Logos

λόγος

Logos : λόγος

Volume III · 2006–2007

Logos : λόγος

Volume III · 2006–2007

Contents

Editor's Introduction

After a two semester hiatus the *Logos* staff is proud to present the newest edition of our journal. The decision to switch to an annual instead of bi-annual publication came after much debate amongst the editors and officers, but we feel that this has allowed us to produce a fuller and more stimulating volume with enough space for several noteworthy articles in multiple fields. This change has also enabled us to sort through an even larger volume of submissions, thus increasing competition and ultimately the quality of our selections.

It has been a year of transition for several other reasons. Our staff has again increased in size, but we must also say our goodbyes to the last remaining founding editor, Peter Goldstein. His input will surely be missed. From now on an entirely new generation of Cornell philosophers will be in charge of the journal's activities, which have expanded along with the staff. In addition to our meetings on submissions we have initiated a series of lectures and discussion sections with the Cornell faculty, and our thanks goes out to Professor Michelle Kosch for her enlightening presentation on Kierkegaard's *Fear and Trembling*. We would also like to thank Professor John P. Hawthorne of Rutgers University for taking the time out of his schedule to stop by for a group discussion late last year.

As always, when selecting our feature articles submissions were evaluated on the basis of quality of research, depth of philosophical inquiry, ingenuity, and clarity. However for this volume we were also looking for papers that might go well together thematically. To this end we have included two essays on Plato, two on the Philosophy of Language, and two with a refreshing Continental slant to them, in addition to one intriguing paper on Philosophy of Mind. Two of these papers came from fellow editors, but because we feel very strongly that they can and should stand on their own we decided to feature them as main articles instead of incorporating them into a staff submission's section as we have done in years past.

Following tradition we have included an interview with a very well known and well loved Cornell professor. This year Richard W. Miller talked with us about what to expect from his anticipated new book, *Globalizing Justice*. Having written previously on our

duties to the global poor, the relationship between transnational duties and duties towards compatriots, the ethics of war, moral problems of globalization, global climate change, and the moral implications of American power, his latest work is meant to bring all these topics together with the unifying argument that "demanding responsibilities to people in other countries reflect relationships of power that reach across borders". The interview traces the progression of his thought from two of his earlier works up to the present.

Flora Lee writes about the teachability of virtue in Plato. First she uncovers and explains the problems posed by several key puzzles left unsolved in *Protagoras* and *Meno*. Next she suggests that a clever solution to them lies in a careful reading of *Republic*. If we understand the two aporetic dialogues as a kind of preparation for his later work then we come to the conclusion that Plato had, from the very beginning, thought that virtue is, in fact, teachable.

Thomas Vacek examines the internal consistency of *Republic* and argues that it is necessary to distinguish between two different theories of justice in order to resolve contradictions that arise from attempting to apply a single conception of it to both the city and the individuals in it. He does an extraordinary job engaging the original Greek text and offers us his own translation of key sections.

David Friedell provides a clear presentation of the essential concepts involved in Robert Brandom's account of singular terms then examines the account's plausibility. After considering several futile objections he offers his own novel critique and argues that Brandom must reject the view that definite descriptions always carry existential import. This article originally appeared in *Harvest Moon,* Berkeley's campus-wide undergraduate journal, and we are very grateful for the opportunity to publish it for an international audience.

Will Lanier analyzes and assesses various accounts of intensional transitive verbs. In particular, he focuses on Forbes' attempt to provide a viable alternative to both Quinean paraphrasing and Richard's take on Montagovian semantics and notional readings. Lanier shows how Forbes is able to accomplish this by combining Montague's higher-order intensional type theory syntax/semantics and Davidson's event semantics.

Larry McGrath tackles a very difficult text, Theodore Adorno's *Aesthetic Theory*. Once philosophically and historically contextualized, he gives an in-depth interpretation of some of the major points advanced by Adorno given the socio-historical conditions

of the work's production. True to the tradition of critical theory, after interpreting *Aesthetic Theory* the author critically engages the text by reflecting on the role of philosophy given Adorno's claims.

Gavin Arnall explores the major surrealist literary piece *Nadja* by one of the movement's leading figures, André Breton. Combining philosophy, literary criticism, and intellectual history, Arnall demonstrates how *Nadja* represents the surrealist movement's dialectical progression and philosophical/political revolution from Hegelian idealism to Marx's dialectical materialism.

Derrick Ward examines one of the more compelling arguments against the reliability and veridicality of standard visual perception. His piece aims to explicate the inverted spectrum theory and examine its implications for more general perceptual theory.

We would like to conclude by thanking everyone who contributed over the past few semesters, all the authors who dealt with long delays, the editorial staff for its continued commitment, our newest members for their lively interest, the Cornell University Student Assembly Finance Commission, the department secretaries, and, of course, the faculty and staff of the Sage School who never stopped believing in us.

Peter Goldstein
Christopher Post
Logos Editors-in-Chief
Ithaca, New York

The Teachability of Virtue in Plato's *Protagoras, Meno,* and *Republic*

Flora Lee, University of Pennslvania

I. Introduction

The teachability of virtue (*arete*) is a central topic explicitly raised and explored in two of Plato's so-called aporetic dialogues[i]: the *Protagoras* and its direct thematic sequel, the *Meno*.[ii] Here, the "virtue" (or "excellence") of a human being refers to that—whatever it may be—which makes him or her the best kind of human being. The question of whether such virtue can be taught was a topic of great importance in upper class Athens, especially with the emergence of Sophists, who claimed to be professional teachers of it.

Although Plato gives an explicit treatment of this topic in the two dialogues, his position in them is far from being clear. Instead, manifold puzzles arise. First, why does Socrates doubt that virtue is teachable while holding that all virtue is one (essentially reducible to knowledge)? This may seem puzzling because Socrates is also committed to the view in both dialogues that virtue is teachable if and only if it is knowledge (*Protagoras* 136b1–6, *Meno* 87b9–c3). Second, what is the significance of the distinction that Socrates introduces in the *Meno* between virtue as knowledge and virtue as true opinion? Third, what are we to make of Socrates' repeated argument against the teachability of virtue that turns on the empirical premise that there are no teachers of it? Last and most important, what really is Plato's position on the teachability of virtue?

In this paper, I offer an exposition of Socrates' arguments for and against the teachability of virtue found in the two dialogues, and articulate the puzzles that they raise. I shall then suggest that these puzzles can be solved when we understand the two aporetic dialogues as Plato's preparation for the *Republic*, which I will argue gives an implicit "Yes" to the question "Is virtue teachable?"

II. The Teachability of Virtue in the *Protagoras*

The *Protagoras* situates Socrates' discussion of the teachability of virtue in the highly apt setting of a heated debate with the famous sophist Protagoras, with other sophists, students, and admirers listening in. In summary, Socrates expresses his commitment to the following three views, which I shall discuss in turn:

(1) Virtue is not teachable (319b2–320b10, 316–d).

(2) All virtues (namely justice, temperance, piety, courage, wisdom) are essentially one, unifiable or reducible to one and the same knowledge (329c–334c, 349b–351b, 358d–360e).

(3) Virtue is teachable if and only if it is knowledge (361b1–6).

(1) Asked by Socrates what it is that he professes to teach to young men such as Hippocrates, Protagoras answers:

> What I teach is sound deliberation, both in domestic matters—how best to manage one's household, and in public affairs—how to realize one's maximum potential for success in political debate and action (318e5–319a3).

Socrates recapitulates that Protagoras "appear[s] to be talking about the art of citizenship, and to be promising to make men good citizens" (319a4–6), and Protagoras agrees. Understanding virtue as such, Socrates gives two arguments for thinking that it is not teachable.

First, says Socrates, the Athenians, who are surely wise, seem to presuppose for the most part that it is not teachable (319b5–e1). His argument here is that

(A) "for everything that is considered learnable and teachable" (319c1–2), the Athenian Assembly turns *only* to (recognized) experts for advice on that matter (319b7–d1);

(B) but, the Assembly turns to *anybody* for advice on virtue (319d1–5);

(C) so, the Assembly does not think that virtue is something learnable and teachable (319e1).

Second, continues Socrates, "the wisest and best of our citizens" (319e2–3), i.e., great statesmen and generals, are unable to transmit their virtues to their sons, even though they surely must try very hard to do so (319e1–320b5). Socrates' point here seems to be that if a recognized possessor of virtue cannot teach it, then the only

plausible explanation for it is that virtue is not something teachable. His argument can be re-constructed as follows:

(i) virtue is teachable if its possessor is able to teach it. (an implicit premise)

(ii) "the wisest and best of our citizens" (319e2–3) possess virtue (an assumption Socrates makes at 319e2–320b4).

(iii) "the wisest and best of our citizens" (319e2–3) are unable to teach virtue (319e2–320b4).

(iv) So, virtue is not teachable.

Protagoras, who is committed to the view that virtue is teachable, responds to these two arguments in the passage that is now commonly referred to as "Protagoras' Great Speech" (320c–328d). Upon hearing the speech, however, Socrates does not bother to consider the replies given in it. Instead, he changes the subject and asks Protagoras to clarify "one little thing" (329b6):

> You said that Zeus sent justice and a sense of shame to the human race. You also said, at many points in your speech, that justice and temperance and piety and all these things were somehow collectively one thing: virtue. Could you go through this again and be more precise? Is virtue a single thing, with justice and temperance and piety its parts, or are the things I have just listed all names for a single entity? This is what still intrigues me (329c3–d3).

(2) This request is the segue into Socrates' project of eliciting and arguing against Protagoras' view that justice, piety, wisdom, temperance, and courage are different and separate from each other. Socrates expresses his own view that these virtues are essentially one and reducible to knowledge, and tries to convince Protagoras of it by getting him to see that different pairs of virtues can each be shown to be identical (329c–334c, 349b–351b, 358d–360e).

(3) At the end of the dialogue, Socrates presents his thesis that virtue is teachable if and only if it is knowledge.

> It seems to me that our discussion has turned on us, and if it had a voice of its own, it would say, mockingly, 'Socrates and Protagoras, how ridiculous you are, both of you. Socrates, you said earlier that virtue cannot be taught, but now you are arguing the very opposite and have attempted to show that everything is knowledge—justice, temperance, courage—in which case, virtue would appear to be eminently teachable (361b1–4).

From this passage, we get:

> (a) If virtue is knowledge, then it is teachable.

Socrates goes on to say:

> On the other hand, if virtue is anything other than knowledge, as Protagoras has been trying to say, then it would clearly be unteachable (361b4–6).

So, now we have:

> (b) If virtue is not knowledge, then it is unteachable.

We can restate this as:

> (b¹) If virtue is teachable, then it is knowledge. (by *modus tollens*)

Putting (a) and (b¹) together gives us the following bi-conditional:

> (c) Virtue is teachable if and only if it is knowledge.

The above three views of Socrates relevant for assessing his position in the *Protagoras* on the teachability of virtue raises the following puzzle: why does Socrates—who thinks that (2) virtue is knowledge and that (3) virtue is teachable if and only if it is knowledge inconsistently maintain that (1) virtue is nonetheless not teachable?

Examining Socrates' two arguments against the teachability of virtue, it is unlikely that he truly takes seriously his first argument, for its conclusion is about what

the masses believe. His second argument against the teachability of virtue appears to be the more important of the two. In the *Meno*, he is committed to the view that if something is teachable then there must be teachers of it (*Meno* 89d5–7), which resembles his implicit premise in the *Protagoras* that virtue is teachable if its possessor is able to teach it. However, there is a problem with Socrates' arguments that hinge on the premise that there is no one who is able to teach virtue, which will be explained at the end of the next section.

III. The Teachability of Virtue in the *Meno*

What we find in the *Meno* are Socrates' arguments both for and against the teachability of virtue. Taking the "pro" stance, he gives two arguments for the claim that virtue is knowledge, in order to argue that it must therefore be teachable. Then, taking the opposite stance, he argues that because there are no teachers of virtue, virtue is not teachable (and hence not knowledge either), and advances a second argument against the teachability of virtue towards the end of the dialogue, relying on the distinction between knowledge and true opinion. This section lays out these arguments, and then articulates some of the puzzles that they, in conjunction with Socrates' views in the *Protagoras*, raise.

Now, the dialogue opens with Meno's question, "Can you tell me, Socrates, can virtue be taught? Or is it not teachable but the result of practice, or is it neither of these, but men possess it by nature or in some other way?" (70a1). In order to answer this question, replies Socrates, one must first know what virtue is, which echoes the ending of the *Protagoras* where the same principle of the priority of definition is expressed, albeit very subtly, at 361d2–4. Subsequently, Socrates goes on to spend a great part of the dialogue explaining to Meno that they must first know what virtue is before they can inquire after its qualities including its teachability (71a2 ff.).

However, despite Socrates' efforts and their joint failure to define virtue (71e–80a), Meno blithely asks him again at 86c–d whether virtue is teachable. Socrates protests again for the same reasons (86d3–e1), but so as not to reiterate the discussion since this question was first posed at the beginning of the dialogue, he propounds a different

approach to investigating into the teachability of virtue—"by means of a hypothesis . . . the way geometers often carry on their investigations" (86e3–4).

III.A. An Argument For the Teachability of Virtue

Socrates asks Meno what sort of thing virtue must be if it is teachable (87b8–9). Taking for granted the view that something is teachable if and only if it is knowledge, Socrates gets Meno to agree that their hypothesis shall be this: virtue is a kind of knowledge (87b9–c3). So, in the *Meno*, we see Socrates first of all advancing the following argument for the conclusion that virtue is teachable:

(1) Hypothesis: Virtue is a kind of knowledge (87b9–c3).

(2) Something is teachable if and only if it is knowledge (87b9–c3).

(3) So, virtue is teachable (87c4–5).

This argument turns on the truth of the hypothesis just posited, so the next thing Socrates does is to argue for (1) above. He seems to present two different lines of argument, which I will call A and B, respectively. They both begin from the premise that virtue is something good—a second and higher hypothesis.

Argument A for (1):

(1) Virtue is something good (87d2–3).

(2) If something is good, then it is a kind of knowledge (87d4–7).

Argument B for (1):

(1) Virtue is something good (87d2–3).

(2) If something is good, then it is beneficial (87e2–3).

(3) Hence, virtue is beneficial; i.e., virtue benefits its possessor (87e4).

(4) What benefits us is understanding, wisdom, or knowledge that
enables us to make a right use of those "neutral things" that are
in themselves neither beneficial nor harmful, but are capable of

being either—e.g., health, strength, beauty, wealth, and all the qualities of the soul (moderation, justice, courage, intelligence, memory, munificence, etc.) (87e6–88e).

(5) Virtue is something in the soul (88c4).

(6) Since virtue is something in the soul and beneficial, it is understanding, wisdom, or knowledge that enables us to make a right use of the qualities of the soul (follows from 3, 4, 5; 88c4–e).

(7) In other words, virtue is knowledge (a restatement of 6; 89a3).

We notice that we have encountered in the *Protagoras* the premises that virtue is a kind of knowledge and that only knowledge and nothing else is teachable. What is novel and interesting about the *Meno*'s treatment of the former premise—the more controversial of the two—is Socrates' appeal to a higher hypothesis containing the term "good" to ground this lower hypothesis. One may ask: why does Socrates think that it is acceptable to appeal to this higher hypothesis? I will briefly address this question later in the paper. For now, let us proceed to Socrates' next argument, one against the teachability of virtue.

III. B. An Argument Against the Teachability of Virtue

Much to Meno's surprise, Socrates immediately expresses doubts about the status of virtue as knowledge (89c5 ff.). He goes on to give an argument against the claim that virtue is knowledge, with an intermediate conclusion that it is not teachable either:

Argument against (1):

1. If something is teachable, then there must be teachers (and learners) of it (89d5–7).

2. But there are no teachers of virtue (89e4–96b8).

3. Hence, virtue is not teachable (96c–d3).

4. Hence, virtue is not knowledge either (because of the bi-conditional relationship stated at 87b9–c3 between being knowledge and being teachable).

Premise (1) is granted by Meno. So, it is obviously premise (2) that carries the weight of this argument. In support of it, Socrates says that he has searched for the teachers of virtue with great efforts but could never find any (89e5–6). He adduces sophists and "worthy people"—i.e., great Athenian statesmen and generals—as the only possible candidates (96b6–8), but the former are dismissed by Anytus (91c1–4, 92a5–b3, 92e4–6) who briefly joins Meno and Socrates in conversation, and the latter by Socrates (93a–94e).

III. C. Another Argument Against the Teachability of Virtue

Confused by the result of the last argument, according to which virtue is not teachable, Meno asks Socrates "whether there are no good men either, or in what way good men come to be" (96d2–3). Socrates responds by introducing a distinction between knowledge, and right or true opinion:

> . . . we failed to see that it is not only under the guidance of knowledge that men succeed in their affairs, and that is perhaps why the knowledge of how good men come to be escapes us (96e1–4).

Socrates' point here is that right opinion, just like knowledge, also enables us to make correct uses of the so-called "neutral thing." In other words, "true opinion is in no way a worse guide to action than knowledge" (97b). He proceeds to explain that both knowledge and true opinion do not come by nature—they are acquired—(98d) and that true opinion differs from knowledge in that the person who has merely the true opinion about something is unable to give an account of it (98a). Socrates suggests that perhaps they have overlooked the possibility that virtue is a sort of true opinion, and not knowledge. Hence an alternative to the original hypothesis that virtue is knowledge is proposed as this: virtue is (god-granted) right opinion.

This new hypothesis implies that virtue is not teachable, because of its origin as a gift from the gods, but more importantly because of the defective nature of right opinions (namely, that they lack understanding). What are the merits of this new

hypothesis? Socrates seems to think they are the following: right opinion is different from and inferior to knowledge, but nonetheless close enough to it to make sense of (a) how we could have mistakenly identified virtue as knowledge, and of (2) the fact that we could not find any teachers of virtue. Like the *Protagoras*, the *Meno* ends with Socrates' reminding his interlocutor of the priority of definition: we must first know what virtue is before we can know for certain whether it is teachable or not.

Now that we have seen Socrates' various arguments in the *Protagoras* and the *Meno* on the teachability of virtue, we are in a position to recognize and articulate the manifold puzzles that arise from them:

(1) First of all, what is the implication and significance of the new hypothesis that virtue is a right opinion in light of Socrates' seemingly strong conviction both in the *Protagoras* and the first part of the *Meno* that virtue is knowledge? Is the thesis that virtue is knowledge (or, less strictly, that virtue is somehow closely related to knowledge) really abandoned by Socrates?

(2) Of course, the puzzle that cannot escape us is: what really is Socrates' answer to the question "Is virtue teachable?"

(3) Some of Socrates' arguments against the teachability of virtue depend on the crucial premise that there are no teachers of it. But, as Kahn points out, the fact that

> there [presently] are no generally recognized teachers of virtue . . . does not prove that teachers of virtue are non-existent, much less impossible. Even if there were really no teachers of virtue to be found, that would only mean that virtue cannot be taught under present circumstances, not that it is unteachable (Kahn, 312).

So, the question becomes: "what does Socrates really think of the significance of this premise?"

IV. The Solution in the *Republic*

My thesis is that these central puzzles raised by the *Protagoras* and the *Meno* can be solved when we see that these two dialogues play the role of preparing us for the *Republic*'s affirmative answer to the question "Is virtue teachable?"

A few words about the dialogue are in order: the *Republic* differs stylistically and substantially from the *Protagoras* and the *Meno*, which are, as we have seen, aporetic. The Socrates of the *Republic* expounds his own views much more explicitly and conclusively, and does not hold back so much. That is, he gives us such great depth and breadth of insights into the overarching Platonic philosophy—e.g., Plato's momentous doctrines of the tripartite soul and of the Forms, and his equally significant philosophical views on education and its aim—that I think we can safely infer Socrates' real position on the teachability of virtue in this dialogue and solve, retrospectively from the vantage point of the *Republic*, the central puzzles that have emerged from the two earlier dialogues.

So, let us first turn our attention to the insights the *Republic* has to offer on the topic. The most logical first step, it seems to me, is to follow the highly under-appreciated advice of Socrates in the *Protagoras* and the *Meno*: define virtue first and then inquire into its teachability. So, what is virtue? More precisely, what is the virtue of a human being? I think that Socrates' (or Plato's) answer to this crucial question is given finally in the *Republic* in Book IV, most clearly at 441d4–444e5: it is the justice or health of the soul (*psyche*), which in turn is the harmony or the right ruling-and-ruled relationship of the soul's three parts (rational, spirited, and appetitive).

> . . . isn't to produce justice to establish the parts of the soul in a natural
> relation of control, one by another . . . ?
> Precisely.
> Virtue seems, then, to be a kind of health, fine condition, and well-
> being of the soul . . . (444d8–e1).

IV.A. The Solution to the First Puzzle

This definition of the virtue of a human being helps us, first of all, to solve our first puzzle. Notice, first of all, that the perfect harmony between the three parts of the soul exists when the rational part rules over the other two—the spirited and appetitive parts—that have become willingly obedient (441d7–443e). Next, notice that Books II–VII of the *Republic* can be read as suggesting that there are two ways in which this harmony is sustained: it can be sustained either by true opinions or by knowledge. The former is (i) the virtue hypothesized as (god-granted) right opinion in the *Meno*, and (ii) the virtue of the auxiliary class in the *Republic*. The latter is (i) the virtue hypothesized as a kind of knowledge in the *Meno*, and (ii) the (true) virtue of the philosopher-kings in the *Republic*.

Thus, the answer to my earlier question "Is the thesis that virtue is knowledge (or, less strictly, that virtue is somehow closely related to knowledge) really abandoned by Socrates?" is: "No." This is the correct answer, insofar as the philosopher-kings' knowledge in general (in the full Platonic sense) and their knowledge of the Form of the Good in particular are what *sustain* their virtue and what *enable* the auxiliary class to acquire and maintain their virtue. Retrospectively, we can see that one role played by the *Protagoras* and the *Meno* was that of gradually familiarizing the reader with Plato's conception of virtue as the health of the soul, which comes in two varieties, so to speak.

IV. B. The Solution to the Second Puzzle

The definition of virtue given in the *Republic*—as the justice, health, or harmony of the tripartite soul—shows that the question "Is virtue teachable?" turns on the question "Can this harmony of the tripartite soul be taught, and if so, how?" In a nutshell, I take Plato's answer to be this: yes, it is teachable in theory (to those born with the suitable nature (430a2, 495a4–b)). There is the qualification "in theory" in Plato's answer as I put it, because in Plato's view, the sort of educational system under which virtue can be taught is so far from being realized in Athens as it existed in Plato's time

that a radical re-configuration of the entire city is called for to establish it. Plato's allegory of the Cave in Book VII suggests that he conceives the aim of education as that of harmonizing people's souls—i.e., re-ordering people's desires in such a way that their rational desires rule over their spirited and appetitive ones—and enabling them, to the best of their natural abilities, to make their journey upwards to the intelligible realm.

I shall now explain *how* Plato thinks virtue can in theory—i.e., in his *kallipolis*—be taught. Since I hold that Plato distinguishes between two distinct degrees of virtue in *kallipolis*, I will treat them one by one.

First, we turn to the civic virtue of the auxiliary class. The question I want to answer is this: how does Plato think that a person can harmonize his or her soul and maintain it somehow without understanding? I take Plato's answer to be: through the right habits and intuitions provided by (i) an appropriate purified curriculum in music and poetry—i.e., without the depictions of falsehood concerning the Gods (e.g., the depictions of their doing bad things), unsuitable portrayals of heroes (e.g., the portrayals of their fearing death or indulging in excessive pleasure), and in general anything else that is false or unsuitable—(Books II–III generally; especially 377e5 ff.), and (ii) a complementary living environment—e.g., a common nursery, no private family or property, etc. (Book IV–V generally).

This extremely encompassing educational system, which permeates nearly every aspect of the auxiliary class' way of life, is designed to inculcate in the young—from the very beginning in their most formative and malleable years—those qualities, beliefs, and behaviors that are in agreement with and will help bring about a well-balanced state of the soul (377a10–b7, 401d4–402a4, 522a4–7). This is how "civic" virtue is taught in the *kallipolis*. Plato knows all too well that the later the intervention to straighten a leaning tree, the more difficult and less possible it becomes (378d4–e1, 491e7–493a2); so, he seeks to ensure that trees grow straight from the very beginning.

