, in that the speakers have arranged themselves in a circle. This representation, moreover, would seem to account for a conspicuous feature common to every speech: each speaker praises *eros* for precisely those attributes that each identifies in himself. Eryximachos posits *eros* as a natural-scientific force because he himself is a doctor and natural-scientist; Aristophanes relates the nature of love through *mythos* because he himself is a teller of myths; Agathon assigns the task of revealing the nature of *eros* to the poet because he himself is a poet; Socrates places *eros* between ignorance and wisdom, and describes him as "shoeless and homeless" (203d1) because he himself is a philosopher. Thus construed, we as readers of the dialogue continually view *eros* from a new perspective as we move around it; but, precisely because we cannot advance towards the center, we feel like a man who by some misfortune could not point directly at any one thing, but was constantly forced to circumscribe whatever he wanted to identify. In this way *eros* exerts a kind of gravitational force upon each of the speakers. Each is in love with himself, and the force of this self-love is inescapable. Thus with each speech we merely circumscribe *eros*; our *logos* of it remains somehow always incomplete. This, moreover, would seem to suggest that some aspect of the dialogue's teaching cannot be expressed by words, that something about *eros* compels us to compose *logoi*, but that this "something" must always remain an unsaid "something."

Of course, this representation too has its weaknesses. Socrates seems to synthesize the insights of the previous speakers into his own speech. Also, in representing the dialogue as a circle, we must grant an equal status to the perspective of each speaker; yet we want to acknowledge the ascent of philosophical dialogue. On the other hand, the circle seems to remind us that our *logos*, no matter how refined, nonetheless re-

mains incomplete. In other words, each geometrical representation seems to illuminate something about the teaching of the dialogue. Hence we might seek a different representation of the dialogue, under the condition that it allows the dialogue to ascend in pursuit of the supreme object of *eros* (the pyramid), but simultaneously acknowledge the incompleteness of the *logos* of *eros*.

Might we find this representation in a musical analogue? This is not as implausible as it sounds, considering the comparison of Socrates to the flute-player Marsyas. Music begins just where words leaves off, and in this sense it might supply the "something" that was missing from the *logos* of *eros*. Just as the flute-player enchants us with the beauty of his music, so too does Socrates enchant us with the music of his words. Harmony, moreover, expresses ascension, in the sense that the ear naturally inclines towards the higher pitches in a chord. Incidentally, there is a musical phenomenon that should be familiar to anyone who knows a little bit about harmony. This phenomenon is the acoustic illusion. The acoustic illusion is the result of perfect harmony, occurring when the pitches in a chord are so perfectly attuned that they reproduce themselves in a higher register. Though the pitches are present in the chord, the resonance they create at the higher octave can be regarded as an illusion, since they have no corresponding origin in the instrument that sounded them.

Given the musical analogue of Socrates with Marsyas, we might consider whether the acoustic illusion could be extended to moments in Socrates' speech, specifically the ascent passage (210-211d), where Socrates relates the ascent of the initiate towards more temporally perfect objects of *eros*. In order to produce the acoustic illusion, we must identify two pitches in the passage. By 204d, Diotima has defined the basic structure of *eros*. *Eros* consists in the *epithumia* (desire) for an object which one lacks, and inhabits a space in between ignorance and wisdom. So understood, *eros* consists in a relation between the one who desires and the object which he lacks. In the sections that follow, Diotima attempts to locate the object of *eros*, claiming in different passages that *eros* is of the beautiful, the good, of "engendering and begetting upon the beautiful" (206e5) and "of immortality" (207a4). The key to explaining the variance among the objects of *eros*, and identifying the pitches in the acoustic illusion, lies in the features of the structure of *eros* given above. The lover and the object never unite, but are constantly related *through* love. It is precisely this separation which constitutes the erotic relation; if the object and the lover are united, then the erotic relation collapses into itself. In order to understand the variance among the objects of *eros*, as well as their parallel to the acoustic illusion, this basic structure of the erotic relation must constantly be held in mind.

