The Undemandingness of Progressive Consequentialism

JOSEPH PHILIP FLANAGAN

THE UNIVERSITY OF BRISTOL

Abstract: Michael Slote's Satisficing Consequentialism seems to offer a way to reduce the demands of Consequentialism, by only requiring us to bring about consequences which are good enough. However, it has proved to be a flawed and unpopular theory. Dale Jamieson and Robert Elliot's Progressive Consequentialism, the view that our fundamental ethical imperative is to improve the world, appears to solve some of the problems of Satisficing Consequentialism, including arbitrariness, prevention of goodness and blatant satisficing. This paper shows that despite its apparent successes, Progressive Consequentialism is an implausible theory because it is implausibly undemanding. It does not always require agents to do what they clearly ought to do.

§1. INTRODUCTION

In this essay I will argue that Progressive Consequentialism (PC), the ethical view that right actions are those which improve the world, is implausibly undemanding. Any ethical theory which does not place adequate demands upon agents should be rejected, and thus PC should be rejected.

I will begin (§{2&3}) by explaining the motivations for PC, particularly focussing on its apparent superiority to Satisficing Consequentialism (SC), the view that right actions are those which produce good enough consequences. In the process of motivating PC, I introduce it (§3) and explain how it functions. Next, I present (§4) a new objection to PC, that of undemandingness, to which I consider (§5{a-b}) several responses. I argue (§5) that these responses are implausible, concluding (§6) that PC is implausibly undemanding and so should be rejected.

For the purposes of space, I will assume that what is of intrinsic value in consequences is human wellbeing, so when I refer to the value of consequences, I refer to wellbeing. It should also be noted that I mostly use 'good' in the same sense as 'value'.

§2. SC AND THE MOTIVATIONS FOR PC

Perhaps the most well-known form of Consequentialism is Maximising Consequentialism (MC), the view that we ought to bring about the consequences with the most value. Under MC, producing consequences with less than the most value is wrong, so it is a very demanding view. If working at the soup kitchen produces even slightly better consequences than devoting time to treasured hobbies, then we may be required to give up these sentimental projects, or at least severely limit the time that we allocate them.

One response to the demandingness of MC has been to argue that we only ought to bring about consequences which are 'good enough', for some consequences whose value is less than the best, but still good enough, will suffice. This view is called Satisficing Consequentialism (SC) and was first defended by Michael Slote ("Satisficing Consequentialism").

While it is true that SC can reduce the demands of MC, it faces several problems, namely: arbitrariness, gratuitous prevention of the good, and blatant satisficing.

The primary motivation for PC is that it can avoid these three problems, while still reducing the demands of MC. I will now explain why these problems seem to apply to SC, then showing, in the following section (§3), how PC can avoid them.

Problem 1 – Arbitrariness

SC may be arbitrary because there appears to be no principled way to decide what amount of value is good enough. If we choose some

quantity, perhaps X% of the most valuable consequences possible in the situation, then we must question the significance of X as against X-1 or X+1. If what is good enough is some fixed amount of good Y, then the same issue will arise (Jamieson and Elliot 244).

Problem 2 – Preventing the Good

Bradley ("Against Satisficing Consequentialism") has argued that SC permits cases of gratuitous prevention of the good, where agents are permitted to bring about a good enough outcome, by preventing something much better from happening. An example will help to illustrate this problem:

Essay: To pass his history essay, Milo needs a mark of at least 60%. However, getting higher will give him a great confidence boost and increase his wellbeing. For advice, Milo asks his older sister Effie to look over the essay. Calculating that Milo will receive a mark of 90%, Effie alters the answer, calculating that Milo will now receive 60%. Effie, a Satisficing Consequentialist, claims that she has acted rightly, for while 90% would be much better than 60%, 60% is still good enough.

Having assumed that human wellbeing is of intrinsic value, it is clear that getting 90% would be a consequence of greater value than getting 60%. However, because 60% produces a wellbeing increase which is good enough, Effie is permitted to prevent Milo from getting 90%. We intuitively feel that Effie cannot have done the right thing by going out of her way to bring about a worse outcome, yet SC delivers the opposite verdict.

Problem 3 – Blatant Satisficing

Mulgan (Slote's Satisficing Consequentialism 122-125) argues that SC permits cases of blatant satisficing, where agents are permitted to knowingly bring about less than the most good, even when bringing about the most would come at no additional personal cost or effort. To make this clearer, consider the following example based

on Mulgan's "Magic Game" scenario (Slote's Satisficing Consequentialism 125):

Aphrodite: Aphrodite is in a room with two buttons. Pressing button A clears the debts of N people and button B clears the debts of 1000 people. Once a button is pressed, both are deactivated and the situation ends, with both buttons being equally easy to mechanically activate. Aphrodite is a satisficing consequentialist who knows all this. Aphrodite claims there is some number of people N which is considerably lower than 1000 but which is good enough, and presses button A.