Let us now turn to the higher degree of virtue possessed by the philosopher-kings, the harmony of the tripartite soul based on knowledge. What is required to acquire this virtue is the intellectual contact with the Forms[iii], especially with the Form of the Good, the most important Form of all. The following passages on the importance of seeing the Form of the Good are particularly illuminating:

. . . you've often heard it said that the form of the good is the most important thing to learn about and that it's by their relation to it that just things and the others become useful and beneficial. . . . if we don't know it, even the fullest possible knowledge of other things is of no benefit to us, any more than if we acquire any possession without the good of it. Or do you think that it is any advantage to have every kind of possession without the good of it? Or to know everything except the good, thereby knowing nothing fine or good? (504e6–505b2)

In the knowable realm, the form of the good is the last thing to be seen, and it is reached only with difficulty. Once one has seen it, however, one must conclude that it is the cause of all that is correct and beautiful in anything, that it produces both light and its source in the visible realm, and that in the intelligible realm it controls and provides truth and understanding, so that anyone who is to act sensibly in private or public must see it (517b7–4).

Intellectual contact with the Forms gives the philosopher-kings knowledge in the full Platonic sense; they are now able to give an account of what each of the things that exist is. In particular, through their knowledge of what the Good is, they understand what things are good in the visible world of human beings, and how and why the three parts of people's souls must be harmonized for them to live a good life. (Recall Socrates' appeal in the *Meno* to the higher hypothesis that virtue is something *good*, in order to ground the hypothesis at issue: virtue is knowledge. It is now clear why he felt justified in making such a strategic move there.)

Can a person with a suitable nature be taught to "make [this] ascent and see the good" (519c7–8)? Socrates seems to think so, in that he proposes in Book VII a higher level of education beyond the elementary one of the auxiliary class, including 10 years of training in mathematics—as propaedeutic to dialectic—and 5 years in dialectic—the most important training of all, and tells us that "dialectic is the only inquiry that travels this road, doing away with hypotheses and proceeding to the first principle itself, so as to be secure" (533c6–8).

IV. C. The Solution to the Last Puzzle

Now that we have seen Socrates' educational proposals for the *kallipolis*, I believe that we can solve the remaining puzzle: how are we to understand the following argument against the teachability of virtue that Socrates repeatedly makes?

(1) There are no teachers of virtue.

(2) If something is teachable, then there must be teachers of it.

(3) Thus, virtue is not teachable.

I think that it is clear now, from the vantage point of the *Republic*, why Socrates gives this argument and what he actually means by it. As he sees it, there presently are no teachers of virtue and there never will be any in Athens as it exists. For there to be teachers of virtue, Athens would need to be transformed into the *kallipolis*— a city whose educational systems and entire constitution are carefully designed by the philosopher-kings to function as "teachers" of virtue. So, Premise (1) as held by Socrates actually reads "There *presently* are no teachers of virtue." Accordingly, the conclusion Socrates intends to reach with this argument is also qualified by "presently"—what he actually means by (3) is that virtue is not teachable *under present circumstances*. Most likely, Socrates omitted this important qualification in the *Protagoras* and the *Meno* because he needed to first delineate what the *alternative* circumstances would be (as he does in the *Republic*) before he could introduce it to his unsuspecting interlocutors.

To conclude, a reading of the *Protagoras* and the *Meno* from the vantage point of the *Republic* dissolves all the central puzzles that I have shown to arise from the two dialogues on the teachability of virtue. Such a reading of these three dialogues is essentially an application of the unitarian view of Plato's dialogues presented by Charles H. Kahn in *Plato and the Socratic Dialogue*[iv]. That is, I have tried to argue that a careful examination of an important thematic connection between the three dialogues— the teachability of virtue—shows that an interpretation that does *full justice* to the way Plato treats the same topic in these different dialogues (e.g., the specific context in which a dialogue treats the topic, the manner in which Plato deals with it, the extent to

which the topic is addressed, and so on) is one that posits that he has comprehensively composed his dialogues all under a single, unified philosophical vision that remains unchanged throughout his career. In contrast to the developmental view of Plato's dialogues, according to which Plato fundamentally changes his philosophical view from the time of his early dialogues to the later ones, this unitarian view sees that he deliberately holds back from fully disclosing his view in his early dialogues, plants hints and references to the dialogues to come, and presents different aspects of his ultimately single philosophical view in different dialogues as if showing different faces of the same die. As I have tried to show with the example of the topic of the teachability of virtue, the purpose behind such ingressive exposition of his view is to engage the attention of an exoteric audience with his early, aporetic dialogues such as the *Protagoras* and the *Meno*, and then to progressively introduce it in middle works such as the *Republic* to his full philosophical view—without doubt a radically new view at the time.

Works Cited / Further Readings

Kahn, C. H. *Plato and the Socratic Dialogue*. Cambridge: Cambridge U. P., 1996.
Plato. *Complete Works*. ed. John M. Cooper. Indianapolis: Hackett Publishing Company, 1997.

Endnotes

[i] Aporetic dialogues are inconclusive dialogues, in which Socrates' position on the topic under discussion is (perhaps deliberately) not entirely clear.

[ii] See Kahn, *Plato and the Socratic Dialogue*, p. 212 for an explanation of the thematic connection between the *Protagoras* and the *Meno*.

[iii] Roughly, Plato's Forms are eternal and unchanging beings that exist in a non-sensible realm and that we qualify by the words "what it is."

[iv] See especially "The interpretation of Plato," pp. 36–70.

Justice and the City-Soul Analogy in the *Republic*

Thomas W. Vacek, Saint John's University

Plato's *Republic* enumerates a conception of justice in Book iv which is called formal because it does not specify the motivations and activities one's life should contain to accord with it. However, when Plato analyzes possible contents of lives in Books viii and ix, he thoroughly disparages some of the life contents which are inherent in his formal conception of justice in Book iv. His specific conception of justice in Books viii and ix would appear to condemn large numbers of people to unhappy lives, even in his ideal city. This interpretive problem can be eliminated by distinguishing two different theories of justice in the *Republic*. As additional support to the validity of the distinction, to interpret the *Republic* with a single theory would entail accepting that Plato made a grave logical mistake in his analogy of city to individual.

The underlying theme in the *Republic* is the analogy between justice in a city and justice in a person. Moreover, as the work progresses, the relationship becomes much more than an analogy, for the work makes pronouncements on justice both in individuals and in cities, as well as the relationship between them. However, the shift from the analogous relationship to the causal relationship presents problems of interpretation. The analogy begins from the fact that "justice" can be predicated to both a single individual and a whole city.[i] For the analogy to provide insight the cause of justice in an individual would have to be the same as the cause of justice in the city; therefore, we would expect independent analyses of justice in a city and of justice in an individual, so that we could compare them and decide if the analogy has any meaning. Socrates appears to endorse this approach; after constructing the ideal city and determining where justice lies in it, he offers to modify later the recently-completed theory if justice in the individual turns out different.[ii] Holding out for independence in the two approaches, however, would mean that Socrates would be at exactly the same dilemma in Book iv as he was in Book ii when he proposed the analogy as an aid to understanding justice in the individual. Socrates would need to determine what sort of different causes can be found in the soul. There is nothing logically compelling at this point to show why the soul need be tripartite, like the city.

Shortly afterward, however, Socrates shifts the argument, and even acknowledges the shift: "It is my opinion that we would never understand it accurately from these methods which we are employing in the discussion . . . nevertheless, perhaps we may

understand something worthy of the things having been prefaced and examined" (435d). Socrates argues that when a term is applied to two things, they must be alike in how they are called the same,[iii] and therefore the just person *must* resemble the just city. Earlier, the analogy was tentative, serving as a point of departure to see if there were any similarities between justice in a city and in an individual. Here, however, Socrates wants the analysis of justice in the individual to proceed along the same lines as justice in the city *because* the same term can be applied to both. Rather than independent analyses of individual and of city, Socrates argues that, because a just city has particular characteristics, a just individual must be similar: "And indeed the just man will not differ from the just city according to the form itself of justice, but will be similar" (435b). Finally and most significantly, Socrates attempts to strengthen the assertion that the tripartite analysis of the city applies to the individual by arguing that a given characteristic of a city results from constituents who have that characteristic: "For it would be a laughable thing if someone would think that the spirited part hasn't arisen from the private citizens in the cities who have this characteristic" (435e). Therefore, a person's soul must have the same three dispositions found in the city because the latter could have no other origin than the city's inhabitants.

To summarize, Socrates makes two assertions after 435a: First, because the term "justice" can be applied to both a city and a person, the form (εἶδος) of justice will be the same in both. Second, a city has particular characteristics because its inhabitants have those characteristics. As Bernard Williams points out, these two assertions contain an inner contradiction (Williams, 256)[iv] Inverting the order, we have: (1) "x is F if and only if x has constituent parts which are F," (*Ibid*) and (2) the constituent parts are F in the same way that x is F. Combining these two statements leads to an infinite regress. Are x's constituent parts F because of their constituent parts, and so on? Let us examine more closely how Plato argues that a characteristic of a whole is related to its parts, which, following Williams' terminology, we shall call the "whole-part" relationship.

In general, there are three ways to relate parts to a whole, and we can find all of them in the *Republic*.

    i. x is F if and only if ALL its constituent parts are F

    ii. x is F if some of its constituent parts are F

iii. There is no relation between x being F and its constituent parts being F.

Clearly, statement iii utterly rejects that a part has anything to do with the whole, while statement i holds the relationship as absolute. Statement ii is middle ground, but it seems to demand some further criterion to determine the relationship fully. For instance, most cars have black tires; however, only a fraction of cars with black tires are black cars. On the other hand, black cars may have headlights, hubcaps, or other parts that are not black without calling into question the blackness of the car itself.

Williams argues that there is textual support in other contexts for how statement iii treats the whole-part relationship:

> . . . answering Adeimantus' objection that the guardians get a thin time of it, Socrates says that that a city's being sublimely happy does not depend on all, most, the leading part, or perhaps any of its citizens being sublimely happy, just as a statue's being beautiful does not depend on its parts being severally beautiful. (Williams, 256)[v]

Williams' point is well taken. The ideal city is happy, whether or not its constituents are sublimely happy. On the other hand, the *Republic* is not concerned with the attainment of sublime happiness, but justice. In the *Republic*, the relationship between justice and happiness is not that just people are sublimely happy, but that just people are (comparatively) *happier* than unjust people.[vi] Therefore, we could read the passage simply as Plato trying to avoid having to give a more complicated analysis than is needed for the purpose of the work. Moreover, one would hardly expect the ideal city to be inhabited by unhappy residents. Therefore, statement iii remains as a logical explanation of the whole-part relationship, though it is questionable whether Plato intended it in this way.

The whole-part relationship pertains to the way the ideal city has its virtues. Plato accounts for two of its virtues—wisdom and courage—according to statement ii. The city is wise because a small class within it has the characteristic of being wise: "By the smallest class and portion of the city and the knowledge in it, the part leading and

ruling, the whole city founded according to nature would be wise" (428e). Similarly, a city is courageous not because this quality is held by everyone, but only by some: "And I say, would someone say that a city is either cowardly or courageous looking at a something other than at the part which guards and serves in her behalf? I don't suppose, I said, the other people in her, being either cowardly or courageous, would be the authority that the city is one or the other" (429b). However, statement i explains temperance: "because not as courage and wisdom, each being in some part, made the city either wise or brave, [temperance] does not operate in this way, but has been spread evenly through the whole city through all the people, making them people singing together the same thing, the weakest, the strongest, and those in the middle, on the one hand if you please, in wisdom, and on the other hand if you please, in strength, and, if you please, in number or in goods or in whatever other of these things" (431e.). What relation between the whole and the part does Plato intend for justice? Is a city just because only some of its inhabitants (the ruling class) are just, or does justice, like temperance, extend through the whole population?

Recall that statement ii required some further criterion. What is the criterion that allows the ideal city to be wise or brave when only a few of its members are? Concerning the perverted cities of Book viii, there is no prohibition against democracies and tyrannies having a few logistic and thymoeidic individuals in them (provided they keep a low profile, in the case of a tyranny[vii]). But certainly these cities are not logistic or thymoeidic. So there must be some criterion that allows a whole city to be characterized by only a segment of its population. The text indicates that justice allows each of the ideal city's parts to carry out its work for the benefit of the whole, for Socrates says: "It seems to me, I say, the thing left over from what we have examined, from temperance and courage and wisdom, is this, the thing which provided power to all them so that they may arise and it seems to me that it provides preservation for them having arisen, while it is in the city" (433b). So, justice is the criterion for statement ii, since it explains the whole-part relationship for the other qualities. Now, if we try to account for justice in the whole city according to statement ii and with that criterion, we are faced with manifest absurdity, that a constituency of just individuals makes a city just if the city is just. Clearly this is no explanation, so we must account for justice in a city in some other way than statement ii.

We are left with statements i and iii. On the one hand, statement iii would entail that a city can be just without its leaders being just, but the text contradicts such a reading. Despite the quaintness of the explanation of why the ideal city's guardians become perverted—Socrates blames the inevitable perversion on the guardians failing to heed that the quality of offspring waxes and wanes according to an elaborate schedule[viii], it is clear that the ideal city becomes perverted along with its guardians.[ix] On the other hand, Williams argues that applying statement i to justice, that a city is just because every individual in it is just, leads to a contradiction:

> The δικαιοσύνη of a city, as of anything else, consists in [each of the city's elements doing its job]. So in order to be δίκαιος, a city must have a logistic, a thymoeidic, and an epithymetic element in it. Since it must have an epithymetic element, it must, [since an element of a city is logistic, thymoeidic, or epithymetic if and only if its men are], have epithymetic men: in fact, it is clear from Plato's account that it must have a majority of such men, since the lowest class is the largest. But an epithymetic man is not a δίκαιος man; if he is not, then the city must have a majority of men who are not δίκαιος, which contradicts [statement i]. (Williams, 257)

Williams concludes from this that the similarity between the whole and the part cannot hold. Therefore, he concludes that Plato intends justice to be found only in some individuals in a city. In this case, the criterion that makes the whole take on the characteristic of a part is that these individuals must rule the city. The relationship between the whole and the part becomes that "a city is *F* if and only if the leading, most influential, or predominant citizens are *F*" (Williams, p. 259). Unfortunately, this conclusion winds up looking more unpalatable than the contradiction, for it has the effect that the philosopher kings rule by force an epithymetic, contumacious city. Williams writes:

> So what we have to believe, it seems, is that cobblers are characteristically men of powerful passions—of more powerful passions,

indeed, than soldiers—who nevertheless have enough rational pow-
er to recognize the superiority of philosopher kings when there are
philosopher kings, but become unmanageably volatile when there
are no philosopher kings. (Williams, 262)

In general, the ideal city hardly looks like a happy city with Williams' treatment of it. However, I argue that Williams takes a wrong turn in his analysis, and that his objections are answerable. Statement i does explain how justice exists in the whole city.

To interpret Plato consistently, we shall have to change Williams' interpretation in two ways. First, we must go back to the analogy between city and individual. We can logically say that a city is *F* because its citizens are *F*, and not be plagued with problems of regression, if the component parts of an individual arise and relate to each other differently than their analogous parts in the city. Second, we must examine Williams' *reductio* of statement i, that the just city would have a majority of men who are not δίκαιος. The only way to resolve this contradiction is to hold that every individual in the city, even an epithymetic individual, must be δίκαιος, since we need to say that a city is *F* because *all* its citizens are *F*. However, we need not approach this statement as if we are doing our best to gulp down one logical inconsistency in order to avoid more; rather, we shall see that the *Republic* contains two theories of justice: one which is objective and by which a city is just, and one which is subjective and concerns the individual even if the individual be epithymetic in the scheme of the city. This distinction answers two problems: first, the problem of the regression in the definition of justice, and second, how we could call an individual both epithymetic and δίκαιος.

The objective standard of justice is the justice which Socrates describes in the foundation of the city, in Books ii–iv. From the very beginning, the city is founded on a principle of efficiency in fulfilling basic human needs. Socrates gives two proposals to Adeimantus, first that individuals share and second that individuals could live completely self-sufficiently. Socrates gives the second proposal as follows: "Or is it necessary for [the farmer], neglecting [others], to make a quarter portion of the food—the portion for himself—in a quarter portion of the time, and to spend the three, one for the preparation of the house, one for clothing, and one for cobbling, and not to have

common matters with others, but himself to do his things by himself?" Adeimantus' answer is telling: "Perhaps the first way would be easier (ῥᾷον) than the second" (369e). The objection against the second proposal is not impossibility; both are possible, but one is easier. There are *no* basic needs that can be furnished only in the context of a community, so the formation of a city is not necessary to fulfill an individual's basic needs. Though Socrates states earlier that the individual is in need of many things (ἐνδεὴς πολλῶν (369b)), those needs do not compel the foundation of a city; rather, a city alleviates those needs more easily. Hence, the principle of formation for the city is not compulsion of need, but making life easier for the inhabitants. I believe Plato makes a significant implicit point here, that chaotic diversity in a person's daily activities is a debilitating condition, even if the diverse ends sought are not bad in themselves.[x] Therefore, a city which seeks to make its inhabitants' lives easier will remove anything which is unnecessary for a person's life and work and detracts from them. Therefore, at this early stage, justice is avoiding conflicts of external goods, such as whether a person should grow food or build a house. Organization of these external things allows a person to focus on one meaningful activity. From the principle of inhabitants living easily, Socrates formulates the division of labor because "another shall work upon another task" (370b). Justice, as presented here, is not a problem of reason, for there is no hint at this point in the foundation of the city that reason will be needed to establish justice.[xi]

The foundation of the city has several stages. Socrates first constructs an extremely austere city, called the City of Pigs after Glaucon's assessment of it.[xii] This city provides a simple life for its inhabitants, which Socrates describes at 372a. This city has no need of guardians; it is so simple that there can be no conflicts between the external goods that the people pursue. However, the impossibility of conflicts of goods raises an interesting question. How can we say that the City of Pigs is just? If the City of Pigs were not mentioned in the *Republic*, the text would appear to define justice as organization of conflicting—or at least possibly conflicting—diverse elements. The properly organized city is spirited because spirited individuals defend it and wise because wise individuals rule it.[xiii] However, the City of Pigs has no possibly conflicting elements, whether in the desires of its inhabitants or in classes of inhabitants. Nevertheless,

Socrates describes the City of Pigs as "true" or "real" (ἀληθινὴ) and "healthy" (ὑγιής) (372e). Since the purpose of constructing the ideal city is to construct a just city, arguing that the City of Pigs is not a just city would be absurd. Therefore, Plato's definition of justice must be one that would judge the City of Pigs just. Since the lack of conflicts of external goods is the defining characteristic of it, justice at the level of the city, called objective justice, must be freedom from conflicts of external goods.

The need for the rule of reason, hence guardians, in resolving conflicts of external goods, comes when the city becomes complex enough for competing external goods, which Socrates lists off following 373a. Nevertheless, reason's purpose is still the resolution of conflicts of external goods, a conception that remains through Book iv. "Therefore, it is seemly for the logistic part to rule, it being wise and having forethought over the whole soul" (441e). Reason is valuable for being able to organize the other parts of the soul, but has, as yet, no intrinsic worth. On the other hand, epithymetic desires aren't presented as desires for things evil in themselves, but their fault is that they are insatiable and attempt to rule the other parts of the soul in order to attain pleasures of the body.[xiv]

Objective justice is the resolution of evils caused by conflicting external goods. It is tempting to try to interpret the remainder of the *Republic* based on this conception of justice. We know that different people have different talents. We could say the epithymetic, thymoeidic, and logistic constituencies in the city each have a different class of objects to pursue, and they achieve happiness by organizing their lives around the objects that they naturally strive after. Martha Nussbaum comments on this reading of the *Republic*:

> It appears that the conception of 'rule' by reason articulated in the fourth book of the dialogue is a purely formal conception that makes no attempt to specify the content of the life that gets planned and ordered by reason. All that is required there is that the agent harmonize his or her soul, order his or her life plan, in accordance with some orderly conception of the good...The difference between the reasoning or intellectual part and the other motivational elements of

the person is to be found in this capacity for overall evaluation and selection. Appetites merely reach out for objects, without conceiving of them as overall goods. But there is no reason, according to Book iv's formal conception of rational valuation, why the *content* of a life plan should not include appetitive activities as intrinsically valuable components that get selected and arranged alongside the others. (Nussbaum, 138)

However, we cannot interpret the entire *Republic* by this formal conception. Plato's invective against appetitive activities in the analysis of perverted cities and individuals in Books viii-ix is too strong for anyone to hold to this interpretation solely. However, if we are to solve the problem Williams raises, that the philosopher kings rule epithymetic people against their wills—ruling over a ticking time bomb—we shall need this interpretation.

I call this conception of justice objective because it judges, not from the perspective of the individual, but from the perspective of the city as a whole, which provides for basic needs and streamlines conflicting external goods. One of the principles of the foundation of the city is that different people have different abilities.[xv] The city is no more than the sum of its inhabitants. Some of them will be epithymetic individuals, and some thymoeidic. However, for the city to function, these people have to be recognized as valuable contributors to the city's welfare. A civil war between the rational class and the epithymetic class would preclude the possibility of a happy city. Therefore, if we place objective value in reason alone and equate reason only with the ruling class, an individual would be valuable to the city only to the degree that he or she is a philosopher, and this interpretation would entail that the epithymetic classes are to be lorded over by their rulers. Therefore, the epithymetic and thymoeidic classes must have objective value to the city. How then are we to make sense of Plato's denigration of desire in Books viii–ix?

By Book viii, a person would have a difficult time holding to the objective conception of justice described above. If Plato were to continue it, one would expect that he would treat the epithymetic part of the soul like epithymetic individuals in the

ideal city: individuals whose talents do not pertain to the governance of the city, but necessary individuals nevertheless. We certainly would not expect Plato to blame appetitive individuals for their disposition. However, when it comes to the considering the appetitive parts of a person's soul, Socrates vituperates them, and such treatment can hardly be interpreted other than as a warning to anyone who would give vent to them. Socrates disparages appetitive desires most of all in the analysis of the tyrant, claiming that some desires are lawless (παράνομοι) and present themselves "whenever the rest of the soul sleeps, the rational, civilized, and ruling part of it. . ." Socrates continues:

> . . . the beast-like and wild part, gorged full of food and drink, leaps up and casts off sleep and seeks to go out and to satisfy its instincts. You know that in such recklessness it does everything, as having been released and freed from all shame and prudence. For it doesn't hesitate to try to sleep with its mother, or any other thing of men, gods and beasts, to murder anyone, or not to keep away from any food: and in a word, it falls short in no way either of folly or of shamelessness. (571c)

The foregoing makes it plain that desires cannot be univocally called good. Certainly necessary desires exist,[xvi] but they are placed in the same class as the lawless, turpid aspects of human behavior. Socrates says: "What we wish to know is the following: that some terrible, wild, and lawless class of desires is in each person, even in some of us who seem to be altogether moderate. . ."(572b). The analysis has gone past the objective pronouncement that appetitive desires exist to denouncing some of them as evil in themselves.