At 204d, following Diotima's definition of the basic structure of *eros*, Socrates asks what "use" love could be to mankind. In an effort to guide him to an answer, Diotima asks him what one who loves beautiful things expects to acquire upon possessing them. Socrates, apparently confused, confesses that he cannot answer. Socrates' silence thus indicates something significant about the relation he bears to Diotima; namely, that it is erotic. Socrates lacks answers, and Diotima alone can supply them. So, the substitution of the good for the beautiful is not at all arbitrary on Diotima's part, but

rather her way of increasing the tension in their relation in a way that actually propels Socrates forward.

The object undergoes a second transformation in Diotima's hands (from the good to begetting and engendering upon the beautiful) when she asks whether men who love the good want the good "merely to be theirs" or "theirs always" (206a6). Thus Diotima places a new strain upon the erotic relation. If the lover came into possession of the object, then the erotic relation would collapse; so in order maintain the erotic relation even in the moment of possession, the two terms must be repelled from each by some other means. This new aspect of the erotic relation is temporal: the object must be present at all times.

By stressing the temporal aspect of the erotic relation, Diotima propels the object of *eros* to a higher mode: the lover no longer finds satisfaction in the temporary possession of the object, so he transforms the object and seeks it on a higher plane. "Begetting and engendering upon the beautiful" becomes the new object of *eros* precisely because it promises what the lover lacks, namely, to reduce the strain of the temporal on the relation: through the begetting of offspring, mankind vicariously experiences immortality. (207a1) On the other hand, the transformed object retains a feature of the original by the addition of the words "in the beautiful." This addition would seem to suggest a confusion of the object at this stage, if it were not for the fact that Diotima later reveals that the beautiful provides the *occasion* for begetting: "Therefore when a person is big and teeming-ripe he feels himself in a sore flutter for the beautiful, because its possessor can relieve him of his heavy pangs. *For you are wrong, Socrates, in supposing that love is of the beautiful* [emphasis added]" (206e1-3). Understood as the occasion for begetting, beauty is not restored as the object of love, but instead creates a secondary tension in the erotic relation that propels the lover towards the object (begetting and engendering upon the beautiful).

Since harmony requires at least two pitches, it can be viewed in two different ways: as resolution or as tension. On the one hand, when the two pitches sound together in harmony, they please us, and we say that they are resolved. On the other hand, the pitches must remain at a distance from each other, or else they will converge and thereby become one and the same. In Socrates' account of *eros*, the two terms that are the conditions for the erotic relation are *epithumia* and *chronos*. When *epithumia* comes into conflict with *chronos*, it must constantly fight against *chronos* to come into possession of its object. But if *epithumia* defeats *chronos*—if it finds its object in eternity—then the erotic relation between the lover and the object collapses; likewise, if *epithumia* is annihilated, the relation collapses. Hence the irony of *eros* consists in the fact that if it succeeds, it fails; if *epithumia* defeats *chronos*, the lover's relation to the object is no longer erotic. Expressed in terms of the acoustic illusion, we might identify *epithumia* and *chronos* as two musical voices of Socrates' speech. Their pitches may sometimes vary, and accordingly the harmonies they produce will determine the location of the object; however, they may never converge upon the same pitch.

The pitches of these two voices in the acoustic illusion produce their most brilliant resonance in the ascent passage. As mentioned above, the ascent consisted in propelling

the object of *eros* to a higher temporal plane. Hence we see the same acoustic principles at work in the erotic relation described above in the ascent passage: the object of *eros* at the first step of the ascent is a beautiful body. However, upon acquiring the body, the lover realizes the temporal imperfection of its beauty, whereupon the predicate "beauty" sheds its subject and acquires a new one on a higher temporal plane, namely, the soul. The object continues to ascend in this same manner, with the predicate "beauty" shedding its subject in favor of one more temporally perfect, until it comes upon the essence of beauty, where the harmony between *epithumia* and *chronos* resonates at its highest register. The object of *eros* is thus propelled by the epithumatic and temporal strains on the basic erotic relation until it comes to the stage in its development at which it becomes supremely erotic, precisely because it has no physical or temporal medium.[19] So regarded, the essence of beauty is an acoustic illusion: having no physical or temporal characteristics, it does not correspond to any worldly origin; but emerging from the epithumatic and temporal forces essential to the erotic relation, we perceive it as the resonance of a perfect harmony.

Unfortunately, a thorough application of the acoustic illusion to the entire dialogue is outside the scope of this paper, since this task would require a vast analysis of both the individual speeches and the structure of the dialogue. Nonetheless, I will present a few leads that might be followed in this application. One lead can be found in the speeches of Phaedrus and Pausanias. It is worth nothing that all of the stories in Phaedrus' eulogy of *eros* celebrate the death of the lover.[20] In this respect, what Phaedrus praises in *eros* we might call the immediacy of time, and we could imagine him saying to Achilles: "the past is forever gone, and the future uncertain, except for death; Do not hesitate, Achilles, for at least this moment of glory you have, at least in death you conquer time, and death is but an instant." So for Phaedrus, *eros* actually despairs over time; it tries to abandon time in the immediacy of the moment. Thus we might say that the voice of *chronos* in Phaedrus' speech is least distinct. In Pausanias' speech, however, the distinction between the young and the old appears: heavenly *eros* consists in the proper relation between a man and a boy, with the man undertaking the intellectual development of the boy, and the boy sexually gratifying the older man. So Pausanias recognizes the maturation of personality, and this change invokes time. Thus in Pausanias' speech, the voice of time becomes more distinct. Accordingly, the acoustic illusion might be extended across the structure of the entire dialogue by following the development of the voice of *chronos* until its culmination in Socrates' speech.

Despite the difficulty in extending it over the structure of the entire dialogue, the acoustic illusion retains the aspects that the step-pyramid and the circle taught us about *eros*. Like the pyramid schema, it allows us to ascend in pursuit of the supreme object of *eros*. It denies us the completion of our *logos*, since Socrates can apparently say nothing positive about the essence of beauty; he can only define it negatively. The language Socrates uses to talk about the form of beauty is particularly revealing in this regard. The form of beauty "neither comes to be nor perishes, neither waxes nor wanes; next, it is not beautiful in part and in part ugly, nor is it such at such a time and other a another, nor in one respect beautiful and in another ugly" (211a-b). Ap-

parently, then, the form of beauty not only stands apart from the spatio-temporal world, but also outside the capabilities of *logos*. But this form Socrates expresses musically, by means of the acoustic illusion: when he speaks of forms and essences, we simply hear a harmonic resonance. From the perspective of the acoustic illusion, Alcibiades' speech actually serves as a kind of chamber that sounds the harmonic resonance of Socrates' speech in our ears. Just as Alcibiades undermines Socrates' ascent towards the universal by pointing out Socrates' radical particularity, so too does he attune our ears to the music of Socrates' words. Socrates, like the flute player, enchants us with the music of his words. We listen to him because the tones he produces are so deceptively simple—"he talks of pack-asses, smiths, cobblers, and tanners" (221e)—and yet so perfectly harmonized that we are prepared to follow them, in the pursuit of truth, to their highest and most sublime resonances.

Received December 2008
Revised March 2009

ENDNOTES

1: Apollodorus purportedly relates the version of *The Symposium* that he heard from Aristodemus. The "madness" for the beauty of philosophy is not as implausible as it sounds, considering that Apollodorus has earned himself the title "*manikos*," i.e., "manic" or "crazy."

2: Before the reader continues, I would like to inform him that this paper places more emphasis on the latter part of this question. I am not interested in representing the entire dialogue with a comprehensive schema; such a task I consider, if not impossible, very difficult. rather, I am more interested in the different aspects of *eros* that the dialogue can teach us about if we posit such-and-such a form.

3: *Eidos* is cognate with *eidein*, the second aorist of the verb ὁράω, which means "to see."