The other conception of justice found in the *Republic*, which I call the subjective theory, considers justice from an individual perspective, in which certain actions are rewarded with pleasure and others are punished with pain. I call the individual perspective the personal moral struggle. Concerning subjective justice, Socrates is free to rail on the appetitive desires in Books viii-ix, and to oppose them with the superiority of reason. Plato depicts the moral struggle as being a struggle between two extremes:

reason and lawlessness, or the lack of reason. In the objective conception of justice, three types of people can be found in the city, and the difference between them is qualitative. On the other hand, Socrates' disparagement of appetitive desires accuses them of abandoning 'good sense' (φρονήσεως, or 'wisdom') and scaling the heights of 'unreason' (ἀνοίας). The individual is hardly given a choice between what paradigms of morality to conform to; there is only one morality: the choice between the inimical extremes of reason and lawlessness. There is no qualitative difference between the parts of the soul from the subjective perspective, but a continuity with two extremes. The epithymetic part of the soul is simply an irrational proclivity, and the thymoeidic part is a tendency somewhere between reason and lawless irrationality. In general, we can read the analysis of the perverted individuals in Books viii-ix as a steady progression away from reason to the disorder of desires.[xvii]

Objective justice treats the existence of epithymetic and thymoeidic individuals as positive fact. Though logistic individuals provide the most essential duties for the city (guardianship), all individuals are valuable to the city. However, Socrates' proofs that the philosopher is the happiest person (in the second half of Book ix), combined with his vituperation of appetitive desires (in the first half), would be meaningless if people were determined to be epithymetic or thymoeidic without any ability to choose otherwise. Rather, we must interpret Socrates to be exhorting every person to act according to reason, by praising reason and disparaging the opposite. According to objective justice, epithymetic individuals in the city are a fact of life. They simply exist. Subjective justice treats the desiring and spirited parts of the individual—the causes of epithymetic and thymoeidic citizens' characteristics—as things to be judged. Without two standards of justice, it would be absurd for Plato to execrate the cause (desiring and spirited parts of the person) but not the effect (desiring and spirited people).

Socrates' arguments for justice in Book i are arguments from the standpoint of a personal moral struggle. Let us consider some of Socrates' arguments to see if we can find latent the subjective conception of justice, with its inimical binary extremes. Socrates objects to defining justice as "helping your friends and harming your enemies" on the grounds that a person often mistakes friends for enemies and vice-versa (334c). The implication of this argument is that one can not be just without also having true

knowledge. The argument continues with Polemarchus, and Socrates concludes that the just person never harms anyone (335d). These arguments in Book i invoke a common feature of the Socratic dialogues: the Unity of Virtue. In short, a person cannot be just without being wise, without being good, nor without all other virtues. Furthermore, in the argument with Thrasymachus, Socrates argues that there can't be any particular goods (something which is good *for* something) that conflict with the absolute good.[xviii] Socrates' rationalism evaluates all things according to reason, and a person can be happy or just only to the extent that he or she is wise; as Terry Penner characterizes it, "the only good thing in itself is wisdom. . ." (Penner, 135). Socratic rationalism sets reason as the exhaustive dimension of human behavior, within which everyone acts. The subjective moral struggle in the *Republic* invokes this Socratic rationalism by placing reason and irrational desires at opposite extremes of a single-dimensional moral continuum.

The distinction between an objective and subjective conception of justice in Plato promised to heal two apparent contradictions. Let us now return to those problems to see if the distinction has resolved them. We found that we could read the text without contradiction only by saying that the city is just because all its inhabitants are just. First, we needed to understand the relationship between an individual's component parts in a different way than the relationship between the city's component parts, lest we fall into endless regression. Second, we needed to understand how an individual could be both epithymetic and δίκαιος at the same time.

Justice means different things in the city and in the individual. The beginning analogy, that justice in the city resembles justice in the individual, turns out meaningful because the wise element rules in both the city and the individual. In the city, however, the wise class is not ruling an unruly mob, but people who have arranged their lives according to reason, though perhaps not to the degree of the wise element of the city. The picture is much different with an individual. Reason and desires are enemies, unlike their objective relationship, and every individual uses reason to the degree that her or his life is not completely lawless, just as Socrates argues that even a band of criminals needs justice in order to carry out a joint crime: ". . . they pressed on towards crimes being half-wretched with injustice, since all-depraved and completely unjust people are also completely unable to act. . ." (352c). Therefore, though Socrates can say that the

form of justice is the same in the city as the individual, in that reason rules, the relationship between the elements is vastly different.

Second, we needed to see how a person could be epithymetic and δίκαιος at the same time. Objectively, individuals in the city are categorized according to three dispositions. It was one of the starting points for the city that people are diverse and have different abilities,[xix] and, from an objective standpoint, their dispositions are not a matter of choice. However, since a perfectly unjust person would be unable to act, all the people in the city have some degree of justice—and therefore, according to the standard of subjective justice, reason. The objective definition of an epithymetic individual, as outlined in Book iv, does not entail that she or he be governed by the lawless desires described in Book ix. The two depictions of desire pertain to different contexts, and it is wrong to interpret objective justice according to the subjective perspective.

Finally, let us examine how the objectively just city relates to an individual, whose experience of life is the subjective moral struggle. Certainly an epithymetic individual has lawless desires, but it would be wrong to argue that such a person can do nothing other than act upon them. Rather, there are different levels of the use of reason, from the epithymetic individual to the philosopher. Plato divides pleasures into two categories: need-relative pleasures and true pleasures.[xx] Certainly many epithymetic individuals experience need-relative pleasures, which are only apparent pleasures. Socrates uses the analogy of a man who is led up from some theoretical 'below' to a 'middle', who thinks he has gone above but in fact is only in the middle, whereas philosophers (among others, perhaps) experience the true above.[xxi] However, it would be wrong to conflate the epithymetic individual's pleasures exclusively with need-relative pleasures and to allot pure pleasures strictly to the philosopher. Both kinds of pleasures are the reward of the personal moral struggle.[xxii] In fact, Socrates asks a question about the kind of person who benefits from desires, for which the answer is the affirmative:

> May we be courageous and say that, even concerning the money-
> loving and victory-loving parts, as many desires as there are, they,
> following knowledge and reason and pursuing pleasures by means
> of them, may receive the pleasures which reason governs, and they

will receive the truest pleasures that it is possible to receive when it comes to true pleasures, and, inasmuch as their sort of pleasures follow the truth, they will receive the proper pleasures of their own sort, if indeed it is the best thing for each, and it is the most proper thing? (586d)

The individual described is hardly a philosopher; rather, he or she is the ideal epithymetic individual, whose desires follow reason and so gain maximal fulfillment. This statement ought to remind us of the Unity of Virtue, which the personal moral struggle entails: that no desires will attain fulfillment except to the degree they obey reason. Therefore, Socrates describes the terminus of the personal moral struggle as the point when one has escaped the constraint of competing external goods and may enjoy such goods without need-relativity. These are not the philosopher's pleasures; the philosopher certainly rules his or her desires with reason, yet the philosopher would rather forgo the pleasures of the epithymetic and thymoeidic parts in exchange for the better pleasures that come from knowing the Forms.[xxiii]

Will epithymetic individuals be happy in the ideal city? Plato's conclusion that justice is better than injustice can be sustained on the grounds that a just person is happier than an unjust person, without the need to make a pronouncement about absolute happiness. However, the ideal city cannot contain predominantly unhappy inhabitants, so the question is legitimate if we ask the negative: Are there unhappy people in the city? We know that some epithymetic people will not be unhappy, the same ones who attain their proper pleasures because they regulate their lives with reason. However, no doubt a city will always contain at least a few unruly people, and for them the rule of the philosopher kings will indeed be externally inflicted. Though this may sound oppressive, we have to ask what alternative would be better for such a person; Socrates affirms the philosopher king's rule over unwilling subjects "since it is better for all to be ruled by a wise and more-than-human ruler, most of all by having it private within oneself, and if not, by establishing it from without. . ." (590d). The picture looks bleak if unruly individuals have no hope of ever living willingly in the ideal city; however, we may assume that the ideal city will educate them for the subjective moral struggle

so that they may become willing citizens. This interpretation of Plato reads like John Dewey, whose optimism that education could eliminate any social ill was unshakable. In that vein, the plan to bring the ideal city into reality at the end of Book vii, heartless as it is, amounts to a massive educational program for its future inhabitants.[xxiv]

To conclude, we have seen that reading the *Republic* as containing two conceptions of justice, one objective and one subjective, resolves contradictions that arise from attempting to apply a single conception of justice to both the city and the individuals in it. It could be argued that this distinction damages the analogy between the city and the individual, since the city and the individual are apparently *not* just in the same way, as the analogy would have it; however, the distinction was necessary to preserve the literal reading of the *Republic*—that the work seeks to enlighten its reader concerning justice in both cities and individuals. I think that veracity remains with the analogy, since reason turns out to be basic in the governance of both individual and city. However, I must admit that the similarity is only superficial. Governance in the city is 'horizontal', between competing external goods, whereas the personal moral struggle is 'vertical', between reason and irrationality. Nevertheless if this difference is enough to doom the analogy, I think it would be a worthy sacrifice, for it certainly served its purpose, helping us understand justice in the individual.

Works Cited / Further Reading

Irwin, Terence. *Classical Thought.* Oxford: Oxford U. P., 1989.

Irwin, Terence. *Plato's Ethics.* New York: Oxford U. P., 1995.

Nussbaum, Martha. *The Fragility of Goodness.* Updated ed. Cambridge: Cambridge U.P, 2001.

Penner, Terry. "Socrates and the Early Dialogues" *The Cambridge Companion to Plato.* ed. Richard Kraut. Cambridge: Cambridge U.P., 1992.

Williams, Bernard. "The Analogy of City and Soul in Plato's Republic" *Plato 2.* ed. Gail Fine. Oxford U.P., 1999.

Endnotes

[i] 368e. δικαιοσύνη, φαμέν, ἔστι μὲν ἀνδρὸς ἑνός, ἔστι δέ που καὶ ὅλης πόλεως; (Do we say that there is on the one hand justice of one man, and there is on the other hand somehow justice of the whole city?) I have included the Greek where my argument depends particularly on the text. I omitted it in less

important instances and ancillary discussions, though I tried to err on the side of including too much.

ᵢᵢ 434e. Ὁ οὖν ἡμῖν ἐκεῖ ἐφάνη, ἐπαναφέρωμεν εἰς τὸν ἕνα, κἂν μὲν ὁμολογῆται, καλῶς ἕξει· ἐὰν δέ τι ἄλλο ἐν τῷ ἑνὶ ἐμφαίνηται, πάλιν ἐπανιόντες ἐπὶ τὴν πόλιν βασανιοῦμεν, καὶ τάχ' ἂν παρ' ἄλληλα σκοποῦντες καὶ τρίβοντες, ὥσπερ ἐκ πυρείων ἐκλάμψαι ποιήσαιμεν τὴν δικαιοσύνην· καὶ φανερὰν γενομένην βεβαιωσόμεθα αὐτὴν παρ' ἡμῖν αὐτοῖς. (Therefore, let us ascribe to the individual what appeared to us there, and if on the one hand it agree, it will be fine: But on the other hand if it seems something other in one thing, going back to the city we shall put it to the test, both examining forthwith and testing them alongside one another, so that we make justice flash forth, like from pieces of wood rubbed together to start a fire: And we shall make certain that it becomes evident to us ourselves.)

ᵢᵢᵢ 435a. ἆρ' οὖν, ἦν δ' ἐγώ, ὅ γε ταὐτὸν ἄν τις προσείποι μεῖζόν τε καὶ ἔλαττον, ἀνόμοιον τυγχάνει ὂν ταύτῃ ᾗ ταὐτὸν προσαγορεύεται, ἢ ὅμοιον; ὅμοιον, ἔφη. (Therefore, I say, is a thing which can be both greater and smaller, that someone would call the same term, dissimilar in the way for which it is called the same term, or similar? Similar, he said.)

ᵢᵥ Williams actually makes two arguments concerning the invalidity of the passage. The first argues that, by holding that characteristics of cities are caused by the same characteristics in their inhabitants, Plato accounts for justice in cities. Therefore, it would be pointless to look for a principle of justice in cities, since justice in cities can be reduced to inhabitants. I disagree. There obviously are such things as just and unjust cities. Even if justice in cities is caused by justice in individuals, it would be only wise to examine a city to see if justice is any more easily discerned there than in the individual, and then to apply the conclusions to justice in the individual.

ᵥ Cf. 420b.

ᵥᵢ See Irwin, Terence, *Plato's Ethics*, p. 192.

ᵥᵢᵢ Cf. 567c.

ᵥᵢᵢᵢ Cf. 546a.

ᵢₓ Cf. 547b. The dissolution of the complex city shows that if a city has guardians, a city cannot be just without just guardians. However, the City of Pigs (372d) has no guardians, so it is just without having just leaders. This argument depends on the City of Pigs being a just city, a conclusion I agree with. Socrates says, ἡ μὲν οὖν ἀληθινὴ πόλις δοκεῖ μοι εἶναι ἣν διεληλύθαμεν, ὥσπερ ὑγιής τις . . . (372e). ("It seems to me that the city which we completed is true and so a healthy one . . .") Since the purpose of constructing the ideal city was to construct a just city, the product must be a just city. However, the major implication of statement iii is not about the leaders of the City of Pigs, but all its citizens: the city would be just without any consideration whether its citizens are just. Proving or disproving whether statement iii applies to the City of Pigs depends on the answer to the question, "Are its citizens just?" If yes, the relationship is inconclusive because either statement i or iii could apply to the City of Pigs. If no, then statement iii is proven. I argue for the affirmative so that statement I can explain justice universally. I argue later that justice in the City of Pigs means freedom from conflicts of external goods; by this measure, its citizens are just. It is difficult to conceive of a definition of justice that would judge the City of Pigs just but its citizens not.

ₓ Plato describes the oligarchic man as being "not free of faction; not one, but a double sort of person, lusting after desires, so for the most part he would have better desires prevailing over the worse ones." (554d). Although numerous examples of the deplorable condition of the unjust person can be found in Books viii-ix, they are not all good examples of the point here, for I shall argue that Plato understands his or her desires as evil in themselves (κακὰς ἐπιθυμίας, (554d)). In Book ii, the debilitating condition is presented as the chaotic relationship between several goods, so that evil is not inherent in the goods or the desires themselves. In Books viii-ix, Plato identifies certain desires as evil in themselves, not accidentally.

[xi] See Irwin, Terence, *Classical Thought*, p. 104

[xii] Cf. 372d.

[xiii] Cf. 429a-b.

[xiv] Cf. 442a.

[xv] Cf. 370b.

[xvi] Cf. 558d.

[xvii] The first perversion of the ideal city, the transition from aristocracy to timocracy, happens when guardians yield to appetitive desires, when some guardians begin to move "toward profit and acquisition of land and house and gold and silver..." (547b). Though the things they seek may be better than lawless pleasures, the movement, nevertheless, is only one-dimensional: away from reason and toward appetites.

[xviii] Plato argues that the ruler's advantage cannot differ from the subject's advantage: "And I said, Thrasymachus, no one with any authority, as much a ruler as that person is, either considers or commands advantage for himself, but for the one ruled and for whom the ruler works." (342e).

[xix] Cf. 370b.

[xx] Cf. 584a.

[xxi] Cf. 584d.

[xxii] Although the text might seem to suggest that the only pure pleasures are pleasures from the wisdom-loving part (or cause) in the soul, Plato seems to hedge the question. In the quotation below, Socrates says that the victory- and profit-loving parts "will receive the truest pleasures," though he blunts the assertion by saying "that it is possible to receive."

[xxiii] For consistent interpretation, the philosopher's pleasures are objective pleasures, for they lie outside the moral struggle, which concerns both the pure attainment of external goods and the rejection of lawless desires as a single continuum of rational behavior. The philosopher's pleasures must lie in pure knowledge. If we understand the value of the philosopher's life to be the pure attainment of external goods, knowledge would be need-relative. Moreover, Plato would have innovated very little away from Socratic Rationalism. I would argue that a consistent interpretation of Plato requires that the Forms be purely good without prejudice to their instrumentality in attaining things in the realm of appearance. This interpretation is essential to Plato's argument for the immortality of the soul: οἷον δ᾽ ἐστὶν τῇ ἀληθείᾳ, οὐ λελωβημένον δεῖ αὐτὸ θεάσασθαι ὑπό τε τῆς τοῦ σώματος κοινωνίας καὶ ἄλλων κακῶν, ὥσπερ νῦν ἡμεῖς θεώμεθα, ἀλλ᾽ οἷόν ἐστιν καθαρὸν γιγνόμενον, τοιοῦτον ἱκανῶς λογισμῷ διαθεατέον, καὶ πολύ γε κάλλιον αὐτὸ εὑρήσει καὶ ἐναργέστερον δικαιοσύνας τε καὶ ἀδικίας διόψεται καὶ πάντα ἃ νῦν διήλθομεν (611c). ("To be able to see the truth, it is necessary to see (the soul) not mutilated by the association of the body and of other evils, as we now see it, but to see it become pure, a thing such as to be examined sufficiently by reason, and one shall find it far more beautiful and more visible, and one shall see justice and injustice clearly and all which we now passed through.")

[xxiv] Cf. 541a

Brandom's Account of Singular Terms*

David Friedell, University of California, Berkeley

Introduction

In *Making it Explicit*, Robert Brandom offers a theory of meaning that is a version of "inferentialism."[i] As its name suggests, inferentialism is the view that meaning ought to be explained in terms of inference. It is most likely a condition of adequacy for any theory of meaning, inferentialist or not, that it provide an account of singular terms. Brandom has his own reasons for including such an account (365–366), but this paper focuses not on these reasons but on the account itself—that is, how he explains what singular terms *are*.

A couple of preliminary points are in store: first, although it is commonly thought that singular terms are expressions that purport to refer to particular objects, Brandom thinks that this fact should be used to explain the notion of *an object* rather than the notion of *a singular term* (360). Second, he thinks that proper names (e.g., 'Aristotle') and definite descriptions (e.g., 'the tree outside my window') are both kinds of singular terms. In this way, he rejects a Russellian account of definite descriptions (and any other account on which definite descriptions are not singular terms).[ii] These are both controversial matters that I will not discuss at length. However, for the purposes of this paper, it is important to keep in mind that Brandom rejects a Russellian account of definite descriptions as well as the view that singular terms ought to be explained in terms of purported reference to particular objects. I will thereby grant him both of these points but leave open the possibility that they could be shown elsewhere to be incorrect, or correct.

The central aims of this paper are two-fold: exegesis (Part 1) and criticism (Part 2). I spend most of Part 1 explaining the notions that are needed to spell out Brandom's account of singular terms. I believe that the exegesis offered here is clearer than his own presentation, largely because I do not invoke what he calls "simple material substitution inferential commitments" or "SMSICs."[iii] Instead, I make heavy use of his notion of *material involvement* (sec. 1.3). I begin Part 2 by explaining some of the reasons why his account might seem plausible (sec. 2.1). Next, I continue this discussion by considering some futile objections (sec. 2.2). Lastly, I discuss an original objection and argue that it pressures him to reject a view that he seems to endorse: the view that definite descriptions always carry existential import (sec. 2.3).

Part 1

1.1 Substitution and the Syntactic Criterion

Brandom's account of singular terms includes a syntactic criterion and a semantic criterion. He takes them to be independently necessary and jointly sufficient for an expression to be a singular term. We begin with his notion of *substitution* and then use it to spell out the syntactic criterion.

For Brandom, any substitution involves three expressions with distinct substitutional roles: the component expression being "substituted for," the compound expression being "substituted in," and the shared expression that is a "substitution-frame" (368). For example:

> [1] <u>Socrates is bald.</u>
>   Aristotle is bald.

Here, 'Socrates' is *substituted for*; that is, we substitute another component expression *for* it ('Aristotle'). 'Socrates is bald' is *substituted in*; that is, the substitution of 'Aristotle' for 'Socrates' takes place *in* that sentence, which is a compound expression. Lastly, 'α is bald'[iv] is a *substitution-frame*: an incomplete expression, common to both lines of a substitution, consisting of a function with at least one argument place.[v]

One might hold, in opposition to Brandom, that substitution-frames can be substituted for. For instance, one might think that, in [2], the substitution-frame 'α is a philosopher' is substituted for the frame 'α is bald':

> [2] <u>Socrates is bald.</u>
>   Socrates is a philosopher.

However, Brandom claims that substitution-frames are not substituted for but are instead "replaced" (369). In support of this, he says: "A sentence frame is not a prior constituent of a sentence. . . but a product of analyzing it" (369). That is, substitution-frames are formed via "omission";[iv] in [2], 'α is bald' can be formed only by omitting 'Socrates'. Since 'α is bald' does not exist until we first omit another expression, we cannot omit it and insert 'α is a philosopher' in the same way that we can simply omit 'Socrates' and insert 'Aristotle'.

Now we can understand Brandom's syntactic criterion.

THE SYNTACTIC CRITERION:

An expression is a singular term only if it can be substituted for.

Predicates are taken to be substitution-frames (that can be replaced). We now look at the notions that are needed to spell out the semantic criterion. This will take us through section 1.4.

1.2 Substitution-inferences and Syntactic Kinds

In harmony with inferentialism, Brandom's semantic criterion relies on what he takes to be an inferential difference between singular terms and predicates. More specifically, he focuses on the roles of these expressions in "substitution-inferences." These are inferences that contain one premise and one conclusion, such that the conclusion results from an expression in the premise being substituted for, or replaced by, an expression of the *same syntactic kind* (370). Two expressions are of the same syntactic kind iff no well-formed sentence that contains one of them can be turned into something that is not a well-formed sentence by having the contained one substituted for, or replaced by, the other one (368). Brandom seems to hold that it follows from this that proper names and definite descriptions are of the same syntactic kind—and that all substitution-frames with one argument place (one place predicates) are of the same syntactic kind. I will assume this as well.[vii]

Examples [3] and [4], then, are both substitution-inferences:

[3] <u>Hitchcock is British</u>
The director of *Psycho* is British.

[4] <u>Frank is British</u>
Frank is European

[3]'s conclusion results from substituting the definite description 'the director of *Psycho*' for the proper name 'Hitchcock', and [4]'s conclusion results from replacing the

substitution-frame 'α is British' with 'α is European'. [5], however, is *not* a substitution-inference:

> [5] <u>Jim is running</u>
> It is not the case that Todd is running.

Although the conclusion results from replacing 'Jim' with 'it is not the case that Todd', these two expressions are not of the same syntactic kind. For, some well-formed sentences that contain one of these expressions can be changed into something that is not a well-formed sentence by replacing the contained expression with the other one; e.g., 'Sarah likes Jim' is a well-formed sentence, but 'Sarah likes it is not the case that Todd' is not.

1.3 Inferential Correctness and Material Involvement

Brandom's semantic criterion also appeals to a notion of *material involvement*, which, in turn, relies on a notion of *inferential correctness*. Thus, we will first look at the latter and then use it to spell out the former.

Although unclear on the matter, Brandom hints that the relevant notion of *inferential correctness* involves "status-preservation" (370). This sort of talk echoes Brandom's discussion of "scorekeeping practices" that are required for language and which involve two "deontic statuses": "commitment" and "entitlement." Brandom explains these statuses with a "phenomenal explanation" on which someone is committed or entitled in virtue of being taken and treated by scorekeepers as having such a status (166). Still, it helps to think of entitlement as being very similar to epistemic justification. Commitment cannot be as easily characterized. It is similar to belief, but only in a loose sense, according to which 'Jim believes the axioms of arithmetic' implies 'Jim believes that $114 + 51 = 165$'.[viii]

Brandom focuses on three types of inferential correctness that involve commitment and entitlement: "incompatibility entailment," "commitment preservation," and "entitlement preservation." An inference of the form $\mathbf{A} \mid - \mathbf{B}$ is *incompatibility-entailing* iff everything that is incompatible with $\mathbf{B}$ is incompatible with $\mathbf{A}$, where it is understood that two propositions are incompatible iff commitment to one "precludes" entitlement

to the other. An inference of the form **A | - B** is *commitment-preserving* iff it is the case that, if someone is committed to **A**, then she is committed to **B**. It is *entitlement-preserving* iff it is the case that, if someone is entitled to **A**, then she is entitled to **B** (as long as there are no incompatible commitments).