4: In the commentary to his translation of *The Symposium* C.J. Rowe refers to the apparent ascent of each speech as a kind of "capping," which to him indicates that the speeches "represent a single whole, culminating first in the speech by Agathon [...] and then in Socrates' contribution." While Rowe rejects the possibility of continuous development of a more refined account of *eros*, he nonetheless identifies the temptation to view the speeches in an ascent, as is apparent in his choice of the verb "culminate." The pyramid schema captures just this feature of the dialogue. See: C.J. Rowe, *Plato: The Symposium* (Aris & Phillips, 1998), 8. All Stephanus numbers cited herein refer to this translation.

5: Rational account or speech.

6: Under Pausanias' principle, it seems acceptable to say that murder, so long as it is done excellently, can be good.

7: "For the art of medicine may be summarily described as a knowledge of the love-matters of the body in regard to repletion and evacuation; and the master-physician is he who can distinguish there between the nobler and baser loves, and can effect such alteration that the on passion is replaced by the other; and he will be deemed a good practitioner who is expert in producing Love where it ought to flourish but exists not, and in removing it from where it should not be" (186c-d).

8: *Eros* according to Agathon is "κάλλιστον" (195a6), or "most beautiful" of all the gods, and "requires

a poet such as Homer to set forth his delicacy divine" (295c9).

9: In response to Socrates question at 204c9, "of what use is love to mankind?" Diotima leads Socrates through a series of arguments, at whose conclusion she suggests that just as the broader meaning of *poiesi* is production in general, so too is *eros* "generically […] all that desire of good things and of being happy" (205c-d). Apparently, then, *eros* is useful because it generally directs us towards the good.

10: See Socrates' speech at 209b-c.

11: According to Diotima, at the higher stages in the ascent, the initiate will behold the beauty of various "branches of knowledge (ταϛ επιστήμαϛ)" (210c).

12: Diotima relates the story of the birth *eros* in order to describe its nature as constantly in between mortality and immortality, ignorance and wisdom, etc (203b-e). *Eros*' parents are "resource" (*poros*) and "poverty" (*penia*).

13: See sections 200-201 for the questioning by which Socrates forces Agathon to admit that *eros* is not itself beautiful.

14: I draw this from another insightful comment from Rowe: "It is in any case hard to construct any joint account that might emerge from the sequence from Phaedrus to Agathon. All five are essentially individual contributions, with each attempting to go one better than the one before in an apparently haphazard way" (Rowe 1998, 8).

15: See 199b4-6, where Socrates says "decide then, Phaedrus, whether […] you would like to hear the truth told about love in whatsoever style of terms and phrases may chance to occur by the way."

16: See the dialogue between Socrates and Alcibiades at 214e-215a.

17: According to the myth, Marsyas was skinned alive.

18: "All of these possessions [wealth, fame, beauty] he counts as nothing worth, and all of us as nothing" (216e3-4).

19: The essence of beauty is "ever existent and neither comes to be nor perishes" (211a1) and is "not infected with the flesh and colour of humanity, and ever so much more of mortal trash" (211e3-4).

20: These examples include Alcestis, who died for her husband (179b7), and Achilles, who died to avenge Patrocles' death (179e).

Correspondence

Dear Editors:

I found Ross Wolfe's article on Spinoza and Leibniz ["Substance, Causation and Free Will in Spinoza and Leibniz," *Arché*, II:1, 2008] somewhat misled in rejecting Spinoza's rejection of free will. I disagree with Wolfe's implication that Leibniz is necessary to complete or correct Spinoza's system, as I believe Spinoza is not in need of correction, and that Leibniz does not offer a more compelling account of free will.

The fatalistic interpretation of the *Ethics* focuses on the fact that there is a chain of causes preceding any human involvement in an event. There is something that inevitably causes a given human action, and that cause has its own cause, and so on. Fixating on the causes that precede one's actions does seem to emphasize a certain futility in acting. Even so, Spinoza does not intend to belittle man with his theory. On the contrary, man is of equal importance in the causal system as every other cause within it.

Nowhere in the *Ethics* does Spinoza attempt to argue that man does not cause anything. He does argue, of course, that man cannot be the ultimate cause of an event, but this does not demote man in any way. He is on equal footing with all the other causes in the series; like each of them, he has a cause that precedes him, and like all of them, he has an event following him of which he is the cause. By picking out particular passages, it is easy to reach unpleasant conclusions from the *Ethics*, but a fair treatment of the work as a whole reveals that man has no cause for dissatisfaction with his role.

In IP32, Spinoza states that "The will cannot be called a free cause, but only a necessary one…even if the will be supposed to be infinite, it must still be determined to exist and produce an effect by God…"[1] From the fatalist perspective, this claim makes man out to be impotent. This proposition does appear in the text, but the context in which it occurs should not be overlooked. Just before this, in IP29, Spinoza argues that "In nature there is nothing contingent, but all things have been determined from the necessity of the divine nature to exist and produce an effect in a certain way."[2] With this in mind, it is clear that man should not be humiliated by his necessity. Yes, he is determined, but so is everything else, including the causes that determine him. A negative reaction to this argument amounts to the absurd expectation that man should be an exception to the laws arising from the infinite, divine nature of God.

Furthermore, the fatalistic complaints of impotence are unfounded as well. Man is just as capable of acting as a cause as anything else. In fact, it is equally necessary that he cause some effect as it is that he be caused himself, by IP36:

> Nothing exists from whose nature some effect does not follow.
> Dem: Whatever exists, expresses the nature, or essence of God in a certain and determinate way (by P25C), that is (by P34), whatever ex-

> ists expresses in a certain and determinate way the power of God, which is the cause of all things. So (by P16), from [NS: everything which exists] some effect must follow, q.e.d.[3]

By this argument, man should by all rights be satisfied with his abilities. Not only is he capable of causing an effect, he is incapable of not doing so. If Spinoza were to say, "Whatever exists, with the exception of man, etc.," then perhaps there would be cause for concern and dejection. Instead, man comprises a crucial, necessary part of Spinoza's causal system. If he were absent, the effects determined to occur by God would not come to pass; the link in the causal chain that is man would be missing, the effect he produces would not occur and the system would come to an abrupt halt.

From this perspective, man is misled in despairing of his inefficacy. Spinoza is only arguing that there are many causes and many effects, all of which play a part in the unfolding of the universe, and man is one of them. In light of his equal inclusion, it now seems petty and egotistical for man to bemoan the fact that there are other causes in the system. Man is a part of God's essence and as such expresses his power, by which he produces an effect that is likewise crucial to the structure of the universe. Rather than attempt to situate himself above God, beyond infinity and outside of causality, man would be best served by rejoicing in his inclusion in the intricate, crystalline grandeur of the universe that Spinoza describes.

Leibniz attempts to craft an account in which man's free will and autonomy is preserved while still preserving the importance of God and causality. In so doing, he needlessly fractures the unified system offered by Spinoza. By arguing for autonomous monads, Leibniz commits himself to a multiplicity of individuals that he admits cannot influence one another. At the same time, he cannot deny our "continual dependence" on God. At this point, he owes us an explanation of how causality can take place between these individuals at all and what relationship these supposedly autonomous monads have to God. He attempts to reconcile these issues by crediting God with a vague superiority over the other monads, but Spinoza's explanation is still more coherent and convincing. Spinoza would likely charge Leibniz with creating loopholes in an attempt to preserve the very illusion that the *Ethics* tries to dispel. As Spinoza's work demonstrates, reinforcing this deeply ingrained myth is to the detriment both of man's understanding and of his lifestyle.

Ethan Rubin
B.A. Philosophy, 2010
Boston University

NOTES

1: Benedict Spinoza, *Ethics*. Trans. Edwin Curley. New York: Penguin Putnam Books, 1996, 21.
2: *Ibid.*, 20.
3: *Ibid.*, 25.