It will help to look at the following paradigmatic examples, chosen by Brandom:

[6] <u>Wulf is a mammal.</u>
Wulf is an animal.   (160)

[7] <u>Benjamin Franklin invented bifocals.</u>
The first postmaster general of the United States invented bifocals.   (370)

[8] <u>This is a dry, well-made match.</u>
It will light if struck.   (169)

[6] is incompatibility-entailing, commitment-preserving, and entitlement-preserving. [7] requires a more nuanced analysis: Brandom says of [7] that "in the appropriate context, commitment to the premise involves commitment to the conclusion" (370). The appropriate context, here, is one in which relevant scorekeepers take Benjamin Franklin to be the first postmaster general of the United States. In this context, [7] will count as commitment-preserving (as well as entitlement-preserving). [7] is not incompatibility-entailing, because some things are incompatible with the conclusion but compatible with the premise; e.g., 'The first postmaster general of the United States was Alexander Hamilton, who never invented anything'. [8] is entitlement-preserving but neither incompatibility-entailing nor commitment-preserving. For, there are cases in which a dry, well-made match will not light if struck; e.g., on the Moon. So some things are incompatible with the conclusion and compatible with the premise—and commitment to the premise does not entail commitment to the conclusion.

It is fair for us to assume that, for Brandom, an inference is correct iff it is incompatibility entailing, commitment-preserving, or entitlement-preserving—otherwise it is incorrect. Even if this is an inaccurate interpretation, it is adequate for our goal of understanding *material involvement*. Although Brandom briefly introduces this notion without clearly explaining it (370), the following definition attempts to flesh it out so that the notion of *a SMSIC* is not needed for our purposes:

THE DEFINITION OF MATERIAL INVOLVEMENT:

an expression p is materially involved in an inference A iff there is an expression of the same syntactic kind that can replace, or be substituted for, all instances of p, resulting in a new inference A`, such that the following two conditions are satisfied:

(i)   *either* A is correct and A` is incorrect *or* A is incorrect and A` is correct.

(ii) A` does not involve an intensional context that is not present in A.

It will help to look at some examples. Consider [3] again:

[3] Hitchcock is British.
     The director of *Psycho* is British.

Here, 'α is British' is not materially involved, because [3] is correct (commitment and entitlement-preserving), assuming that relevant scorekeepers take Hitchcock to be the director of *Psycho*, and there are no substitution-frames that can replace 'α is British' to produce an incorrect inference, except for frames that introduce an intensional context. For instance, if we replace 'α is British' with 'α is a good director' or 'α is mysterious' or 'α is overweight' or even 'it is not the case that α is American', the resulting inference is still correct. Accordingly, the following are both correct (commitment and entitlement-preserving) inferences:

[10] Hitchcock is a good director.
      The director of *Psycho* is a good director.

[11] It is not the case that Hitchcock is American.
      It is not the case that the director of *Psycho* is American.

The only way to produce an incorrect inference is to replace 'α is British' with a substitution-frame that introduces an intensional context; e.g., 'Sarah knows that α is a good director'. The resulting incorrect inference is:

[12] <u>Sarah knows that Hitchcock is a good director.</u>
Sarah knows that the director of *Psycho* is a good director.

Other ways of producing incorrect inferences rely on intensional substitution-frames like 'Sarah said that α is a good director' and 'Sarah believes that α is a good director'. But with condition (ii) in place, these kinds of frames are blocked from consideration, and we can safely say that 'α is British' is not materially involved in [3].

Conversely, 'Hitchcock' and 'the director of *Psycho*' are materially involved in [3], because if we substitute, say, 'Spielberg' for either of them (but not both), then the resulting inference is incorrect, and no intensional context is introduced—satisfying conditions (i) and (ii).

It is worth noting that one might think that condition (ii) is a circular way of dealing with a problem that intensionality poses for Brandom. On the one hand, omitting condition (ii) entails that 'α is British' counts as materially involved in [3]—something which Brandom surely does not want. On the other hand, including condition (ii) introduces a notion of *an intensional context*. And a common way of characterizing intensional contexts is to say something like the following: a sentence **A** has an intensional context iff an expression can be substituted for one in **A** that refers to, or applies to, the same object/s, resulting in a sentence **B**, such that the inference **A | - B** is incorrect.[ix] But this presupposes a notion of *an object*, which Brandom intends his account of singular terms to explain in the first place. He opens himself to a similar worry via his discussion of "primary occurrence" and "secondary occurrence" the occurrences of expressions in extensional and intensional contexts, respectively (373-374). He explains these two types of occurrence in terms of SMSICs but also appeals to them in order to explain SMSICs. In this way, Brandom deals with intensionality by having the notions of *a SMSIC* and *primary occurrence* work together to play the role of an unexplained explainer. In my presentation, the only difference is that a notion of *an intensional context* plays this role.

Let's look at another example to further clarify the notion of *material involvement*.

[13] <u>Joe is French.</u>
Joe is European.

'Joe' is not materially involved in [13], because [13] is correct, and we cannot make it incorrect by substituting for 'Joe' an expression of the same syntactic kind. Conversely, the substitution-frames 'α is French' and 'α is European' are both materially involved, because replacing either of them (but not both) with, say, 'α is three years old' results in an incorrect inference without introducing an intensional context—satisfying conditions (i) and (ii).

1.4 Symmetry, Separation, Substitution-pairs, and the Semantic Criterion

Brandom's semantic criterion also relies on a notion of *inferential symmetry*. An inference **A** |- **B** is symmetric iff it and **B** |- **A** are either both correct or both incorrect. Otherwise, it is asymmetric. For instance:

[14] *Jim is a son*
Jim is a male child

[15] *This is red.*
This is colored.

[14] is symmetric, for it is correct, and so is its converse: 'Jim is a male child |- Jim is a son'. [15] is *a*symmetric, for it is correct, but its converse ('This is colored |- This is red') is not.

There are two final notions that are needed to understand the semantic criterion: *separation* and *a substitution-pair* (my terminology). *Separation* is simple: an expression in an inference is separate from another in the same inference iff neither of them contains the other. For instance, in 'Hesperus is bright |- Phosphorous is bright', 'Hesperus' is separate from 'α is bright', 'Phosphorous', and 'Phosphorous is bright' but not from 'Hesperus is bright'. The notion of *a substitution-pair* is simple, too. A substitution-pair for a given substitution-inference is a pair of expressions, such that one is substituted for, or replaced by, the other. Thus, the substitution-pair for 'Clemens is a writer |- Twain is a writer' consists of 'Clemens' and 'Twain'.

We can now understand Brandom's semantic criterion for singular terms.

THE SEMANTIC CRITERION:

> An expression is a singular term only if all substitution-inferences that
> have a substitution-pair that consists of it and another expression,
> such that nothing separate from these two expressions is materially
> involved, are symmetric.

As stated in section 1.1, Brandom takes the syntactic and semantic criteria to be independently necessary and jointly sufficient for an expression to be a singular term. Thus, he takes an expression to be a singular term iff (a) it can be substituted for, and (b) all substitution-inferences that have a substitution-pair that consists of it and another expression, such that nothing separate from these two expressions is materially involved, are symmetric. A critique of this position is offered in Part 2.

Part 2

2.1 The Appeal of Brandom's Account

In Part 1, I presented Brandom's account of singular terms but excluded his notion of *a SMSIC*. In its place, I included (and fleshed out) his notion of *material involvement*. Some might prefer a Brandomian presentation that mentions SMSICs. Either way, the ensuing discussion can be modified to deal directly with such a presentation.

We will take Brandom's account to under-generate if it counts at least one singular term as not being a singular term. We will take it to over-generate if it counts at least one expression as a singular term that is not a singular term. I will try to show here why his account might appear to neither under-generate nor over-generate. It should be kept in mind that we are granting that proper names and definite descriptions are singular terms.

Let's look first at under-generation. It seems obvious that all singular terms can be substituted for. So we will suppose that they all satisfy the syntactic criterion. It is more difficult to see if they all satisfy the *semantic* criterion, but considering examples like [16]–[19] might make it seem as if they do.

[16] *Hesperus is bright.*
Phosphorous is bright.

[17] <u>Hesperus is beautiful.</u>
The tree outside my window is beautiful.

[18] <u>Hitchcock is mysterious.</u>
The director of *The Birds* is mysterious.

[19] <u>The director of Psycho is British.</u>
The director of *The Birds* is British

All of these inferences have a substitution-pair that consists of two singular terms, such that nothing separate from this pair is materially involved. [16]'s pair consists of 'Hesperus' and 'Phosphorous', and 'a is bright' is not materially involved; [17]'s pair consists of 'Hesperus' and 'the tree outside my window', and 'a is beautiful' is not materially involved; [18]'s pair consists of 'Hitchcock' and 'the director of *The Birds*', and 'a is mysterious' is not materially involved; [19]'s pair consists of 'the director of *Psycho*' and 'the director of *The Birds*', and 'a is British' is not materially involved. Moreover, all of the inferences are symmetric; [16] is correct both ways, as long as relevant scorekeepers take Hesperus to be Phosphorous; [17] is incorrect both ways, as long as relevant scorekeepers take Hesperus not to be the tree outside my window; [18] is correct both ways, as long as relevant scorekeepers take Hitchcock to be the director of *The Birds*; [19] is correct both ways, as long as relevant scorekeepers take the director of *Psycho* to be the director of *The Birds*. Thus, considering these examples (and other similar examples) might make it seem as though whenever a singular term is part of a substitution-pair for a substitution-inference, such that nothing separate from that pair is materially involved, the inference is *symmetric*—in accordance with the semantic criterion. Thus, it should be intuitive why it might seem that all singular terms (a) satisfy the syntactic and semantic criteria, and (b) therefore count as singular terms on Brandom's account. This issue will be revisited in sections 2.2 and 2.3, but for now we move onto over-generation.

In order for Brandom's account to over-generate, it would have to count as a singular term at least one expression that is not one. I will discuss here only predicates. His account rightly entails that all predicates are not singular terms. This is not only because predicates are taken to be substitution-frames that cannot be substituted

for, meaning that they fail the syntactic criterion, but also because they fail the *semantic* criterion. This becomes intuitive once we look at examples such as following:

[20] This is blue.
This is colored.

[21] Sally is tall.
Sally is extended.

[20] and [21] are both *a*symmetric, because they are both correct and both of their converses are incorrect. Moreover, [20] does not materially involve 'this', which is the only expression separate from the substitution-pair consisting of 'α is blue' and 'α is colored', whereas [21] does not materially involve 'Sally', which is the only expression separate from the pair consisting of 'α is tall' and 'α is extended'. Thus, these predicates fail the semantic criterion; for, according to it, an expression is a singular term only if all substitution-inferences that have a substitution-pair that consists of it and another expression, such that nothing separate from these two expressions is materially involved, are *symmetric*.

More importantly, for any predicate of the form 'α is F', we can invoke a predicate of the form 'α is F or G' in order to produce an *a*symmetric substitution-inference that (a) has a substitution-pair consisting of these two predicates, and (b) does not materially involve anything separate from them. For instance, following this strategy, we can produce [22], which shows that 'α is bald' is not a singular term on Brandom's account.

[22] Frank is bald.
Frank is bald or thin.

For these reasons, predicates do not count as singular terms on Brandom's account.[1]

2.2 Some Unsuccessful Objections to Brandom's Account

The following would be a counterexample to the semantic criterion: an *a*symmetric substitution-inference that has a substitution-pair consisting of at least one singular term, such that nothing separate from this pair is materially involved. I will now dis-

cuss two objections that attempt, but fail, to produce such a counterexample. This ought to further clarify why it might seem that Brandom's account does not under-generate. The first objection involves [23]:

> [23] Jim is tall.
>   An object is tall.

[23] is asymmetric; it is correct, and its converse is not. One might think that it is a substitution-inference with a substitution-pair consisting of a singular term ('Jim') and another expression ('an object') and that it does not materially involve any expression separate from this pair. But this is not so for two reasons. First, 'α is tall', which is separate from 'Jim' and 'an object', is materially involved; if we replace it with 'it is not the case that α is tall', then we get an incorrect inference: 'it is not the case that Jim is tall |- it is not the case that an object is tall'. Second, 'an object' and 'Jim' are not of the same syntactic kind, because 'I see an object of great importance' is a well-formed sentence, but 'I see Jim of great importance' is not. Thus, [23] is not even a substitution-inference. For these reasons, [23] fails to show that there is an asymmetric substitution-inference that (a) has a substitution-pair consisting of at least one singular term, and (b) does not materially involve anything separate from this pair.

We borrow our second fruitless objection from Brandom. He considers the following:

> [24] The skinniest person in the room can't fit through the narrowest door.
>   The fattest person in the room can't fit through the narrowest door. (388)

[24] is asymmetric; it is correct, and its converse is not. It has a substitution-pair consisting of two singular terms: 'the skinniest person in the room' and 'the fattest person in the room'. But it, too, poses no threat to Brandom. Speaking generally about examples like this one, he says:

> "These examples clearly turn on interactions between the predicates
> used to form definite descriptions and those involved in the sentence
> frame in which the description is embedded. Just by their nature,

such asymmetries do not generalize across sentence frames generally and so have no systematic significance of the sort appealed to in the substitutional account of the difference between singular terms and predicates" (388).

This point can be put more clearly by invoking material involvement: 'α can't fit through the narrowest door' is materially involved in [24], because [24] is correct and becomes incorrect if 'α can't fit through the narrowest door' is replaced with, say, 'α is kind'. Thus, since 'α can't fit through the narrowest door' is materially involved and is separate from both 'the skinniest person in the room' and 'the fattest person in the room', [24] fails to show that there is an asymmetric substitution-inference that has a substitution-pair consisting of at least one singular term, *such that nothing separate from this pair is materially involved*.

2.3 Another Objection to Brandom's Account

This section covers another attempt to produce the sort of counterexample discussed in 2.2. Unlike the previous attempts, however, this one poses a problem for Brandom. I will argue that it pressures him to reject the view that definite descriptions always carry existential import.

Suppose that there is a barn, henceforth referred to as 'the barn', and that all relevant scorekeepers are agnostic as to how many men the barn contains. That is, they have no idea how many men are in the barn. With this in mind, consider the following substitution-inference:

[25] The man in the barn is happy.
      The tallest man in the barn is happy.

[25] is a counterexample to Brandom's account if the following two conditions hold:

(i)   It is asymmetric.

(ii) It does not materially involve any expression separate from both

'the man in the barn' and 'the tallest man in the barn', which constitute a substitution-pair of singular terms.

Condition (i) holds; [25] is correct, but its converse is incorrect. 'The man in the barn is happy' entails 'there is only one man in the barn, and he is happy'.[xi] If there were only one man in the barn, then he would have to be the tallest (because there could be nobody taller or of equal height). Thus, 'the man in the barn is happy' entails 'the tallest man in the barn is happy'. But 'the tallest man in the barn is happy' does not entail 'the man in the barn is happy', because the former, unlike the latter, leaves open the possibility of there being more than one man in the barn. Thus, [25] is asymmetric, and condition (i) holds.[xii]

Although harder to tell, it might seem that condition (ii) holds. With little reflection, we can see that it holds unless 'a is happy' is materially involved—that is, unless 'a is happy' can be replaced by a non-intensional substitution-frame, resulting in an incorrect inference. Intuitively, [25] remains correct when 'a is happy' is replaced by most frames; e.g., 'a is sad', 'a laughs', etc. It might seem that the only exceptions are *intensional* frames, such as 'Jim said that a is sad' and 'Jim believes that a laughs'. If so, then 'a is happy' is not materially involved. Thus, it might seem that condition (ii) holds and that [25] is a counterexample to Brandom's account.

Brandom seems to overlook examples that are like [25]. Still, how he might respond can be gauged by looking at his argument that aims to show that all expressions that can be substituted for satisfy the semantic criterion (378–381). Roughly put, he supposes that there is an asymmetric substitution-inference **Pa** |- **Pb**, such that **Pa** |- **Pb** is correct and **Pb** |- **Pa** is incorrect. (**P** is a substitution-frame; **a** and **b** are expressions that can be substituted for.) He then attempts to show that **P** is materially involved, by showing that there are non-intensional substitution-frames that can replace **P** to produce an incorrect inference. He discusses two types of frames that fit this description: those that involve negation, and those that involve a conditional. For instance, if we replace **P** with ~**P**, then we get ~**Pa** |- ~**Pb**; and since **Pb** |- **Pa** is incorrect, this inference must also be incorrect, according to Brandom. Additionally, if we replace **P** with a frame that results in (**Pa** → **r**) |- (**Pb** → **r**), then this, too, must

be incorrect, according to him. Thus, Brandom is confident that, by invoking frames that involve either negation or a conditional, he can show that any frame is materially involved in an asymmetric substitution-inference with a substitution-pair consisting of singular terms.

Can Brandom employ this general strategy to show that 'α is happy' is materially involved in [25], thereby showing that condition (ii) does not hold? That is, can he show that 'α is happy' can be replaced by a non-intensional substitution-frame that involves negation or a conditional, resulting in an incorrect inference? Let's first consider negation and see what happens if 'α is happy' is replaced by 'it is not the case that α is happy', resulting in [25a]:

[25a] <u>It is not the case that the man in the barn is happy.</u>
It is not the case that the tallest man in the barn is happy.

My intuition is that [25a] is correct; I believe that its premise entails 'there is only one man in the barn, and he is not happy'. And this undeniably entails the conclusion of [25a]. Still, Brandom might argue that [25a] is incorrect. This can be so only if its premise entails not 'there is only one man in the barn, and he is not happy' but merely the following disjunction:

D: *Either* there is only one man in the barn, and he is not happy *or* it is not the case that there is only one man in the barn.

There are many possible scenarios in which (a) there is more than one man in the barn (satisfying D's second disjunct, thereby making D true), and (b) there is a tallest man in the barn, and he is happy (making the conclusion of [25a] false). Thus, if the premise of [25a] entails merely D, then [25a] is incorrect.[xiii]

In order for Brandom to hold that the premise of [25a] entails not 'there is only man in the barn, and he is not happy' but merely D, he must reject the following principle:

EI: All definite descriptions carry existential import. In other words, a proposition that contains a definite description of the form 'the F' entails 'the F exists'.

What I have said about 'it is not the case that α is happy' is true for other re-placements of 'α is happy' in [25] that involve negation; e.g., 'it is not the case that α is shy' and 'it is not the case that α runs'. In each case the resulting inference is incorrect only if its premise, of the form 'it is not the case that the man in the barn is F', does not entail a proposition of the form 'there is only one man in the barn, and he is not F'. And this is so only if EI is wrong.

Having covered negation, let us consider replacing 'α is happy' with a substitu-tion-frame that contains a conditional. Replacing it with 'if α is happy, then it is sunny' results in [25b]:

[25b] <u>If the man in the barn is happy, then it is sunny.</u>
       If the tallest man in the barn is happy, then it is sunny.

Similar to [25a], whether or not [25b] is incorrect boils down to whether or not its prem-ise entails 'there is only one man in the barn'. If so, then it is correct. If not, then it is in-correct. And, in order for Brandom to maintain that the premise of [25b] does not entail that there is only man in the barn—and thus that [25b] is incorrect—he must reject EI.

What I have said about 'if α is happy, then it is sunny' is true for other replace-ments of 'α is happy' in [25] that involve a conditional; e.g., 'if α is rich, then the stock mar-ket is doing well' and 'if α is healthy, then he exercises'. In each case, the resulting inference will be incorrect only if its premise, of the form 'if the man in the barn is G, then p', does not entail 'there is only one man in the barn'. And this can be so only if EI is wrong.

It should be noted that it would be telling for Brandom to reject EI, because, in chapter 7 of *Making it Explicit*, his discussion of Frege suggests that he endorses it. Although inexplicit, he seems to agree with Frege that being committed to a proposition that contains a definite description of the form 'the F' entails two commitments: (1) an existential commitment—i.e., a commitment to there existing an F, and (2) a uniqueness commitment—i.e., a commitment to there being only one F (432-434). It seems, there-fore, that Brandom actually endorses EI.

To bring my argument all together: [25] is a counterexample to Brandom's account, unless 'α is happy' is materially involved in [25]. This is so only if there is a non-intensional substitution-frame that can replace 'α is happy' to make an incorrect

inference. There is a frame of this sort that involves negation or a conditional only if EI is wrong. Thus, assuming that the only possible candidates are frames that involve negation or a conditional (and I see no other possibility), Brandom can maintain that [25] is not a counterexample only if he rejects EI (which he seems to endorse).

At this point, I will introduce two examples that are similar to [25]:

[26] *The February 29ᵗʰ of the rainiest year in history is a Monday.*
The day before March 1ˢᵗ of the rainiest year in history is a Monday.

[27] *The 100ᵗʰ person in line is funny.*
The person in line whose numerical position is closer to 100 than is the numerical position of any other person in line is funny.

These are a little more complex than [25], but they serve the same general purpose. The main idea with [26] is that it is asymmetric, if relevant scorekeepers have no idea whether or not the rainiest year in history is a leap-year. Its premise entails its conclusion, simply because any February 29ᵗʰ of any year must also be the day before March 1ˢᵗ of that year. But theconclusion does not entail the premise, because if the rainiest year in history is not a leap-year, then the day before March 1ˢᵗ of that year is a February 28ᵗʰ. Likewise, [27] is asymmetric, if relevant scorekeepers have no idea how many people are in line. Its premise entails its conclusion, because its premise entails that there is a 100ᵗʰ person in line—and if that were so, then that person's numerical position would be closer to 100 than would be the numerical position of anyone else. But the conclusion does not entail the premise, because if there were, say, only 40 people in line, then the 40ᵗʰ person would have a numerical position that is closer to 100 than is the numerical position of anyone else. Brandom might respond to [26] and [27] by replacing their respective substitution-frames with ones that employ either negation or a conditional. Nevertheless, unless he rejects EI, he must accept that [26] and [27] are counterexamples.

Conclusion

I have discussed Brandom's account of singular terms as involving the claim that an expression is a singular term iff (a) it can be substituted for, and (b) all substitution-inferences that have a substitution-pair that consists of it and another expression, such

that nothing separate from these two expressions is materially involved, are symmetric. I have outlined some of the reasons why his account might appear to neither over nor under-generate, covered some futile objections, and raised an objection that pressures him to reject EI: the view that definite descriptions always carry existential import.

I will close by briefly sketching another available response for Brandom. Rather than rejecting EI, he could bite the bullet and accept that [25], [26], and [27] are counterexamples to his account. In doing so, he could allow for a special class of singular terms that would include expressions such as 'the tallest man in the barn', 'the February 29th of the rainiest year in history', and 'the 100th person in line'. These differ from "unspecial" definite descriptions such as 'the man in the barn' and 'the tree outside my window' in that they appeal to a kind of ordering, whether it be height-ordering, the ordering of the days on a calendar, or a numerical kind of ordering. While conceding that this special class of definite descriptions does not cohere with his account, perhaps Brandom could revise his account so that all other singular terms do.[xiv]

Works Cited / Further Reading

Brandom, Robert. *Making it Explicit*. Cambridge: Harvard U. P., 1994

Dummett, Michael. *Frege: Philosophy of Language*, 2nd ed. Cambridge: Harvard U. P., 1981

Russell, Bertrand. "On Denoting" *Mind*. 14, 1905, 479–493.

Soames, Scott., *Philosophical Analysis in the Twentieth Century, Volume 1: The Dawn of Analysis*. Princeton U.P., 2003.

Endnotes

[i] Unless stated otherwise, all citations in this paper are from Robert Brandom. *Making it Explicit*.

[ii] Bertrand Russell thinks that definite descriptions are not singular terms but quantifiers. See Russell, "On Denoting".

[iii] For Brandom's discussion of SMSICs see pp. 370–376.

[iv] As a matter of convention, Brandom represents substitution-frames with the help of Greek letters such as 'α', and so will I in this paper.

ᵛ It might help the reader to consider how Frege characterizes "sentence-frames." See Dummett, (Chapter 2).

ᵛⁱ I borrow this term from Dummett.

ᵛⁱⁱ However, perhaps proper names and definite descriptions are not of the same syntactic kind. Thanks to John MacFarlane for suggesting that 'the celebrated astronaut John Glenn ran for president' is a well-formed sentence and that 'the celebrated astronaut the first man to orbit the earth ran for president' is not. If so, then 'John Glenn' and 'the first man to orbit the earth' are not of the same syntactic kind, even though they are codenoting! Since this matter bears not on what Brandom takes to be the essence of a singular term but merely on whether or not all singular terms are of the same syntactic kind, I will for simplicity's sake set aside this worry and assume, as I believe that Brandom does, that definite descriptions and singular terms are of the same syntactic kind.

ᵛⁱⁱⁱ It might also help the reader to consider that the fundamental thing that triggers sanctions in Brandom's scorekeeping model is being committed to a proposition without being entitled to that proposition (178–180).

ⁱˣ A similar characterization of intensionality is offered in Soames (*Philosophical Analysis*, 357). However, Soames' characterization is in terms of sentential truth rather than inferential correctness.

ˣ It remains unclear why sentences do not count as singular terms on Brandom's account. This becomes salient when one considers that (a) Brandom acknowledges that sentences can be substituted for in embedded contexts (369), and (b) he argues that all expressions that can be substituted for satisfy the semantic criterion (see section 2.3 of this paper). He suggests an answer when he claims that singular terms are "essentially subsentential," which likely means that singular terms can have semantic content only as sentence-parts (400). This distinguishes singular terms from sentences, which can have semantic content on their own—when, in Brandom's words, their utterance "has the significance of performing a speech act. . ." (367). Perhaps, then, his account actually has three criteria: the syntactic criterion, the semantic criterion, and what can be called 'the essentially subsentential criterion'. However, since I do not adequately understand Brandom's position on this matter, I have set aside this issue. Let it suffice that the ultimate explanation of why sentences do not count as singular terms on Brandom's account, assuming that there is such an explanation, has no bearing on the central objection of this paper, offered in section 2.3.

ˣⁱ For this discussion, a statement of the form 'A entails B' means that **A** |- **B** is correct.

ˣⁱⁱ Note that [25] is still asymmetric if we suppose that the relevant scorekeepers take there to be more than one man in the barn. What is important is that we do not suppose that they take there to be only one man in the barn, because then they would take the tallest man to *be* the only man—and [25] would be symmetric.

ˣⁱⁱⁱ Russellians would allow that merely D is entailed, for they think that 'it is not the case that the man in barn is happy' is ambiguous due to a kind of scope-ambiguity. It can mean either 'Not: [the x: x is a man in the barn](x is happy)' or '[the x: x is a man in the barn](Not: x is happy)'. This explanation seems unavailable to Brandom, given that he takes definite descriptions to be singular terms.

ˣⁱᵛ This response was suggested to me by Brandom himself with one important difference: he suggested that 'the tallest man in the barn' be excluded from this special class of singular terms, as he felt that it was more problematic than 'the February 29ᵗʰ of the rainiest year in history' and 'the 100ᵗʰ person in line'. It seems plausible to me, however, to group 'the tallest man in the barn' with the other two on the basis that all three appeal to some kind of ordering. Perhaps Brandom's hesitance is related to the fact that, while [26] and [27] rely on facts about the calendar and facts about our system of numbers, [25] relies merely on the meaning of the word 'tallest'.

Intensional Transitive Verbs

Will Lanier, Cornell University

A transitive verb is a verb that usually takes a direct object.[i] An *intensional* transitive verb (ITV) is both transitive and displays one or more of the following three peculiarities: (i) substitution problems, (ii) relational/notional readings, or (iii) existence-neutrality. In this essay, I will explain Graeme Forbes' recent account of intensional transitive verbs, why such an account is necessary, and whether Forbes' account has plausible alternatives.[ii]

We frequently use ITVs in everyday speech. The following table shows examples of ITVs:

Category	Example of verbs
Desire verbs	want, desire
Search verbs	search (for), looking (for)
Depiction verbs	draw, sculpt, imagine
Evaluative verbs	respect, admire, disdain
Emotion verbs	lust (after), fear[iii]

Each of the three hallmarks of ITVs—substitution problems, relational and notional readings, and existence-neutrality—has its distinct difficulties. Substitution problems are well known in their connection with propositional attitudes. In the direct object of most verbs, we can interchange co-referential names without affecting the truth-value of the sentence. For example, if I hit Superman, then I hit Clark Kent. Yet if I am looking for Superman, it is not necessarily true that I am looking for Clark Kent.

The second hallmark of ITVs, relational and notional readings, borrows terminology from Quine's "Quantifiers and Propositional Attitudes". ITVs, like propositional attitudes, often admit of two different readings—the relational and the notional. 'I'm looking for someone', for example, could be read two different ways. It could mean that there is a particular someone, and I am looking for that someone. Or it could simply mean that I am seeking relief from loneliness; there is no one particular person I am looking for. Quine called the former reading relational, and the latter reading notional.[iv]

The third hallmark of ITVs, existence-neutrality, often accompanies relational/notional ambiguities. If I say that I'm hitting the Loch Ness monster, the Loch Ness mon-

ster must exist for my sentence to be true. In general, the truth of sentences with transitive verbs requires the existence of their direct object. Intensional transitive verbs, however, do not. I might fear the monsters under my bed even if no monsters actually exist.

ITVs pose problems for accounts of language because the simplest compositional accounts do not account for their peculiarities. Thus, philosophers attempt to expand their semantics to explain ITVs' features. One acute problem is the second hallmark of ITVs, the relational/notional reading. In propositional attitudes, this feature is often explained away as a scope ambiguity—the *de re/de dicto* ambiguity.' For example, take the sentence 'Lois Lane believes that Clark Kent can fly.' Assuming, of course, that Lois Lane, and Clark Kent/Superman are real people about whom true and false things can be said, this sentence can be true or false. On a relational reading, I might be saying something like 'Clark Kent is such that Lois Lane believes he can fly'. In this sense, the sentence would be true, and the names Clark Kent and Superman are interchangeable. On a notional reading, I might be saying something like 'Lois Lane believes the thought expressed by the following sentence. Clark Kent can fly.' In this reading, the sentence is false. Lois Lane does not think that the nerdy Daily Planet reporter with glasses can fly.

While appealing to a scope ambiguity appears to help with propositional attitudes, it alone does not solve the problem with ITVs. The main difficulty is that notional readings of ITVs with a quantified noun phrase (QNP) as a direct object are not well-formed under the formation-rules of first-order language. For example, suppose I am looking for a sloop. Relationally, this glosses to the following sentence: there exists a sloop such that I am looking for it, or more formally, $\exists x$ (IsASloop (x) & LookingFor (I, x)). Notionally, however, the QNP is an argument of the function represented by 'looking for'. The following attempted formalization of this reading is not allowed in first-order language: LookingFor $(I, \exists x(\text{IsASloop}(x))$. Seen from another light, the problem is that we have an existential quantifier, yet our sentence is neutral concerning the existence of what we are looking for. In the literature, there are two main responses to this problem. First, following Quine, we can attempt to paraphrase the verbs so that QNPs are no longer arguments of functions. Second, following Montague, we can abandon first-order accounts in favor of a higher-order syntax/semantics that allows such a syntactical analysis and explains its semantic value.

Quine attempts to solve the notional problem of quantifiers by paraphrasing the ITVs so that the QNP is no longer an argument of the function. For example, 'I'm looking for someone' would gloss to 'I'm endeavoring to find someone'. Further refinements by Parsons and Larson replace this with 'I'm looking in order to make true the proposition that someone is such that I find it'.[vi] Forbes questions whether these and other paraphrases preserve the meanings of the verbs; nevertheless, he admits that Quinean solutions may account for some ITVs. His motivation behind supporting the second, higher-order solution is that he believes Quinean paraphrases do not work for *all* ITVs. Thus, he focuses on the most problematic ITVs for paraphrasing, puts forth a higher-order account for those, then expands his account to cover all ITVs, arguing that one must either (i) accept a piecemeal approach where some ITVs have a Quinean paraphrase and others a Forbesian analysis or (ii) adopt a higher-order semantic analysis for all ITVs. I will focus on the problematic ITVs for paraphrasing, explaining Forbes' motivation for appealing to Montague's semantics and avoiding the issue of choosing between (i) or (ii) for all ITVs.

Paraphrasing becomes more problematic with depiction and emotion verbs. Consider the following sentence: Kirstin imagines a unicorn. The most plausible paraphrase is Parsons' "Hamlet ellipses": Kirsten imagines a unicorn to be. Forbes objects that Parsons' account faces problems with negation. Compare the following sentences:

(i)   Kirsten does not imagine a unicorn
(ii) Kirsten does not imagine a unicorn to be

I'm not sure what (ii) would mean, but it seems to have different readings than (i). Kirsten could imagine a unicorn, without imagining it to be—say she is responding to the command, "imagine in your head a unicorn but don't pretend that it exists.' Even if responding to this command is impossible, and (ii) implies (i), this would presumably be due to mental facts about imagining. A transposition of this account to other depiction verbs creates difficulties. Compare Kirsten drawing a unicorn and Kirsten drawing a unicorn to be. To me, 'Kirsten drawing a unicorn to be' suggests a fairy tale where everything Kirsten draws comes to life. Perhaps a case could be made for the

equivalence of the two expressions. The correct analysis of depiction ITVs, however, seems unclear, perhaps due to difficulties in imagining and depicting themselves. Regardless, Forbes' best example of a failure of paraphrasing is not depiction ITVs but emotion ITVs.

The notional reading of 'Joe fears something' seems very resistant to paraphrasing. Various proposals do not preserve meaning. 'Joe fears encountering x' does not necessarily imply 'Joe fears x', and similarly for the converse. Perhaps Joe fears Bill's communicable disease, and thus fears encountering him. Nevertheless, Joe would not fear Bill himself. Likewise, 'Joe fears that x will do something undesirable to him' is not the same as 'Joe fears x'. He could simply fear the undesirable action, but not the agent of that action. In general, the problem with paraphrasing emotion verbs is that the emotion might be caused by the extra material added in the paraphrasing, and not by the direct object of the non-paraphrased sentence. Perhaps there are as yet unthought-of paraphrases that surmount this problem. Nevertheless, such difficulties suggest considering Forbes' alternative, even if Forbes does not conclusively shown its necessity.

In "The Proper Treatment of Quantification in Ordinary English", Montague proposes a higher-order intensional type theory syntax/semantics that Forbes extends and supplements with Davidson's event semantics to account for ITVs. Montague's syntax/semantics is a type theory—each expression is of a certain type, and the type of an expression explains what syntactical role it can play and what kind of semantic value it has. Here's a recursive definition of kinds of types:

1. There are expressions of type i, whose semantic values are individuals

2. There are expressions of type b, whose semantic values are boolean constants (i.e., true or false)

3. For any type x and y, there is an expression of type '$x \rightarrow y$', i.e., an expression whose semantic value is a function that maps the semantic values of expressions of type x onto the semantic values of expressions of type y.

The simplest Montagovian account would proceed as follows: names are type i, and sentences are type b. The semantic value of a name is an individual; the semantic value of a sentence is a truth-value. Predicates are type $i \to b$. In other words, 'is hunting sloops' takes an individual (denoted by an expression of type i) and maps it onto a truth-value (denoted by an expression of type b). So 'John is hunting sloops' would look like this: IsHuntingSloops(John).

Montague's account diverges from this simple account when he considers quantification. Quantified noun phrases (e.g., 'a sloop') are of type $(i \to b) \to b$ (or of type $(ib)b$ if we abbreviate the arrows). They denote a second-order function that takes a first order function (denoted by a predicate, an $i \to b$) and returns a truth-value (denoted by a sentence, a type b). For various reasons, Montague wants names to be of the same type as QNPs. So 'John' is not an expression of type i, and does not denote an individual. Through an operation[vii], 'John' denotes instead a second-order function that takes a first-order function, perhaps the function denoted by the predicate 'is hunting for sloops', and returns a truth value. In other words, 'John' becomes type $(ib)b$ just like QNPs such as 'a sloop'. Thus, 'John is hunting sloops' is really analyzed as the following: John(IsHuntingSloops). Furthermore, since QNPs/names are of type $(ib)b$ and predicates are of type ib (again, abbreviating the arrows), the verb 'hunts' has a complicated type. In 'John hunts a sloop', the predicate 'hunts a sloop' is composed of a verb, 'hunts', and a QNP, 'a sloop'. The entire predicate must be of type ib, and the QNP is of type $(ib)b$, so the verb must be of type $((ib)b) \to (ib)$. In other words, the verb takes an $(ib)b$, the QNP, to produce a predicate, type ib. In sum, Montague analyzes the notional reading of 'John hunts a sloop' in the following manner: 'John' is of type $(ib)b$. The function denoted by 'John' takes the function denoted by the predicate ('hunts a sloop', of type ib), and returns a truth-value denoted by the sentence 'John is hunting sloops'. The predicate is composed of a verb ('hunts') that denotes a third-order function which takes the second-order function denoted by the QNP ('a sloop', of type $(ib)b$) and returns the first-order function denoted by the entire predicate.

Forbes objects that transitive verbs shouldn't be this type. He finds it counterintuitive that ITVs and other verbs denote third-order functions, and claims that any philosophical argument to the contrary is "just stipulative" (*Attitude Problems*, 84). Neverthe-

less, Forbes himself suggests a motivation to adopt Montagovian syntax/semantics—the apparent failure of Quinean paraphrases. Also, Forbes gives some expressions unintuitive types that seem required once further philosophical analysis is done, e.g., 'and' is originally of type $b(bb)$ but often "type-shifts" to type $(ib)((ib)(ib))$ or type $((ib)b)(ib)$ (*Ibid*, 28). In fact, Forbes' later account of search verbs is almost identical to Mark Richard's theory—an attempt to explain Montagovian semantics for ITVs that I will discuss later in the paper. Forbes' borrowing here weakens his complaint that any attempt to justify ITVs as third order functions is "a philosopher's gloss", i.e., not what the layperson would give when asked what he or she meant by the expression. Thus, I do not see a problem with $((ib)b)(ib)$ semantics for ITVs, if the only objection is that they are counter-intuitive. Another of Forbes' objections seems to be that the Montagovian account is incomplete—it doesn't explain why notional readings of sentences with ITVs are existence-neutral concerning their direct object. In my reading of Montague, however, the existence-neutrality of ITVs follows from Montague's claim that ITVs take the intension of the direct object, instead of its extension. If an object has a non-empty intension, it does not follow that the object has a non-empty extension in this world. Thus, the existence of the direct object would not follow from the truth of a sentence in which it falls under the scope of an ITV. More work would have to be done, however, to distinguish between terms that require finer-grained distinctions than intensions, e.g., 'unicorn' and 'gorgon', as they both have the same intension—their extension is the empty set in all possible worlds, according to Kripke's standardly accepted account in *Naming and Necessity*. As I will show below, other philosophers (namely Richard) agree with this incompleteness of Montague, but think you can work within the Montagovian framework to explain this.

Forbes' final objection to Montague's semantics is the following:

> This brings a more fundamental problem into view. . . [Montague's notional analysis of a sentence containing an ITV] has a term for a property of properties as input to [the ITV], so its gist is not transparent. Where NP (noun phrase) is singular, we understand '*x* seeks NP' to mean that *x* is in the seeking relation to the individual to whom

NP refers. We cannot understand ['seeks' with $((ib)b)(ib)$ semantics]
in any different way, given that 'seek' is univocal with singular and
quantified NP-compliments. But to understand 'x seeks QNP (quanti-
fied noun phrase)' in this way is to have x seeking the meaning of a
quantifier (perhaps by looking it up in a dictionary). Montague as-
signed type (ib)b to singular terms as well, but matters are hardly clar-
ified by *extending* the scope of the puzzling property-of-properties
semantics. (*Ibid*, 71)

Forbes falls back on this objection as "the main problem with the Montagovian ap-
proach" (*Ibid*, 82) when his other objections appear weak. Yet, I think that Forbes'
objection fails once one considers other aspects of Montague's account of ITVs. Forbes
starts with an analysis of 'x seeks NP' taking names to be of type i, ignoring Montague's
lambda abstraction that converts them to second-order properties of type *(ib)b*—the
same type as QNPs. Once we focus on Montague's actual semantics for names, Forbes'
problem disappears. If names also are of type *(ib)b*, then the Montagovian analysis of 'x
seeks QNP' as ascribing a second order-property (the subject) to a first-order property
(the predicate) is exactly the same analysis given to 'x seeks NP'. At most, Forbes' argu-
ment shows that QNPs and NPs should be given the same semantics—providing further
motivation for Montague's *(ib)b* treatments of all NPs. Forbes' only argument against
considering Montague's *(ib)b* semantics for names is that "matters are hardly clarified by
extending the scope of the puzzling property-of-properties semantics." Yet "properties-
of-properties", i.e., second-order properties, are not puzzling in and of themselves. A
second-order property is merely a property of a first-order property, which is a property
of an object. There is nothing puzzling about 'a sloop' being a property held by any
property true of at least one sloop. Similarly, there is nothing puzzling about converting
the name 'John' that denotes an individual into a type *(ib)b* expression denoting the
second-order function true of all first-order functions true of John.

Nevertheless, while Montague's semantics allow for a notional reading of
quantified noun phrases without appealing to paraphrases, Montague gives no ac-
count of what relation each ITV is supposed to pick out, i.e., how the extension of an

ITV is determined, or the satisfaction conditions for sentences involving ITVs. Richard attempts to fill in the gaps while still using *((ib)b)(ib)* type ITVs. Forbes gives a similar account but abandons this analysis of ITVs.

In "Seeking a Centaur, Adoring Adonis: Intensional Transitives and Empty Terms", Richard attempts to flesh out a Montagovian treatment of ITVs. Focusing mainly on search verbs, Richard argues that an attribution of a searching makes a claim about the agent of the search's goals. Thinking along the lines of paraphrasing 'hunt' as 'bringing about to find', Richard claims that the agent's intentions are integral to the satisfaction conditions of the sentence. Richard creates "success structures" (ss's) that are similar to Montagovian intensions. The ss's are sets of ordered pairs <w, s> where w is a possible world and s is the extension of the goals of the search, i.e., everything that would constitute a successful search. A search sentence of the form '*x* searches for *y*' is true iff there is a search *e* such that *x* is the agent of *e* and such that every *ss* <*w,s*> of *e* is found in the extension of the intension of *y* at world *w*. After tweaking various formulations of satisfaction conditions for search verbs, Richard then addresses the issue of empty terms that have vacuous intensions, e.g., unicorn. If John is hunting for unicorns, Richard appeals to John's personal idiolect and beliefs, i.e., his own conceptions of what unicorns are (for example, a type of animal, distinct from trolls, etc.) to differentiate between 'John hunts for unicorns' and 'John hunts for trolls'—two huntings whose direct objects have the same vacuous intensions.

In contrast with Richard, Forbes does not use Montague's *((ib)b)(ib)* semantics for ITVs. Like Richard, he supplements Montagovian semantics with Davidson's event semantics[viii], although he favors slightly different conditions than Richard for what count as successful searches. Under Forbes' and Richard's framework, ITVs are predicates of events: a searching, a wanting, etc. In *Events in the Semantics of English*, Parsons combines Davidson's event semantics with the notion of semantic roles. For Parsons, each event has various roles, and an attribution of an event is a conjunction of various roles. For example, a biting is an event with a "biter" and a "bitee". 'The snake bit me' glosses to 'There was an event *e* and *e* was a biting, and the snake was the biter and I was the bitee'. Applying this to our Jack example, the relational reading of 'Jack wants a Lexus' is 'For some event *e*, *e* is a wanting and Jack is the "wanter" of *e* and for some Lexus *x*,

x is the "wantee" of *e*. One of the benefits of Davidsonian semantics that Forbes lauds is its elimination of a need for a special ITV logic. You can derive the existence of a Lexus by applying &-Elimination to the analysis of the relational reading, separating the existentially quantified phrase at the end.

To explain the notional reading of ITVs, Forbes argues that the quantified noun phrase characterizes the search, instead of providing a semantic role for the analysis. 'Jack wants a Lexus—but no particular one' roughly glosses to 'For some event *e*, *e* is an a-Lexus-characterized wanting, and Jack is the wanter of *e*'.[ix] The direct object cannot be a semantic role, because then a simple application of &-Elimination would enable you to derive the DO's existence from the analysis.

Forbes attempts to explain the characterization of ITVs by appealing to other quantified, hyphenated phrases that adjectivally describe a noun, e.g., a 'three-story building' or 'two-man bobsled'. In both phrases, the hyphenated adjective describes the following noun. Also, 'two-man bobsled' is intensional in the sense that a two-man bobsled does not require the existence of two men, just as a Lexus-characterized wanting does not require the existence of a Lexus. Forbes then describes the characterized phrase using common sense, arguing that any further explanation would be an illegitimate "philosopher's gloss". For example, a two-man-characterized painting is simply a painting depicting two men. The immediate problem, of course, is that this seems obviously circular. 'A painting depicting two men' needs to be able to be read notionally. Otherwise, it implies that there are two men such that they are depicted by a painting. Yet then the correct analysis of 'a painting depicting two men' is 'a two-men-characterized painting'. But 'a painting depicting two men' was supposed to explain the characterized-relationship, not rely on it. In light of this worry, Forbes resorts to the same basic, Richardian framework of defining characteristics of a search in terms of the agent's intensions and conditions for a successful search. Forbes then concludes that each type of ITV will require its own corresponding *ad hoc* account of satisfaction conditions, and he delegates this job to the "interpreter" (*Ibid*).

Forbes' account thus provides an alternative to Quinean paraphrasing and Richard's theory of how Montagovian semantics might explain notional readings. To its advantage, Forbes' account easily explains troublesome ITVs for Quine, e.g.,'fears'.

Along with these advantages, however, come the numerous complications of Forbes' account in explaining all the various peculiarities of ITVs. Furthermore, it is unclear whether Forbes presents any independent reasons to support Montagovian syntax/ semantics or Davidsonian event semantics. A complete evaluation of Forbes' theory would have to weigh the complexity and assumptions of these frameworks he relies on. Also, Forbes' case against Richard's usage of Montague strikes me as weak. Since Forbes ends up borrowing successful search conditions for search verbs, and similar philosophical glosses of his own to explain the semantics of his "characterize" relation, perhaps Forbes' account should be seen as an alternative rather than a superior theory to Richard.

Works Cited / Further Reading

Chierchia, Gennaro and Sally McConnell-Ginet. *Meaning and Grammar: An Introduction to Semantics*, 2nd ed. Cambridge: MIT Press, 2000.

Davidson, Donald. "The Logical Form of Action Sentences" *Essays on Actions and Events* Oxford: Clarendon Press, 2001.

Forbes, Graeme. *Attitude Problems: An Essay on Linguistic Intensionality*. Oxford: Clarendon Press, 2006.

Forbes, Graeme. "Intensional Transitive Verbs" *The Stanford Encyclopedia of Philosophy*, Winter 2004, ed. Edward N. Zalta, http://plato.stanford.edu/archives/win2004/entries/intensional-trans-verbs/.

Kripke, Saul. *Naming and Necessity*. Cambridge: Harvard U.P., 1980.

Larson, Richard K. "The Grammar of Intensionality" *Logical Form and Language*. ed. G. Preyer and G. Peter Oxford: Clarendon Press, 2002.

Montague, Richard. "The Proper Treatment of Quantification in Ordinary English" *Formal Philosophy*, New Haven: Yale U. P., 1973, 247–270.

Parsons, Terrence. *Events in the Semantics of English: A Study of Subatomic Semantics*. Cambridge: MIT Press, 1990.

Richard, Mark. "Seeking a Centaur, Adoring Adonis: Intensional Transitives and Empty Terms" *Midwest Studies in Philosophy* 25.1, 2001, 103–127.

Quine, Willard, "Quantifiers and Propositional Attitudes" *Journal of Philosophy* 53, 1956, 177–187.

Endnotes

[1] As defined by Forbes, Graeme. 'Intensional Transitive Verbs', *Stanford Encyclopedia of Philosophy*. This is the definition Forbes uses, so I follow his lead in the paper. I have also heard a transitive verb defined as any verb that can take a direct object.

[2] Forbes, Graeme. *Attitude Problems: An Essay on Linguistic Intensionality*.

[iii] This is a concise summary of the most common examples in the literature and this essay. Forbes provides a full 'taxonomy' of over fifty ITVs in *Attitude Problems* p. 37.