Dear Editors,

Juliet Johnson's essay ["The Sound of Nietzsche's 'Long Bright Silence': The Interpretation of *Zarathustra* as an End in Itself," *Arché*, II:1, 2008] seems to me a flawed philosophical project. Johnson's paper is threatened by that which many "holistic" accounts of Nietzsche's work fail to avoid: we cannot, by claiming that Nietzsche championed a doctrine of non-adherence to doctrine, eliminate that first degree of systematization. We cannot attempt to escape the commitment to non-interpretive experience of his writing by simply shifting the focus of our readings to a personal critique; this is functionally equivalent to a philosophical program of analysis.

Allow me to illustrate my point with a famous koan: "What is the sound of one hand clapping?" How, I ask you, can one hand clap? Leaving aside the question of whether my response is correct according to the Zen tradition, pointlessness is the point. Johnson's discussion is an attempt to systematize a non-system, and it doesn't work. *Zarathustra* is self-contradictory, both within and between its poetic and philosophical elements; why must we force them together? Is not the very act of dividing poetry and philosophy, for the purposes of their later reunion, the sort of project Nietzsche would reject? Are we not, in other words, bound to evoke the laughter of Zarathustra when we strain to hear the musicality of what he has already revealed as silence?

One of Nietzsche's major campaigns was waged against what he called "other-worldliness." The "dogmatic morals" which Johnson makes reference to were reviled by Nietzsche because they separated man from the immediacy of his own being for the sake of a false "other-world." One passage in this paper is a particularly hermeneutical and philosophical, and therefore other-worldly, approach to *Zarathustra*:

> Thus Spoke Zarathustra is un-hearable, incomprehensible nonsense, until one separates oneself far enough from analytic tradition and lifts oneself high enough above modern divisions to see the whole. (26)

One needn't lift oneself anywhere to understand *Zarathustra*, for understanding is already present in its incomprehensible wholeness.

Dylan Rose
B.A. Philosophy, 2010
Boston University

Contributors' Notes

RACHEL BAYEFSKY is a senior at Yale University majoring in Ethics, Politics, & Economics. As a Rhodes scholar, she will begin graduate studies towards an MPhil in Political Theory at Oxford University next fall. Rachel is interested in political philosophy, epistemology, and the philosophy of religion. She is grateful to have been a student in Professor Keith DeRose's epistemology class.

DOUGLAS KREMM is a junior at the University of Pittsburgh, double majoring in Philosophy and English literature and pursuing a certificate in Russian and East European Studies. He is particularly interested in moral philosophy, value theory, and practical reason. Other interests include 19th-century British and Russian literature, literary theory, and the relationship between literature and ethics. He plans to continue studying philosophy at the graduate level after completing his undergraduate degree.

WILLIAM J. BRADY is a senior philosophy and psychology double major at the University of North Carolina at Chapel Hill. His philosophical interests, broadly construed, are in the areas of mind, cognitive science and epistemology. Within these areas, he is interested in cross-disciplinary approaches, and has focused specifically on the nature of human memory and modern cognitive science methodologies. William plans to pursue graduate study in philosophy.

PETER MOORE is a senior at Boston University. In addition to majoring in philosophy, he has earned a minor in Modern Greek Studies. His primary philosophical interests are in ancient philosophy, though he also has a deep interest in Kierkegaard. He intends to pursue ancient philosophy and classical studies, but will postpone further studies in order to find work in Athens, where he hopes to perfect his knowledge of Greek.

’ARCHÉ’

ἀρχή

IS NOW ACCEPTING SUBMISSIONS.

Undergraduate authors are invited to submit all types of philosophical discourse: research papers, critical analyses, conference proceedings, interviews, translations, letters and commentary on texts previously published in *Arché*, and original philosophical works.

Submissions and correspondence should be sent as Word or .rtf attachment, with “Submission” in the subject heading, to:

buarche@gmail.com

Please include a 100-word abstract. The name of the author(s) should appear on the first page only. Citations should follow Chicago style, with all references occurring as endnotes. Authors are encouraged to query if their submission has not been acknowledged in a timely manner.

ἀρχή

BOSTON UNIVERSITY

http://www.bu.edu/arche