[iv] Forbes argues for a third, "non-committal" reading when either a relational or notional reading captures the meaning of the expression (*Ibid*, 47–51). Forbes' example concerns a painting of a dog. Suppose two museum curators disagree whether the painting was of a specific dog in particular. One curator thinks that the painter painted an actual dog; the other curator thinks that the painter had no specific dog in mind when he painted the painting. The first curator with the "actual dog" thesis agrees with the relational reading of 'The painter painted a dog'. The second curator agrees with the notional reading of 'The painter painted a dog'. Forbes claims that there is a proposition they agree on—a third, non-committal reading of 'The painter painted a dog'. Such a reading would be non-committal in that it takes no stance on whether or not the painter painted a particular dog, i.e., it makes no commitment between a relational or notional reading. The existence or non-existence of such a reading has little bearing on this essay, so I refrain from discussing it.

[v] Some people define the dr/dd ambiguity solely as a scope ambiguity. Others use it as synonymous with Quine's relational/notional distinction. With such a definition, of course, the dr/dd distinction might not be a scope ambiguity

[vi] Quine, "Quantifiers and Propositional Attitudes". Larson, "The Grammar of Intensionality".

[vii] lambda abstraction—see Forbes, *Attitude Problems*, pp. 23–26; Chierchia & McConnell-Ginet, *Meaning And Grammer*, ch. 7

[viii] Davidson, Donald. "The Logical Form of Action Sentences" *Essays on Actions and Events*, pp.105–148

[ix] It is '*e* is an a-Lexus-characterized wanting' and not '*e* is a Lexus-characterized-wanting' because Forbes wants quantifiers in the characterized relationship to distinguish between, e.g., a two-dog-characterized painting and a three-dog-characterized painting.

(Un)Doing Critical Philosophy: Reflections on Adorno's *Aesthetic Theory*

Larry McGrath, University of California, Berkeley

An Introduction to Adorno's Thought and a Clarification of Terminology

Theodor Adorno (1903–1969) advanced critical theory as a form of historically self-reflexive and sociological critique. As a mode of philosophical-sociological inquiry, critical theory takes aim at the ideological, scientific, and economic conditions of post-industrial capitalism. These conditions reflect the development of refined instruments of social control that prolong and intensify mass unfreedom and absorb avenues of resistance, such as those originally articulated by Marx.

My interpretation of Adorno's thought thus foregrounds its Marxist elements. These elements coalesce in the task of critical theory: to imagine the world differently by subjecting it to critical negation. As I argue in this essay, *Aesthetic Theory* advances an understanding of art against the backdrop of a world in which this task becomes increasingly difficult to fulfill. *Aesthetic Theory* thus integrates the principles of critical theory with an account of art in its contemporary context. What is art? How does art affect us in the modern era? And, most significantly, how should philosophy engage art? These questions guide *Aesthetic Theory* in its effort to galvanize the power of philosophy and art in an era that blunts the critical potential of both.

Terms:

1. Decline of metaphysics ("a post-metaphysical world"): refers to an ontology of becoming, which this paper credits to Friedrich Nietzsche's assault on metaphysical certainty. This view holds that there exists no abstract (Platonic) realm of stability beyond the physical world; instead, the world is composed of struggle among competing forces. Thus, physics best describes our world, as one in constant motion, reducible only to the activity (rather than substances, objects, or things) of which it is composed.

2. Instrumental rationality: can be understood as a mode of thought that converts the inherent value of something into its use-value. This is made possible by identifying objects and their variety of relations with a conceptual system. Money, for ex-

ample, is purely instrumental within a monetary system; its only value is its ability to purchase other items, as no bill is worth anything beside the instrumental use to which it is put. For Adorno, instrumentalization is the dominant mode of thought in late modernity.

3. Violence of the concept: refers to the subsumption of all objects, and what is inherent of them, under the universal rules constitutive of concepts. This is an epistemological counterpart to instrumental rationality, which imposes the concept of identity[i] on things to render them know-able. We come to know society and nature through identity-thinking when all that is real must harmonize with our own conceptual system.

4. Negative dialectics: attempts to limit the dominance of identity-thinking by reflecting on the limits of our concepts. It therefore aims to think what is un-thought. The task is to negate the identity between object and concept, and therefore think of what falls outside identity. This is a reflective, rather than projective exercise, whereby we question and explore the conditions of out thought. For Adorno, negative dialectics is the central task and driving thrust of critical theory.

The beginning moments of *Aesthetic Theory* offer the author's reflections on the contemporary status of art, that "nothing concerning art is self-evident anymore, not its inner life, nor its relation to the world, not even its right to exist" (*Ibid*, 1). My project takes Adorno's reflections as its point of departure. I will begin by identifying the socio-historical conditions that efface art's self-evidence: a post-metaphysical world circumscribed by instrumental rationality. It is in response to these two conditions that Adorno crafts his negatively dialectical aesthetics. The question I am interested in asking is how philosophy responds to these same conditions. In order to answer this question I probe the style of the philosophical text. My analysis will uncover in *Aesthetic Theory* a negatively dialectical style of philosophical construction that grounds its coherence in the philosophical conditions of a post-metaphysical world and resists its instrumentalization by the demands of instrumental reason. This will unfold as I trace the development of the central theses of *Aesthetic Theory*, analyzing its style, which opposes the traditional logical and narrative form of philosophical texts.

I.

Historically, philosophy has built theories of aesthetics upon an ambivalent relationship between their constituents, philosophy and art. On the one hand, art, in its production and reception, is often understood to be inherently subjective. The subject's aesthetic experience of its engagement with an artwork seems to be confined to the particularity of that experience. On the other hand, philosophy aims to conceptualize what is universal in that experience. Adorno recognizes "the fundamental difficulty, indeed impossibility, of gaining general access to art by means of a system of philosophical categories." At the same time "aesthetic statements [such as "this work is beautiful"] have traditionally presupposed theories of knowledge" (*Ibid*, 332). Thus, aesthetics must negotiate these twin dimensions, the philosophical demand to articulate universal categories and the particularity of the work of art. This duality motivates a dialectical aesthetic that mediates between the philosophical concept and the work's resistance to conceptual consumption.

Adorno's dialectic mediation between these oppositional dimensions owes its groundwork to the seminal aesthetics of Kant and Hegel. Kant's contribution to the aesthetic tradition is his transcendental critique of aesthetic judgment. In the *Critique of Judgment* judgment is an inherent faculty of knowledge; it functions as "the ability to think the particular as contained under the universal" (*Ibid,* 18). Aesthetic judgment, in particular, grounds itself in the *a priori* faculty of taste, confirmed by "the fact that whenever we judge any beauty at all we seek the standard for it *a priori* in ourselves, and that the aesthetic power of judgment itself legislates concerning the judgment as to whether something is beautiful or not" (*Ibid,* 225). Thus, the subject who deems a work of art beautiful does so on the basis of the universal faculty that endows the work with its aesthetic objectivity. Hegel's' philosophy of art responds to Kant's transhistorical account of aesthetic judgment. For Hegel, art has a history of its own. It's understanding is not guided by universal categories; instead, spirit speaks through individual works within their historical era. In its opposition to the Kantian primacy of the subject, whereby the concept subsumes the particular work of art, Hegel's aesthetics reveals the work's own cognitive comportment.

Drawing upon these contributions, Adorno crafts an aesthetics that builds upon the autonomous historicity of the work of art, yet recognizes that the subject who engages the work cannot dispose of its conceptual apparatus. The work is a product of history, but comprehension of its uniqueness depends upon its subjective mediation. *Aesthetic Theory* orients this subjective mediation in new directions attuned to the intrinsic temporality of the work. He writes of aesthetics, "as an investigative procedure subsumption never reveals aesthetic content, but if subsumption is rejected altogether, no content would be thinkable" (*Ibid,* 18). Hence, what is necessary is an aesthetics that curbs the conceptual subsumption of the work in an effort to reflectively yield to its own objectivity. Dialectics, in its movement between the work and the concept, responds to this demand to preserve the particularity of the work in the face of the violence done to it by the concept: "Aesthetics is not obliged, as under the spell of its object, to exorcise concepts. Rather, its responsibility is to free concepts from their exteriority to the particular object and to bring them within the work" (*Ibid,* 181).

This is a daunting project that Adorno takes up. The trick is to deploy philosophy successfully against its own medium—the concept. But before moving to how

Adorno succeeds in his task, we should note the socio-historical circumstances that an aesthetics must also address. These are the historical conditions philosophy now faces in a post-metaphysical world and the barbarity imposed on that world by the logics of late capitalism, consisting of the mass promulgation of unreflective and anti-critical modes of art by the culture industry. In the face of such conditions, aesthetics finds its task to "free concepts from their exteriority" more demanding.

The decline of metaphysics marks the rise of a world wherein the stable ground upon which to found an aesthetics dissolves. Aesthetics can no longer ground itself in the lofty Kantian position of a transcendental subject. This is because faculties of knowledge do not submit themselves to transhistorical investigation. Nor can Adorno work within the logics of the "end of history," in which Hegel's dialectic culminates. Theory must dispense with the search for a stable starting point from which investigation of the artwork proceeds. This has become the case following Nietzsche's dismantlement of the truth of metaphysics in his revelation that "the 'apparent' world is the only true one: the 'true' world is merely added by a lie" *(Ibid,* 481). Heidegger's reflections on Nietzsche illustrate the world philosophy must now address:

> [I]f the world were constantly changing and perishing, if it had its essence in the most perishable of what perishes and is inconstant, truth in the sense of what is constant and stable would be a mere fixation and coagulation of what in itself is becoming: measured against what is becoming, such fixation would be inappropriate and merely a distortion. . . A knowledge that—as true—takes something to be "being" in the sense of constant and stable restricts itself to beings, and yet does not get at the actual: the world as a becoming world. *(Ibid,* 64)

A world of becoming is one in which philosophy cannot content itself with the stability of the given. The sense of truth toward which philosophy strives undergoes a radical transformation. The correspondence between an object and its predicates, as held by the Cartesian tradition, gives way to a sense of truth that slips through the grasp of the postulates of systematizing philosophies. Historical contingency resists the thrust toward

universality which motivated the aesthetics of Adorno's predecessors. Adorno's project reorients this thrust in philosophy, as he writes, "The great philosophical aesthetics stood in concordance with art to the extent that they conceptualized what was universal in it; this was in accordance with a stage in which philosophy and other forms of spirit, such as art, had not yet been torn apart" (*Ibid*, 334). Aesthetics must now attune itself to a historical understanding of truth that grounds itself and the individual work. Both reside in the socio-historic conditions within which the subject engages the work of art.

Hence, "art" is neither a stable category nor a catalog of exemplary works. Aesthetics cannot begin with reflections on art, but must ground itself in the truth of the individual artwork. But this truth does not exhaust what appears internal within the work; its elements unfold in layers belonging to a historical world of becoming: "The artwork is a process essentially in the relation of its whole and parts. Without being reducible to one side or the other, it is the relation itself that is a process of becoming." This process does not pertain to the shifts in the work's reception through time, but the activity inherent to the work itself. Whatever we can label the "totality" of the work cannot be a "structure that integrates the sum of its parts" (*Ibid*, 178).

The immediate consequence for aesthetics is the crisis of art's self-identity. If the post-metaphysical worlds strips artworks of their idea, understood as their aesthetic verity resonating beyond time, aesthetics cannot aim to reach behind the work to capture its truth. Aesthetics must critically engage, and not blindly surrender itself to a fractured reality. The relationship between subject and work is not immediate, nor can philosophy penetrate what truths hold in this relation. Aesthetics must mediate the intertwinement of the work, society, and history. This mediation is necessary because no aesthetics grounded in a systematic conceptual apparatus can do justice to the demands of the individual work. If the world is becoming, then theories of aesthetics must relinquish their reliance upon systematicity; to understand the work is not the same as pumping it full of philosophical concepts. In short, form must give way to experience. Only the latter is equipped to address art in a world of becoming:

The exertion of cognition is predominantly the destruction of its usual exertion, of its using violence against the object. Knowledge of

the object is brought closer by the act of the subject rending the veil
it weaves about the object. It can do this only when, passive, without
anxiety, it entrusts itself to its own experience. In the places where
subjective reason senses subjective contingency, the primacy of the
object shimmers through; that in the object which is not a subjective
addition. (*Critical Models*, 254)

While the "subject is the agent," aesthetics cannot allow it to be "the constituent of
object." (*Ibid*, 254) The requirement that the subject rend "the veil it weaves about the
object" demands more than that the abdication of aesthetic systems. The subject, too,
finds itself interwoven within the forces of power and history. Thus a dialectical aes-
thetics performs the requisite task by submitting the reification of systematic thought
to critical revision. It must resist positivity in its construction of concepts, and dispense
with identity thinking which represents the object in the image of the universal con-
cept. The violence done by the concept compels Adorno to posit as a criterion of an
aesthetics' success, its capacity to draw from the artwork those elements that resist its
comprehension in the form of concepts.

This imperative to advance socio-historical critique heightens against the
backdrop of the universal instrumentalism late capitalism imposes upon life. Aesthetics
must not only pry itself loose from the rigidity of conceptual systems, but also save the
work of art from the spell of its commodification. I would like to suggest that this is the
central objective of *Aesthetic Theory*, and the work of the Frankfurt School generally:
resistance to the valueless fungibility we face in a world circumscribed by instrumental
rationality. The latter takes as its operating principle the reduction of all aspects of life,
including art, to their use value. Under these conditions, the work of art is reduced to a
unit of pure exchangeability, a commodity circulated in the market. As a result, Adorno
recognizes that "art no longer has a place" in our society. Under instrumental reason,
"art fragments on one hand into a reified, hardened cultural possession and on the
other into a source of pleasure that the consumer pockets and that for the most part has
little to do with the object itself" (*Aesthetic Theory*, 15).

The modern era marks the culmination of rational-Enlightenment thought,

whereby the empirical world succumbs to the Kantian transcendental subject—object becomes subject. As a consequence, "thought makes itself mere tautology" (*Dialectic of Enlightenment*, 20). Late capitalism embodies the socio-economic concretization of instrumental reason, which reifies consciousness as identity thinking. The market, through its rational-economic modes of thought, has seized the subject from the world, thereby sterilizing the subject's critical relationship to the world. The subject now unknowingly becomes the object of a world in which one-dimensional thought dominates. This makes the subject increasingly unable to perform the task Adorno demands of aesthetics, to subject reality to critical revision through the artwork, to see the world otherwise. Instead, everywhere the subject goes, it confronts only itself: "The man of science knows things to the extent that he can make them. Their "in-itself" becomes "for-him." In their transformation the essence of things is revealed as always the same, a substrate of domination" (*Ibid*, 20). The complete administration of society instrumentalizes all spheres of life to such a degree that knowledge hypostasizes into the operations of a machine. Experience has completely given itself away to form. Adorno takes these consequences as the central target to which dialectical thinking must respond, which lays the foundations of negatively dialectical thinking:

> To grasp existing things as such. . . to think of them as surface, as mediated conceptual moments which are only fulfilled by revealing their social, historical, and human meaning—this whole aspiration of knowledge is abandoned. Knowledge does not consist in mere perception, classification, and calculation but precisely in the determining negation of whatever is directly at hand. Instead of such negation, mathematical formalism, whose medium, number, is the most abstract form of the immediate, arrests thought at mere immediacy.
>
> (*Ibid*, 20)

Adorno, in collaboration with Max Horkheimer, thus crystallizes the project of a dialectic aesthetic theory. This notion of dialectic, however, differs from its original Hegelian and Marxist variations. Dialectics must be negative. This means it neither submits itself to the

positivity of idealism's synthesis (the sublation of thesis and antithesis), nor does it operate according to the objective laws of historical materialism. Instead, dialectic thought must dwell reflectively in the "determining negation of whatever is directly at hand."

It is from the objectives set forth by Adorno's critical project that I take this essay's investigative point of departure. *Aesthetic Theory* offers a response to a completely administered post-metaphysical world—it presents a reflective and dialectical aesthetics that dislodges the artwork from its social appropriation by instrumental rationality. The question I am posing is how does the philosophical work, in particular that of Adorno, respond? How does the philosophical text resist its appropriation by a reified world that strips objects of their inherent value, reduced to market commodities? I will argue that *Aesthetic Theory* provides an answer in its philosophical style, achieved though a constellational and negatively dialectical construction. Adorno does this by modeling his philosophy in aesthetic experience. This is not to say that *Aesthetic Theory* is an artwork, but that it integrates elements unique to art that preserve its critical capacity. These elements, however, cannot be brought out independent of the text's internal development of its ideas. *Aesthetic Theory* is not a work of literature that subjects itself to aesthetic analysis. The text offers a theory that unfolds through its manner of presentation, but the theory reflexively shapes the text's presentation. Thus, my investigation will begin by drawing from the text what it means for a philosophical work to resist its instrumentalization. I will first argue that the work must overcome its utility as a method, from which I will uncover how *Aesthetic Theory* performs this task.

II.

Artworks retain a critical dimension given their situation both within and outside the world. This is what distinguishes artworks from inert objects—their resistance to the world within an explanative context. Adorno writes, "Only by virtue of separation from empirical reality. . . does the artwork achieve a heightened order of existence" (*Aesthetic Theory*, 4). Great artists Adorno looks to, such as Rembrandt, Beckett, or Beethoven, were among those whose "sharpest sense of reality was joined with estrangement from reality" (*Ibid*, 9). Hence the necessity for a dialectical aesthetics arises

from what is artistic, and therefore critical, in the work. For aesthetics to tend to the work, it must preserve art's autonomy—its own negative participation within reality: "That art on the one hand confronts society autonomously, and, on the other hand, is itself social, defines the law of its experience" (*Ibid*, 348).

But how can aesthetics engage a work that is both within and outside reality, given the position of the theorist within society? The subject must understand the work as an antagonistic movement between its participation within and resistance to reality. If the work's own determination opposes the conceptual systems through which we identify reality as real, it cannot be understood as a self-enclosed identity. A world of becoming precludes the possibility of aesthetic comprehension that exhausts the work's meaning. Rather, "a relation, not identity, operates between the negativity of the metaphysical content and the eclipsing of the aesthetic content" (*Ibid*, 358). The question thus persists, a relation between what? The traditional aesthetic categories such as form-content or universal-particular cannot capture this relation; they must give way to an experiential aesthetics. No framework of binaries can capture the antagonistic *processes* in the work, as Adorno specifies:

> Whatever may in the artwork be called totality is not a structure that integrates the sum of its parts. Even objectified the work remains a developing process by virtue of the propensities active in it. Conversely, the parts are not something given, as which analysis almost inevitably mistakes them: Rather, they are centers of energy that strain toward the whole on the basis of a necessity that they equally perform. The vortex of this dialectic ultimately consumes the concept of meaning. (*Ibid*, 178)

That the parts of the work are "not something given" forecloses their capacity to ground an aesthetics. There exists nothing for the subject to grasp a hold of in the work, despite instrumental reason's claim to do so. Adorno instead couches the work's elements in a discourse of becoming: "developing process," "centers of energy," "vortex of dialectic." But what allows Adorno to assert such claims about the artwork? Could we not say that he has injected the work with a dialectic method of process, thereby sacrificing experience to form?

Quite the contrary, Adorno's account presupposes an aesthetic experience on the part of the reader without proactively pointing to a particular work. The philosophy we read does not take its object as a given starting point. It suspends reliance upon ground and thereby engages a world of becoming. All that Adorno can rely upon is the experience of artworks, and not solely what is determinate within them. Hence he turns to our experience, upon which he invites us to reflect. What Adorno points out is what he sees taking place as he experiences the work, as if to say, "there it is; do you see it too?" In short, *Aesthetic Theory* offers a philosophy of negatively dialectical reflection and not conceptual projection. This is what allows Adorno to posit the sorts of reflective identity statements we continuously find in the text: "The artwork is X, art does Y," and so forth.

This means that Adorno dispenses with the task of demonstration. He does not begin with the work's elements and move outward in order to identify the work any more than he injects the work with concepts external to it. The movement is not unidirectional, but mediates between the aesthetic insights concepts afford us, and the resistance to those concepts that we experience within works. This style is confirmed by, on the one hand, the position of the above passage within the text, and on the other, the examples of art the passage references. Adorno's conceptualization of artworks as processual movements rests between a preceding paragraph that spans three pages, a fragmented discussion of Mozart, Beethoven, "Stockhausen's concept of electronic works," and "Picasso's rayonism." Simultaneously, the experiences of these artworks inform Adorno's aesthetic insights. Mozart's compositions polemically distance themselves from a musical tradition that has refused their revolutionary form. What is artistic in Beethoven's *Appassionata* as well as modern electronic music perish by the nature of their temporal movement as works of art; for history affords works their life as well as their death. And finally, Picasso's pastiche of haute couture elements offers an art form as turbulent as fashion itself. In consequence, the passage is a brief moment of clarity caught within a discussion whose movement is suggestive of that inherent to works of art. The textual movement of the passage and the aesthetic movement constitutive of individual artworks both proceed by way of their antagonistic elements. Just as the artwork unifies its objectivity against the historical movement it inhabits, so to

does Adorno's text dwell in a similar friction, between the concepts contained within his *Aesthetic Theory* and the temporal objectivity of the artworks the text considers. Adorno's insight is not the conclusion of a logically developed argument but an instant of reflection that allows the antagonisms constitutive of the artwork to shine.

Adorno's style allows his dialectic aesthetics to reflect on an experience of the artwork. The text's construction is not a hierarchical presentation of concepts; it instead subjects itself to dialectic thought. Thus *Aesthetic Theory* models itself in aesthetic experience. Indeed, if the artwork is a movement of antagonisms, then aesthetics' response must attune itself to this processual experience. Aesthetics cannot rigidify itself anymore than the work of art can abdicate its internal movement. Thus art calls for an aesthetics that is dialectical, and thereby allows the subject to engage the work without appropriating it:

> To whoever remains strictly internal, art will not open its eyes, and whoever remains strictly external distorts artworks by a lack of affinity. Yet aesthetics becomes more than a rhapsodic back and forth between the two standpoints by developing their reciprocal mediation in the artwork itself. (*Ibid*, 350)

This means the subject's consciousness must "remain constantly mobile both internally and externally to the work" (*Ibid*, 350). We witness this mobility unfold through the stylistic strategies Adorno deploys in the development of his theory. The text's ideas linger, they do not explain away the meaning of either art or their theory. What makes the latter in fact dialectical is not just its capacity to reflect upon the particularity of the work in its non-identity; a dialectical aesthetics must maintain its coherence through its own dialectic mediation. And this occurs in the fragmentary style in which *Aesthetic Theory* envelops its aesthetics.

A dialectical aesthetics thus jettisons conceptual systems in order to afford the work its autonomy. This allows the work to exist apart from the world that struggles to pull it back. Instrumental reason endeavors to harden the work in order to fetishize its value as a commodity within the market. At the same time, a post-metaphysical world of becoming renders conceptual aesthetic methods untenable. Adorno recognizes, "the

tendency of philosophical aesthetics toward those abstract rules in which nothing is invariable. . . is transient; the claim to imperishability has become obsolete" (*Ibid*, 339). Yet what prevents the petrifaction of Adorno's dialectic aesthetics into "abstract rules"? In other words, how does an anti-systematic aesthetics, amid an instrumentalized world, ward off its conversion into method, thereby offering itself as a dispensable tool to the art critic?

This concern is eminent for Adorno, who warns, "the over-valuation of method is truly a symptom of the consciousness of our time. . . this tendency is related to the nature of the commodity: to the fact that everything is seen as functional, as a be-ing-for-another" (Goldmann, 129). Thus, dialectical aesthetics must turn its negation back on itself. That is, it must submit its conceptual content to reflection by way of artworks and their inherent resistance to conceptualization. Yet Richard Wolin suggests *Aesthetic Theory* fails in this task. Wolin asserts that Adorno's aesthetics "remained un-dialectically wedded to the concept of an esotericized, autonomous art as an absolute model of aesthetic value" (Wolin, 45). Consequently, *Aesthetic Theory* "runs the risk of a false sublation of autonomous art, whereby a crucial refuge of negativity and critique would be prematurely integrated with facticity as such" (*Ibid*, 45). Yet Wolin's claims that Adorno remains "undialectically wedded" to autonomous art undervalues the criti-cal comportment of such art, and moreover, Adorno's incorporation of similar elements into his own work. I would like to suggest that criticism such as Wolin's neglects this latter point, the critical style of *Aesthetic Theory*, which dialectically preserves its co-herence in the face of its instrumentalization. It is toward this insight that I direct the following section.

III.

Nowhere in *Aesthetic Theory* do we find a definition of what it means for an aesthetics to be negatively dialectical, at least not one that exhausts the multiplicity of its dimensions. What we find is instead a myriad of art-situations that call for critical thought. It is each work's truth that issues this call, a truth that emanates from the flow of its movement that draws our attention as well as our critical reflection. This feature secures the complexity of the text, its multiple layers and hypnotic affect in its series of

digressions and philosophical excursions. But rather than dispel the text of its power, these qualities ensure its successful resistance to its instrumentalization as a method.

Aesthetic Theory, in short, is enigmatic, a term Adorno uses to describe the artwork's autonomous position both within and outside society. But the artwork's autonomy, its internal antagonistic movement, does not reduce it to a unit of chaos. Rather, within its enigmaticalness the subject encounters the artwork's critical dimension. This is what allows Adorno to contend, "the idea of a conservative artwork is inherently absurd." This is because the artwork occupies a critical posture beyond the limits of its social inception, "By emphatically separating themselves from the empirical world, their other, they bear witness that that world itself should be other than it is; they are the unconscious schemata of that world's transformation" (*Aesthetic Theory*, 177).

How is it that the work's 'enigmaticalness', in its dynamic presence in and absence from this world, critically negates society? Adorno's answer is its truth-content, which the work possesses as its own cognitive capacity, what allows the work to remain an object not subsumable by the subject's concepts. The work's truth-content orients the movement of its internal parts, which depends upon philosophy for its self-actualization. Works aim toward the "determination of the indeterminate" in their resistance to reality, but in so doing they simultaneously pose a problem, that of their negative dimension. This is why we do not look at artworks and immediately think, "revolution!" Instead, the work "achieves meaning by forming its emphatic absence of meaning" (*Ibid*, 127). No interpretation will reveal the work "as a new immediacy" because the work's "enigmaticalness outlives the interpretation that arrives at the answer." (*Ibid*, 125). Thus it becomes the objective of a dialectical aesthetics to no longer "explain away the element of incomprehensibility" but instead "understand the incomprehensibility itself" (*Ibid*, 347).

In order to perform this task, the subject allows itself to be disciplined by the truth-content of the work. Thus it becomes the task of aesthetics to reflect on the work's truth: "By demanding its solution, the enigma points to its truth-content. It can only be achieved by philosophical reflection. This alone is the justification of aesthetics" (*Ibid*, 128). Philosophy becomes the midwife of the work's truth; it allows the work to instruct us how the world, beyond its identification with instrumental concepts, can be thought in a radically different way. But this only occurs when the subject allows the work to

speak for itself. Philosophy must dispense with its metaphysical quest for truth and submit to the truth of the work. Martin Jay clarifies, "Truth for Adorno was not. . . merely correspondence between propositions and an external referent in the current world, but rather a concept with normative resonances as well, referring to a future 'true' society" (Jay, 159). The question thus arises, how does a dialectical philosophy perform this task insofar as there exists nothing stable within the work to comprehend?

Philosophy draws forth from the work its truth: its artistic particularity that suspends and re-imagines our instrumentalizing identity relations that seize hold of the world. We do not encounter the work's truth as immediately present content, but instead, existent within the work's form and its relation to the historical conditions of its materialization. This relation, however, does not offer it's meaning conceptually. In response, philosophy takes up a reflective task that is experiential; it extracts from the work a moment of insight into our concepts' indeterminate other. This other resides in the blind spots of conceptualization: the experience of both what is outside the determinations of instrumentalizing reason, as well as the conditions of our concepts' emergence. For instrumental reason severs the experience of its own emergence from thought; its reification perceives the universal concept, and its corollary identity-relations, as given things unaltered by the historical world of becoming of which it is a part. The truth-content of the artwork elicits this reflective exercise; it provides the antipode of the concept toward which dialectic thought moves. As the subject lingers in the work's truth-content, it makes manifest the violence of the philosophical concept. A dialectical aesthetic experience paralyzes the subject-object distinction, which undergirds the process of conceptualization. In its paralysis, we do not take the distinction between subject and object for granted. The universal and subjective concept cannot penetrate what is lawful of the work's truth-content without acknowledging its inadequacy. This lawfulness is not determinate as something juridical, but constitutes the work's objectivity that commands the subject's obedience. The subject obeys with the aid of a dialectical aesthetics, and together submits our conceptual representation of reality to the possibility of its other. Philosophical reflection on the work's truth-content interweaves reified consciousness and reflective self-consciousness, thereby estranging the subject from its hypostatized relation to objects.

This experience re-orients the way in which the subject engages the object. Philosophy does not content itself with the grasping of concepts, but must experience the object's resistance to conceptualization. The following series of subsequent passages reflect the text's negatively dialectical movement—just as soon as Adorno's aesthetic concepts appear to reach their concrete fruition, the text returns to particular aesthetic experiences:

> Even by artworks the concrete is scarcely to be named other than negatively. It is only through the nonfungibility of its own existence and not through any special content that the artwork suspends empirical reality as an abstract and universal functional nexus. Each artwork is utopia insofar as through its form it anticipates what would finally be itself, and this converges with the demand for the abrogation of the spell of self-identity cast by the subject. No artwork cedes to another. (*Aesthetic Theory*, 135)

The objectivity of the artwork has its own agency, yet it acts through that which it is not. Only against the suffering of reality does the work revive its own singularity, its nonfungibility, and thereby direct its movement toward imagining the world otherwise. In *Aesthetic Theory*, philosophy responds by negatively presenting the contradictory movement of the work. The text does not explain away this movement, but participates in it. Indeed, the "artwork is utopia," yet simultaneously utopia is not entirely of the work; it is anticipated and dependent upon the subject's "abrogation of the spell of self-identity." The demand the artwork places on philosophy is simultaneously fulfilled by the dialectical style of the text. This movement continues through the passage:

> The nonfungibility, of course, takes over the function of strengthening the belief that mediation is not universal. But the artwork must absorb even its most fatal enemy—fungibility; rather than fleeing into concretion, the artwork must present through its own concretion the total nexus of abstraction and thereby resist it. (*Ibid*, 135)

The text shares a movement of resistance with the artwork it portrays. The artwork's extirpation of fungibility folds within the "nonfungibility" of its "own existence." Deep within the artwork's "nexus of abstraction" and resistance to it, dwells the call for a mode of thinking that pleasures before what lies beyond identity relations. *Aesthetic Theory* heeds this call as it unfolds through a "sequence of dialectical reversals and inversions" (Richter, 101). Much like the dialectical aesthetic experience, Adorno's philosophy resists its own positivity; it develops in a manner that is non-identical. The movement of the text halts before the full concretion of its aesthetic theory and returns to the experiences of art objects upon which it depends. We can see in the continuation of the passage:

> Repetition in authentic new artworks is not always an accommodation to the archaic compulsion toward repetition. Many artworks indite this compulsion and thereby take the part of. . . the unrepeatable; Beckett's *Play*, with the spurious infinity of its reprise, presents the most accomplished example. The black and grey of recent art, its asceticism against color, is the negative apotheosis of color. (*Ibid*, 135)

The works mentioned do not demonstrate Adorno's theory, for the theory does not develop apart from the aesthetic experiences of which it consists. The works mentioned belong to the movement of the text, which rides their aesthetic experience. This movement is confirmed in the ensuing moment:

> But because for art, utopia—the yet to exist—is draped in black, it remains in all its mediations recollection; recollection of the possible in opposition to the actual that suppresses it; it is the imaginary reparation of the catastrophe of world history; it is freedom, which under the spell of necessity did not—and may not ever—come to pass. Art's methexis in the tenebrous, its negativity, is implicit in its tense relation to permanent catastrophe. (*Ibid*, 135)

The textual movement of *Aesthetic Theory* slides from its theory of aesthetics to particular works and back to aesthetics in a seemingly fragmentary manner. It posits moments that do not develop the next. Much like the artwork, the text feels tenebrous in its negativity, yet resists its own catastrophe by its very negativity. The experiences of particular artworks, such as Beckett's *Play* or "black and grey" ascetic art, are elements in the development of Adorno's aesthetics: the aesthetics and individual works dialectically inter-relate. This experience grounds an aesthetic theory in our post-metaphysical world that has dismantled all stable foundations; yet the experience is incomplete, "draped in black," without a complementary aesthetics. Hence, the text moves in a non-deductive manner; it submits its own grounds, that of aesthetic experience, to dialectical reflection. The text reads like a musical composition, but not like a pop song or Jazz piece, both of which Adorno disdained for their rationalized predictability. The text's form does not dictate the ordering of its contents, but instead, like an atonal Schoenberg composition, it weaves through the friction of its own elements.

This strategy is made possible by the text's constellational style. The constellation ties together conceptual moments in such a way that resists the Kantian distinction between subject and object. Whereas for Kant, the universal categories of the understanding subsume particulars in order to render them knowable, in constellations particulars (the experience of art works) and universals (the concepts belonging to aesthetics) co-exist alongside each other. *Aesthetic Theory* is constellational in that its aesthetics retains its coherence as a theory, but not apart from the aesthetic experiences out of which it is composed. Susan Buck-Morss clarifies this epistemological strategy:

> Cognitive knowledge. . . was achieved by means of abstraction: the particular entered into the concept and disappeared. But in [constellations] the particulars, although conceptually mediated, reemerged in the idea. . . they *became* the idea in the conceptual arrangement of their elements. The role of the subject, to draw connections between the phenomenal elements, was not unlike that of the astrologer, who perceived figures in the heavens. (Buck-Morss, 92)

The constellations that comprise *Aesthetic Theory* retain aesthetic experiences as non-hierarchical monads in an inter-connected web. Each moment, like a single star of a constellation, contains the totality, its own picture of the world, yet remains distinct from the other moments. The concepts belonging to Adorno's aesthetics do not subsume particulars; the latter sustain their particularity in the former.

Adorno's mediation of artworks in *Aesthetic Theory* preserves non-identity, and in so doing unwinds particulars from their conceptual reification. This explains how the text does not offer an aesthetic method, but instead submits itself to the particularity of the artwork. Where instrumental reason congeals the world of contradictions into systems of identities, *Aesthetic Theory* unfolds through constellations that render visible what is antagonistic of reality. It therefore mobilizes art in resistance to the instrumentalizing dictates of the market, wherein commodities operate according to the principles of abstraction, identity, and reification, which Buck-Morss describes as the "ossification of the object as a mystifying fetish by splitting it off from the process of its production" (*Ibid*, 98). In contrast, dialectical aesthetics employs what instrumentality severs. The individual work only becomes comprehensible in relation to the historical whole to which its truth-content responds.

IV.

The truth of the artwork lies beyond reality in its estrangement from the society in which we receive it. But how does the work retain its negativity, its openness to a world that does not yet exist despite its existence within a reified world? It is in mimesis that art resists the empirical identification of its truth-content: "By pursuing its own identity with itself, art assimilates itself with the nonidentical: This is the contemporary stage of development of art's mimetic essence" (*Aesthetic Theory*, 134). The work's "mimetic essence" compels a reflective, and not rationalizing, response to the work. It makes possible the work's openness to a primal world, hypothetically existing prior to the subject's abstraction from nature through reason, and the subsequent concretization of its conceptual apparatus. Mimesis is what cannot be spoken in the work; it is a mystic openness to a world not translatable into language. Hence

Robert Kaufman describes mimesis as that which is "grasped not as transcription but as an attempt provisionally to know something of the otherness outside the subject" (Kaufman, 201). It is by way of mimesis that the subject comes in contact with the otherness beyond its concepts. This is why mimesis, as Frederic Jameson clarifies, "can be said often to function as a more adequate substitute for the primal relationship of subject and object." What is primal of mimesis "forestalls dualistic thinking by naming the dualism as such" (Jameson, 105). Thus, mimesis forces the concept back on itself; the concept recognizes that it cannot ascertain what is mimetic in the artwork. Rather than appropriate the work, the concept must approach the work by way of dialectical reflection—it must experience that which cannot be named using the subject's tools of reason.

The nature of mimesis is one of non-identity, which occupies a world that is both primal and of the future. On the one hand, mimesis belongs to a world that exists prior to the ascendance of the bourgeois subject and its domination of nature. Adorno writes, "Art is imitation exclusively as the imitation of an objective expression, remote from psychology, of which the sensorium was perhaps once conscious in the world and which now subsists only in artworks" (*Aesthetic Theory*, 112). In its mimetic dimension, the artwork offers a glimpse into how subjective consciousness has become alienated from nature. Yet on the other hand, mimesis opens a more free world not yet realized, beyond the subjective logic of identity:

> Only the autonomous self is able to turn critically against itself and break through its illusory imprisonment. This is not conceivable as long as the mimetic element is repressed by a rigid aesthetic super-ego rather than the mimetic element disappears into and is maintained in the objectivation of the tension between itself and its antithesis. (*Ibid*, 117)

The artwork's mimetic comportment, which Adorno goes on to describe as "the plenipotentiary of an undamaged life in the midst of mutilated life," is not recuperated in a nostalgic past, but imitates a world not yet actualized within our own. Thus, mimesis

assumes different contexts, both of the past and future, throughout the text. *Aesthetic Theory* does not advance an extractable notion of mimesis; its account, instead, can be said to unfold mimetically. That is to say mimesis works against the constraints of language. Our understanding of mimesis develops in fragments, each of which unfolds within the contours of a particular constellational moment.

The text's fragmented account of mimesis illustrates its constellational structure and movement. In certain moments, mimesis is that which is remembered in the artwork; at other monuments it speaks to the work's effort to imagine a future world. These accounts are, of course, not mutually exclusive, as Adorno states, "The trace of memory is mimesis, which every artwork seeks, is simultaneously always the anticipation of a condition beyond the diremption of the individual and the collective" (*Ibid*, 131). In it's fragmented and divergent moments, the text charges mimesis with different social, historical, and aesthetic valences, but never does so in a univocal manner.

I would like to suggest that *Aesthetic Theory* contains a mimetic dimension that allows it to resist its instrumentalization. The mimetic element of the artwork allows its truth-content to speak for itself, and thereby counters the violence wrought by instrumentalizing concepts' claim to speak for it. It occupies a purely experiential world that escapes its petrifaction in language. Yet we cannot say the text is mimetic insofar as its medium is language, which necessarily obfuscates what is intrinsically unspeakable in mimesis: "By virtue of its double character, language is a constituent of art and its mortal enemy." Adorno goes on to clarify that "compared to significative language" the expression of mimesis "is older though unfulfilled" (*Ibid*, 112).

Mimesis, what affords the artwork its dimension of negativity, resists its inclusion within philosophy. Yet philosophy depends upon a mimetic dimension in order to preserve the primacy of the object apart from its conceptualization. Philosophy thus confronts a paradox: in order to resist its instrumental concretization it must not subsume its object, yet this requires mimesis, which resists linguistic translation. Thus philosophy must work against language by way of language, and this is what *Aesthetic Theory* accomplishes through its constellational style.

The text demonstrates its resistance to language in its constellational account of mimesis. As we have seen, mimesis unfolds in fragmentary moments. Each moment

belongs to a constellation, yet no moment provides the complete picture. These moments are self-contained, yet simultaneously bleed into one another. Contradictory accounts are assembled alongside each other. They are not smoothed out and narrated, their development successive, but rather manifest the work of conceptualization. As disunited assemblages, the divergent accounts of mimesis expose a glimpse into a world in which thought lingers yet does not conceptualize its object.

There exists within the text an account of mimesis, its idea moves through the text's constitutive constellations, yet it never abstractly presents itself before us. This is because the nature of mimesis is its conceptual impossibility. We can only "know" mimesis in its constant rebuttal to the definitions we impose. *Aesthetic Theory* employs language to provide an account of mimesis, but never names it directly. Instead, the text invites us to perform this task as we piece together their fragmentary moments. What we piece together is not a unified whole, but a constellation. Consequently, we proceed through the text in a dialectical manner. It's idea of mimesis is absently present; it is there, but not in its self-identity.

Philosophy assumes a mimetic comportment of its own in its attempt to offer an unknowable alternative to the world it critically negates. Thus philosophy mobilizes itself against itself: it becomes dialectical not only in its content, but also in the style through which its content develops. *Aesthetic Theory* is certainly philosophical. But it preserves itself, in resistance to its instrumentalization as method, in its opposition to the concept, upon which philosophy has traditionally relied. In the face of a post-metaphysical world, the text presents itself as an experience, which has otherwise been denied by the grip in which instrumental reason binds our world.

Works Cited / Further Reading

Adorno, Theodor. *Aesthetic Theory*. Trans. Robert Hullot-Kentor. Minneapolis: University of Minnesota Press, 1997

Adorno, Theodor. *Critical Models: Interventions and Catchwords*. Trans. Henry W. Pickford. New York: Columbia University Press, 2005.

Buck-Morss, Susan. *The Origin of Negative Dialectics: Theodor Adorno, Walter Benjamin, and the Frankfurt Institute*. New York: The Free Press, 1977.

Goldmann, Lucien. *Cultural Creation*. Trans. Bart Grahl. Oxford: Basil Blackwell & Mott Ltd, 1977.

Heidegger, Martin. *Nietsche vol. 3 &4*. Trans. Joan Stambaugh, David Farrell Krell, Fran A. Capuzzi. Ed. David Farrell Krell. New York: Harper & Row, 1982.

Horkheimer, Max and Theodor Adorno. *Dialectic of Enlightenment*. Ed. Gunzelin Schmid Noerr. Trans. Edmund Jephcott. Stanford, California: Stanford University Press, 2002.

Jameson, Frederic. *Late Marxism: Adorno, Or, The Persistence of the Dialectic*. New York: Verso, 1990.

Jay, Martin. *Adorno*. Cambridge: Harvard University Press, 1984.

Kant, Immanuel. *Critique of Judgment*. Trans. Werner S. Pluhar. Indianapolis: Hackett Publishing Company, 1987.

Kaufman, Robert. "Lyric's Expression: Musicality, Conceptuality, Critical Agency." *Cultural Critique* 60, 2005: 197–216.

Nietzsche, Friedrich. "Twighlight of the Idols." Trans. Walter Kaufman. *The Portable Nietzsche*. Ed. Walter Kaufman. New York: Penguin Books, 1982. 463–564.

Richter, Gerhard. "Aesthetic Theory and Nonpropositional Truth Content in Adorno ." *New German Critique* 33, 2006, 119–135

Wolin, Richard. "Utopia, Mimesis, and Reconcilliation: A Redemptive Critique of Adorno's Aesthetic Theory." *Representations* 32, 1990, 33–49.

Endnote

[1] Identity is that which renders an entity definable and recognizable using its concept. This is an axiom of logic, according to which the identity relation holds only between a thing and the concepts of itself: $x = x$.

Nadja, or, the Dialectical Progression of André Breton and the Surrealist Movement

Gavin Arnall, Logos *Staff*

Originally published in 1928, André Breton's *Nadja* represents a turning point in Surrealist thought in which the movement was, "crossing over the gap that separates absolute idealism from dialectical materialism" (*What is Surrealism?*, 157). In this paper, parallels will be drawn between Marxist theory and *Nadja* in an attempt to elucidate the aforementioned transition as presented within this literary work. Specifically, the concepts of dialectical progression, revolution, work as labor, and sensuous activity will be analyzed as presented by Breton, either directly or indirectly through Nadja.

Although integral to both Hegelian idealism and Marxist materialism, the type or kind of dialectical progression represented in *Nadja* demonstrates the influence of Marxism on Breton. Nadja metaphorically describes thought through her image of the fountain, illustrating dialectical movement:

> Those are your thoughts and mine. Look where they all start from, how high they reach, and then how it's still prettier when they fall back. And then they dissolve immediately, driven back up with the same strength, then there's that broken spurt again, that fall. . . (*Nadja*, 86)

With neither matter nor mind given primacy, both a Hegelian dialectic of the Idea and a Marxist dialectic of the (material) fountain could be applied as an interpretation of this passage. Perhaps the ambiguity corresponds with the transitional period of Surrealist thought in general. With that said, a Marxist might question the fountain as metaphor due to the horizontal orientation of the dialectic it describes and its subsequent lack of progression. Instead of a forward movement, the dialectics of the fountain seems, like the protagonist in *The Myth of Sisyphus*, to be destined to repetitive movement in the same sphere, devoid of any reconciliation or progression. As though siding with this interpretation, Breton reflectively criticizes Nadja's view of the world, in which, "everything so rapidly [assumes] the appearance of a rise, a fall" (*Ibid*, 135). In other words, Nadja admirably acknowledges the presence of numerous contradictions in society but loses sight of the possibility of synthesis.

Breton's own notion of dialectics and how it relates to revolution is similar, yet different, in that he "saw the dialectic as a technique for thinking about the possibilities

of revolutionary change, while he condemned the System as a reactionary attempt at totalization or closure" (Roth, 12). In other words, dialectical materialism is a method to be applied to effect change, but it is not a forecast of bright skies in the future.[i] Beyond rejecting the position that history ended with Hegel, Breton also rejects any reified concept of revolution or reconciliation by asserting that these events are only the preconditions for, "the liberation of the mind, the express aim of surrealism" (*What is Surrealism?*, 155). In this formulation, Breton is not making a normative claim on the greater importance of one kind of liberation over the other, nor is he replacing one fetishized teleology with another. Through Nadja, Breton clarifies his view by alluding to Leon Trotsky's Theory of Permanent Revolution thus avoiding the anti-dialectical trap of reification. Nadja tells Breton that she chose her name, "because in Russian it's the beginning of the word hope, and because it's only the beginning" (*Nadja*, 66). This moment, combined with the passage a few pages prior about Breton's purchase of Trotsky's most recent work, suggests that Breton was influenced by Trotsky while writing these sentences. Advanced by Trotsky in 1928 as a response to Stalin's assertion of "Socialism in one Country", the Theory of Permanent Revolution could be altered from its macro-level call for international revolution to the micro-level of personal revolution within each pre-critical revolutionary subject.[ii] This reworking and broadening of the theory's application enables Breton to stay true to his view of dialectics as method. The mind would be liberated to *continue* to develop; the subject would not become a *completed, reified thing.*

Further indication of Breton's call for revolution, both material and personal, is demonstrated during his walk toward the Opéra, during which he analyzed peoples' ". . . aces, their clothes, their way of walking. . ." and decides, "No, it was not yet these who would be ready to create the Revolution" (*Ibid*, 64). A Marxist might question how such considerations would be relevant to a proletarian revolution, for the way a man walks seems irrelevant if the historical conditions are ripe. Perhaps Breton feels that the demeanor of those he is observing reflects their class/revolutionary consciousness; however, his emphasis on the subject in this passage may indicate some lingering elements of idealism.[iii] Breton then points out that Nadja, "carried her head high, unlike everyone else on the sidewalk" (*Ibid*, 64). This contrast between the average passerby and Nadja relates back to the name she chose for herself in that Nadja represents hope,

as she is the personification of the possibility of change, and a beginning, as she is one of the many needed for a revolution. This reading of Nadja, as one of the many needed for change, shifts emphasis away from the subject as individual to the subject as a social being. Instead of the Hegelian historical figure altering the course of history, man as species-being shall effect this change in solidarity with his fellow man.

Breton's inclination towards personal and material revolution is heavily driven by two aspects of life integral to a Marxist critique of society. One aspect of this dual liberation involves his vehement rejection of work. Breton reports that, "there is no use being alive if one must work" (*Ibid*, 60). This point may elicit doubt about Breton's sincerity regarding his involvement with his own work as part of the Surrealist movement; however, it is to be understood that by work, Breton means *wage labor*. Labor is man's alienated life-activity, that which makes "life itself [appear] only as a *means to life*" (*The Marx-Engels Reader*, 76). Marx explains that work like Breton's was not labor because "a writer does *not* regard his works as means to an end. They are an end in themselves; so little are they 'means', for himself and others, that he will, if necessary, sacrifice his own existence to their existence."[iv] As though verifying Marx's account of the writer, Breton wrote in *What is Surrealism?* that he would accept the consequences of his refusal to work, that he has, "agreed to pay the price for [his] non-slavery to life" (*What is Surrealism?*, 218).

Interestingly, although Breton calls for the "better, that is . . . fairer, division" of labor, he states that "life's grim obligations make it a necessity" (*Nadja*, 59). It seems that Breton's assertion of the necessity of alienated labor is partly correct in that certain production is necessary for the continuation of life, yet he makes no reference to the possibility that the alienation derived from this labor could be transcended in some way. One way to surpass this alienation would be through totally alienating such activity from man qua man via "total automation" (Marcuse, 156). This idea, advanced by Herbert Marcuse in *Eros and Civilization*, stems from Marcuse's own rejection of any intrinsic value of the laborious and un-free nature of work-as-labor. With the proper organization, any work performed by man would be more adequately described as play—that is free, creative, and spontaneous activity. The basic necessities would be handled by automated machinery while man could develop his individual capacities and satisfy other pleasures that were primarily ignored, or rather *repressed*, during the period of wage-slavery. As Ernst Bloch said in his essay

"Marxism and Poetry", "the time will come. . . when having an imagination is no longer a crime" (Bloch, 158). In other words, activities like libidinal creativity through art and sexual interaction as libidinal and social (rather than productive) would not be seen as a waste of time, hindering the further development of man, or merely recreational. Rather, such activity would actually *enable* man to *continue* to develop.

Another aspect of Breton's dual liberation of man involves the incomplete and repressed development of the senses. When Breton first meets Nadja, he asks, "'who are you?' And she, without a moment's hesitation: 'I am the soul in limbo.'" (*Nadja*, 71). Although multiple interpretations could be applied to this quote, it is imperative to look to the role of identity as definition in this passage. When confronted with the question of defining herself, Nadja opts-out of definition, of "reified is-ness" for the not-yet, the soul-in-limbo (Holloway, 62). Even the very signifier she chose for herself is only the *beginning* of a word. Although not made available to the public until after this book was published, a parallel can be drawn between the aforementioned concepts expressed by Breton through his presentation of Nadja and the early work of Marx. In his 1844 *Economic and Philosophical Manuscripts*, Marx wrote that, "The *sense* caught up in crude practical need has only a *restricted* sense. For the starving man, it is not the human form of food that exists, but only its abstract being as food" (*The Marx-Engels Reader*, 89). Later in that same work, Marx says,

> The transcendence of private property is therefore the complete *emancipation* of all human senses and attributes. . . precisely because these senses and attributes have become, subjectively and objectively, human. The eye has become a *human* eye. . . (*Marxism and Art*, 59)

In other words, Nadja and others are incomplete in their current, transitory state—they are in limbo. As Terry Eagleton comments in *The Ideology of the Aesthetic*, "if communism is necessary, it is because we are unable to feel, taste, smell and touch as fully as we might. . ."(Eagleton, 201). Our objects, like our activity, take on solely an instrumental role in our lives. Food becomes a means to an end as our senses are subjugated for the prolongation of existence. The freeing of the senses would entail the advent of a

new dimension, the opening of a path by which man may continue to progress.

As noted by Walter Benjamin, Pierre Naville referred to the politicization of the Surrealist movement as a dialectical development (Benjamin, 209). *Nadja* is the material reality of this development-as-transition produced by one of the movement's leaders. Breton's portrayal of Nadja, along with secondary accounts of their relationship, alludes to her influence on Breton as the spark for his personal revolution of thought.[v] As though following the footsteps of its founder, Breton came to dialectical materialism through another.

Works Cited / Further Reading

Benjamin, Walter. "Surrealism: The Last Snapshot of the European Intelligentsia." *Selected Writings: Volume 2: 1927–1934*. Trans. Rodney Livingstone *et al*. Ed. Michael W. Jennings, Howard Eiland, and Gary Smith. Cambridge: The Belknap Press of Harvard UP, 1999.

Bloch, Ernst. *The Utopian Function of Art and Literature: Selected Essays*. Cambridge: MIT Press, 1988.

Bottomore, Tom. *A Dictionary of Marxist Thought*. Malden: Blackwell Publishing, 1991.

Breton, André. *Nadja*. New York: Grove Press, 1960.

Breton, André. *What is Surrealism?* New York: Pathfinder, 2006.

Eagleton, Terry. *The Ideology of the Aesthetic*. Malden: Blackwell Publishing, 1990.

Holloway, John. *Change the World Without Taking Power: The Meaning of Revolution Today*. Ann Arbor: Pluto Press, 2002.

Marcuse, Herbert. *Eros and Civilization: A Philosophical Inquiry into Freud*. Boston: Beacon Press, 1955.

Marx, Karl. "Economic and Philosophic Manuscripts of 1844", ed. Robert C. Tucker, *The Marx-Engels Reader*, New York: W. W. Norton & Company, 1978.

Marx, Karl. "The Formation and Emancipation of the Senses", ed. Maynard Solomon, *Marxism and Art*, Detroit: Wayne State University Press, 2001.

Rosemont, Penelope. *Surrealist Women: An International Anthology*. Austin: University of Texas Press, 1998.

Roth, Michael S. *Knowing and Reason: Appropriations of Hegel in Twentieth-Century France*. Ithaca: Cornell University Press, 1988.

Solomon, Maynard. *Marxism and Art*. Detroit: Wayne State University Press, 2001.

Tucker, Robert C. *The Marx-Engels Reader*. New York: W.W. Norton & Company, 1978.

Endnotes

[i] See Bottomore, Tom, *A Dictionary of Marxist Thought*, 224–238.

[ii] See Holloway, John, *Change the World Without Taking Power: The Meaning of Revolution Today*, 140–154.

[iii] Precursor elements of the Italian Autonomist Marxist movement can also be seen here.

[iv] Marx, Karl. "Economic and Philosophic Manuscripts of 1844", as quoted by Terry Eagleton, *The Ideology of the Aesthetic*, (Malden: Blackwell Publishing, 1990), 204.

[v] see Rosemont, Penelope, *Surrealist Women: An International Anthology*, xlv.

The Inverted Spectrum as Grounds for Perceptual Relativity

Derrick Ward, Logos *Staff*

In the debate on perception, many contemporary philosophers have taken a stance known as representationalism. This stance avoids naïve realism by relying on the idea of intentional content of objects. Such content is most basically defined as ideas with no properties of their own, which instead suggest, or are "about" something else. Harman provides an example of intentionality through mentioning a painting of a unicorn. The physical nature of the painting is nothing more than canvas and various paints; however, the painting is recognized by the perceiver as depicting a unicorn. The paint on the canvas has suggested the image of a unicorn, which would therefore make the unicorn the intentional object of the painting[i]. Representationalism hinges upon the notion that sameness of intentional content should equate to sameness in the phenomenal character of the experience[ii]. More simply, this idea is to say that if two experiences offer the same suggestions, then they must yield the same response as to what the experience "was like". Perhaps the most challenging obstacle raised for this way of looking at perception is known as the inverted spectrum theory, which nullifies claims about the necessity for equation between intentional content and phenomenal character, as will be explained shortly (Block). The goal of this paper is to examine the implications of spectrum inversion and to find the most reasonable conclusion to deal with these implications. I intend to argue that the inverted spectrum theory, when held as both possible and rife, leads to the indeterminate yet inevitable conclusion of perceptual relativity.

The inverted spectrum theory, on a basic level, is illustrated through the experience of two unique perceivers and the functionality of that experience. In the theory, as Harman explains it, these two perceivers both look at the same object; however, one of the perceivers' visual color spectrum has been inverted. Both subjects look at the object and identify it as having a certain color, perhaps red. Despite both parties functional similarity, the ability to identify the object with the color word red, they differ in their mental accounts of color experience. The "normal" perceiver has a red experience, while the inverted perceiver has a green experience. To clarify, suppose that Normal Nancy and Inverted Ivan are both looking at the same fire engine. When asked the color of the fire engine both Ivan and Nancy verbally identify the fire engine as being red; however, since Ivan's color spectrum is the exact opposite of Nancy's, Nancy is identifying her red experience as the color red while Ivan identifies his green experience as

being the color red. Nancy and Ivan are functionally identical since both offered the same verbal description of the object. For all of Nancy's life she had been taught to verbally identify red experiences, associated with fire engines, blood, and other things that produce such an experience, as being and possessing the color red. Similarly, Ivan had been taught to verbally identify his green experiences, also associated with items such as blood and fire engines, as being and possessing the color red. Both Ivan and Nancy were presented with the same object, which implies that they both received the same intentional content; therefore, according to the aforementioned assumption by representationalists, both Ivan and Nancy should have experienced identical phenomenal character associated with the object.

Since Ivan, who had a green experience, and Nancy, who had a red experience, differed in the phenomenal character of their experience, the representationalist has a problem. In order to remain objective (mind-independent) about color experience, as is the goal of the representationalist, the assumption is made that the object, experienced as red by Nancy, is experienced as such because the fire engine possesses intentional qualities representative of, or suggestive of red. The problem for the representationalist is obvious. Ivan had a green experience while looking at the same object. This difference in phenomenal character of color between Ivan and Nancy is challenging because if the color experience were truly objective, and intentional content caused both experiences, then the experiences would presumably be the same.

Another problem posed by this theory lies in the question of functionality. Both parties are functionally identical in the sense that they both identify a given object (the fire engine) with the same color word (red); however, the parties can be seen as functionally different when one considers what each individual meant by the applied color word. When Ivan describes the vehicle as red he means that it has the same phenomenal character as experiences associated with blood and ripe tomatoes. Unfortunately, due to the inversion, the phenomenal character that Ivan is describing is the same as the phenomenal character that Nancy experiences when she looks at grass. This difference implies that at least one party is either failing to use color vocabulary to express beliefs about color, or is failing to use phenomenal experience in formulating beliefs about color. More simply put, Ivan is either saying red when he means green,

or he is ignoring his experience and applying the terms of color as learned. Now the position is one of perceivers who, when presented with identical stimulus, differ in the phenomenal character of their experience and thus lack functional similarity, huge problems for a representationalist.

This being said, it is now possible to explore the possible implications of this problem. In order to resolve the issue, the representationalist must now find a plausible explanation for the breach in intentional and phenomenal equation. Luckily, several avenues of approach are available for such a solution. The simplest, and most obvious reaction to solve such a problem is to deny the possibility of spectrum inversion altogether. If such an inversion were impossible then the concerns addressed above become irrelevant and the representationalist can continue in his/her belief without worry; however, since no evidence exists against the possibility of such inversion, it would seem illogical to rule spectral inversion out. This is especially true considering the feasibility of such a theory[iii].

Tye takes the next logical step in refuting the problems raised by spectrum inversion by calling such a perceptual difference misperception. This claim would hold that any perceiver whose phenomenal content differed from "normal" or standard interpretations of intentional content, would in some way be wrong. Such a claim is made by stating that the function of a color experience is to indicate the presence of a color; therefore, if red is present and a person has a green experience Tye would see this as a malfunction. By writing these troublesome differences in color perception off to error, representationalism remains intact since "mutant" misperceptions about color lack veridical representation, and, can therefore be disregarded. This claim can easily be refuted assuming, as Block does, that spectrum inversion is rife. If a fifty-fifty split existed between "normal" and inverted perceivers, then no reason exists to see the inverted as wrong or abnormal since their concept of a particular color experience is just as prevalent as the other concepts.

Given the unconvincing nature of the past two objections, the next move is to reject representationalism and its implied objective nature. Objective view about color, assuming the prevalence and possibility of spectrum inversion, seems, at this point, irrational. Representationalism had remained objective because the perceiver had no

active mental role in experience, and, rather, was simply affected by physical properties and intentional content. Now a subjective view on color can be taken due to the rife nature of perceiver dependent color experience. To clarify, because half the population supposedly sees things one way, while the other half sees things another way, an element of subjectivity cannot be avoided. Dispositionalism can now be applied. This theory of perception divides qualities of objects into two groups: primary and secondary qualities[iv]. Secondary qualities are qualities defined by their disposition to cause a certain reaction. For instance, something is said to be fragile if and only if it is disposed to break when struck, thereby defining fragile solely in terms of an objects disposition. If color were to be considered a secondary quality, an object could be called red if and only if it were disposed to cause a red sensation or appearance. Color cannot then be considered a primary quality, which are defined in ways other that their disposition to cause an experience. Weight and size serve as examples of primary qualities because they can be quantified. Colors fit nicely into the category of secondary qualities because, aside from the disposition just discussed, they cannot be more simply defined. It seems strange at first, but color words are completely basic and one cannot describe "blue" without use of the word "blue" itself or through reference to an object exhibiting that color.

Because the inverted spectrum has been assumed to be more prevalent than isolated cases, neither "normal" nor inverted color experience can be considered more correct that the other. Due to the subjective nature of color experience, implied by dispositional secondary quality classification, color experience becomes relative to the perceiver. If one individual experiences something as green, while another individual experiences the same object as red, neither party can effectively argue that their experience is somehow more correct. Any argument used by one perceiver could be used with equal strength against the other, and vice versa. This is the idea supporting perceptual relativity, which is the conclusion that both parties are right in their own respect. The question is no longer one of which "red" experience is correct, but rather a question of "red" for whom. An individualistic account of color experience such as this proves to be somewhat inconclusive if the overall goal is to find the "true" nature of color experience, because truth is now in the hands of the perceiver.

By abandoning representationalism, and the objectivity that goes along with it, and adopting the subjective solution posed by dispositionalism, a sound yet inconclusive (about "true" color experience) conclusion has been provided for the perceptual questions the inverted spectrum theory raises. Despite revised theories about spectrum inversion, such as those offered in Block's works, no other theory of perception can account for the issues discussed as well as dispositionalism and perceptual relativity. Short of resigning thinking about perception to a concept where every object contains every possible disposition, an obvious fallacy, perceptual relativity about color appears to be the only plausible option.

Works Cited / Further Reading

Block, Ned. "Inverted Earth" *Philosophical Perspectives*, 4, 1990, 53–79.

Harman, Gilbert. "The Intrinsic Quality of Experience" *Philosophical Perspectives*, 4, 1990, 31–52.

McGinn, Colin. *The Subjective View: Secondary Qualities and Indexical Thoughts*. Oxford: Oxford U. P., 1983.

Tye, Michael. "Of Colors, Kestrels, Caterpillars, and Leaves", *Journal of Philosophy* 48.9, 2001, 469–487.

Tye, Michael. "Visual Qualia and Visual Content Revisited" *Philosophy of Mind: Classical and Contemporary Readings*, ed. D. J. Chalmers. Oxford: Oxford U. P., 2002.

Endnotes

[i] See Harman, Gilbert, *Philosophical Perspectives*, pp. 31–52.

[ii] See Block, Ned, *Philosophical Perspectives*, pp. 53–79.

[iii] See Tye, Michael, *Journal of Philosophy*, pp 469–487 and Tye, Michael, *Philosophy of Mind*, pp. 447–456.

[iv] See McGinn, Colin, *The Subjective View*.

Interview

by Chris Post

On His Anticipated New Book, *Globalizing Justice*: An Interview with Professor Richard W. Miller

Professor Miller joined the Cornell University Sage School of Philosophy in 1975 after teaching at M.I.T. His primary interests lie in social and political philosophy. His writings and courses on social justice have covered the moral foundations of egalitarianism, the moral status of group loyalties and the proper political role of evaluations of ways of life.

(1) You've written about such topics as our duties toward the global poor, the relationship between transnational duties and duties towards compatriots, the ethics of war, moral problems of globalization, global climate change, and the moral implications of American power. Which of these can we expect to see in your upcoming book, Globalizing Justice?

All of them. The unifying thought is that demanding responsibilities to people in other countries reflect relationships of power that reach across borders. Duties are shaped by these relationships in all the areas you've mentioned.

(2) In Beneficence, Duty and Distance *you introduced a powerful criticism of Peter Singer's "principle of sacrifice" and "radical conclusion", arguing that "adequate reflection on our most secure convictions would lead most of us to embrace a less demanding principle of general beneficence"[1]. Can you remind us what, in your view, was wrong with Singer's principle, and what a less demanding principle would look like?*

Singer says that whenever you have an opportunity to prevent something very bad from happening by making a morally insignificant sacrifice, you should make

that sacrifice, even in the absence of any significant relationship with the one in need. Because of global poverty and international agencies, this could require giving up all luxuries and frills. But most of us have worthwhile goals that cannot be pursued enjoyable and well if we do this. To express my sense of style in how I dress, for example, I have to sometimes buy a better-than-basic shirt, rather than saving up this money for an Oxfam donation. So sticking to Singer's principle on each occasion could worsen my life. And one has no duty to adopt an underlying concern for neediness that is so demanding that this worsens one's life. Rather, you should look at your underlying attitudes and their impact on your life as a whole, and ask, "Could I have greater underlying concern for neediness and fulfill my other responsibilities, without worsening my life?" If the answer is "Yes" (it is for me!), you should be more responsive to neediness. But unless the goals with which you identify are very inexpensive, you will have reason to sometimes buy luxuries or frills, rather than giving to Oxfam. Singer's principle only looks at the sacrifice on each occasion, when what is given up is small. Mine looks at the impact on one's life as a whole of an underlying attitude, which must allow for more self-concern on some occasions to avoid significant worsening of one's life as a whole. This due concern for one's own life-goals doesn't involve disrespect for others or valuing their lives as less important, any more than my wife and I valued our daughter's life less than the lives of children in Cambodia when we spent money on her college education that might have saved Cambodian kids from death at an early age. Our choice expressed an appropriate valuing of our relationship to her. Similarly, appropriate self-concern expresses an appropriate valuing of one's relationship to oneself, not a disvaluing of others.

(3) In Moral Closeness and World Community, *you said "ordinary moral thinking about aid to needy strangers discriminates in favor of the political closeness of compatriots and the literal closeness of people in peril who are close at hand"[ii]. One of your goals in that article was to defend our normal biases but at the same time to show that a proper understanding of their justification establishes substantial, even if less demanding duties to help the foreign poor. Can you give us a quick and dirty summary of the argument before we move on to discussing your new book?*

Take the ordinary view that fellow-citizens have a strong, special duty to help the disadvantaged among them. What ties me in this way to someone in Los Angeles whom I have never met and never expect to meet? For one thing, the fact that the political choices that I make as a citizen shape her prospects of self-advancement; this creates a responsibility of special concern for her if her life prospects are inferior. But as a citizen of the sole super-power, I also take part in a political process with considerable impact on life-prospects abroad. So I also have special responsibilities to the people whose lives are shaped. For example, I have these responsibilities to those whose lives have been changed by structural adjustment policies and trade policies shaped by my government, to those living under client regimes propped up by my government and those whose countries are devastated by U.S. firepower. Among those at home in the territory of a government, power creates responsibility. Once the connection is appreciated it extends to foreigners whose lives are shaped by the same government.

(4) Have any of your goals or positions changed since writing Beneficence, Duty and Distance, *or* Moral Closeness and World Community?

In those essays, I was most concerned to show that the mere distribution of resources and needs did not create responsibilities and to resist exaggerated demands that Singer and others based on this circumstance. Sometimes, this project led me to underestimate the extent of the moral demands that result from actual transnational relationships. For example, I now think that morality might demand more strenuous U.S. political action to help needy people abroad than to help needy Americans. I didn't quite deny this in the earlier essays, but I came close. I needed to move on and look more closely at how the U.S. government and U.S. firms take advantage of people in developing countries. I also needed to distinguish different senses of "patriotism." I tended to think of it as a matter of special responsibilities to compatriots. But, strictly speaking, patriotism is love of one's country. Thinking about the specific nature of this attitude, I have come to believe that enduring tendencies of U.S. foreign policy make it desirable to unlearn American patriotism in favor of more cosmopolitan ties. I didn't think much about this at first, because love isn't something we think about much in academic political philosophy.

(5) Were you responding to any particular critics when you changed your views on international responsibilities?

People with views very different from my own such as Peter Singer and Dick Arneson forced me to clarify my claims and premises. I am also very grateful to students in my classes. They have taught me a lot in resisting my arguments, and insisting on the superiority of writings that I assign which advance positions different from mine. For example, many are attracted to Henry Shue's approach, and have forced me to clarify our similarities and differences and to justify my views where they differ.

(6) How do you plan to argue for your current positions?

In the book that I am completing, I look at specific ways in which people and governments in developed countries take advantage of people in developing countries by taking advantage of their weaknesses. "Exploitation in transnational commerce," "inequity in the framework for commerce," "irresponsibility of the American empire" point in the three main directions. Then I describe how we would express a proper valuing of genuine cooperation by changing these relationships and their outcomes. General moral principles about respect for autonomy and specific interpersonal examples play a role. But empirical inquiry into current transnational interactions is very important, as showing the nature and depth of these processes of taking advantage and suggesting ways of responsibly relating.

(7) Many people in the field talk about the duties generated to foreigners by the existence of an 'American Empire'. Is there an 'American Empire', and if so, what sort of duties do we owe?

Of course, there isn't a literal empire. There is no official assertion of authority (apart from temporary limited cases, such as the Coalition Provision Authority in Iraq.) But the U.S. has enormous special prerogatives to get its way, for example, borrowing prerogatives due to the role of the dollar. It has enormous influence due to the power

of its threats, for example threats to exclude goods from its vast internal market. And it has enormous destructive power, exercised to eliminate hostile regimes and gain access to resources, as in Iraq. Even when this influence isn't exerted directly and bilaterally, it is exerted at second hand, through influence over such institutions as the World Bank, WTO and IMF. In many parts of the world, this influence is deep and much greater than any other power's. So one can speak, say, of Guatemala, Egypt and the Philippines as metaphorical territories of the American empire. To respect the autonomy of people in this empire, the benefits from manipulating their neediness (say, in structural adjustment) should be used to advance their interests; the U.S. should take on the unmet responsibilities of client regimes it props up; and it should make good damage from the exercise of destructive power. Of course, the U.S. is not alone in the process of domination. It is just the leader and, so bears the leading burden of responsibility.

It is also important to appreciate that "empire" is just a metaphor and not to exaggerate U.S. influence. But this metaphor ties together lots of important projects of current social movements. Given facts of power summed up in the metaphor, letting immigrants from the developing world into the U.S. isn't a gracious favor. U.S. immigration barriers are like the pass laws under apartheid. Once in Iraq, the U.S. was bound to shed oceans of blood, because of imperial interests it pursued and the need to preserve fearsomeness by avoiding a quick retreat (a retreat that Kerry and Obama opposed as well as Bush.) The U.S. government will resist equitable solutions to global climate change, fearing the more rapid loss of empire in the face of the rising power of the major developing countries, especially China. By the same token, alliance with people around the world to reduce excesses of the American empire is a fruitful way to make the world a bit better, while laying the groundwork for future global solidarity.

Endnotes

[1] Miller, Richard W. "Beneficence, Duty and Distance" *Philosophy & Public Affairs* 32.4, 2004, 357–383.
[2] Miller, Richard W. "Moral Closeness and World Community" *The Ethics of Assistance: Morality and the Distant Needy*. Ed. Deen K. Chatterjee. Cambridge: Cambridge U.P. 2004, 101.

Submit to Logos

Logos is now accepting submissions for subsequent issues.

Guidelines:

- Submissions can address any philosophical issue.
- Submissions should be under 30 pages in length.
- Please enclose an abstract for your paper.
- All papers should follow MLA formatting.
- Include your name, year of study, college, and major on a separate title page. Do not include author information on the individual pages of the paper.

A cash prize will be awarded to the author of the best submitted paper.

Papers should be submitted by email attachment (.doc) to:

logos-mailbox@cornell.edu.

For more information, including contact information and submission deadlines, visit:

www.rso.cornell.edu/logos