Clearing the debts of 1000 people would greatly increase wellbeing and comes at no more effort or cost to Aphrodite than clearing N peoples debts. We intuitively feel that Aphrodite cannot have acted rightly by knowingly bringing about this lower value, when bringing about more would come at no greater effort or cost, yet SC delivers the opposite verdict.

It should be noted that a different intuition is being violated than in problem two. In two, the intuition being violated is that we ought not to go out of our way to *prevent* good outcomes, while in this case the intuition is that if we can bring about more good at no extra effort or cost, then we should.

I do not claim that these problems provide a reason to reject SC wholesale, although they likely contribute to its current relative unpopularity – Slote himself, the originator of the view, has rejected it (*Morals From Motives*). I will now introduce PC and indicate how it might avoid these problems while still reducing the demands of MC.

§3. PROGRESSIVE CONSEQUENTIALISM

PC, first proposed by Jamieson and Elliot, is the view that an action is right if and only if it improves the world (244-245). But what does it mean to improve the world? From an initial consequentialist interpretation, improving the

world might mean bringing about a future world which has a higher value than the past one. However, if factors outside of an agent's control mean that the value of the world will vastly increase regardless of how they act, then the agent will not be required to do anything and will in fact be permitted to have a negative impact on the world (Jamieson and Elliot 246).

As this is implausible, we will need a more nuanced way to define improvement, which Jamieson and Elliot (247) provide[1]: An act improves the world if and only if the value of the world after the act is greater than the baseline for improvement, where the baseline is the value of the world at T2 (after the act) on the counterfactual assumption that the agent does not exist at T1 (the time of action). At first seemingly complex, this baseline is actually intuitive and easy to apply. We consider what the world would have been like at T2 if the agent did not exist at T1 and compare it to the actual world at T2. This should be made clear by example:

Cookies: Simon can give Koko some cookies or no cookies. Koko likes cookies, so having some would increase her wellbeing. Should Simon give her some cookies at lunchtime? Suppose that Simon does not exist at lunchtime. Then presumably Simon cannot give Koko cookies. Alternatively, if he gives her cookies at lunchtime, then she will have a higher wellbeing than in the world where he doesn't exist. Thus, by comparing these two worlds, we can see that giving Koko some cookies at lunchtime improves the world, while not doing so does not.

This initial formulation of PC seems to avoid two of the problems faced by SC, those of arbitrariness and prevention of the good. First, as Jamieson and Elliot (244) note, a requirement to improve the world is no more arbitrary than the MC requirement to maximise value, so PC doesn't appear to suffer from the same level-setting problems as SC.

Secondly, we can see how the formulation avoids the problem of prevention of the good by applying PC to the *Essay* example from section two. If Effie did not exist at the time of action, then she could not have altered Milo's essay, so he would have received 90%. Therefore, the baseline for improvement would be the value of the world plus the wellbeing boost that Milo would receive from getting 90%. As getting less than 90% would produce a lower wellbeing improvement, Effie's action does not improve the world against the baseline, so is wrong.

More generally, if an agent does not exist at T1, then they do not act at T1, and hence, any value in the world at T2 that arises from the agent's inaction is included in the baseline.

While this initial formulation of PC can answer objections of arbitrariness and preventing the good, it is, in its current form, guilty of permitting blatant satisficing. Recall the example *Aphrodite* from section two. If Aphrodite did not exist then neither button would be pressed, so pressing either button would be an improvement upon this baseline. Thus, this formulation of PC permits either button being pressed, just like SC.

In response, Jamieson and Elliot (245-46) complete their formulation of PC with an efficiency requirement (ER). Jamieson and Elliot argue that PC requires agents to ensure that no other action of the same effort level improves the world more. They note that this is not a demanding requirement, for agents will not be required to expend *more* effort, just to be efficient in its expenditure. Clearly, if Aphrodite is required to be efficient[2] in her effort, and both buttons require the same effort, then she ought to press the one which produces consequences of more value.

Finally, it should be clear that PC will be less demanding than MC. Suppose I have £1000 to spend as I like, which I do not particularly need. MC will require me to use the money to bring about the best possible outcomes: donating to the most effective charities and causes, while PC merely (roughly) requires that I bring about a net good.

Through showing how PC might avoid the major problems faced by SC, while still reducing the demands of MC, I have provided both an explanation of, and motivation for PC. I will now end this section by stating PC in its full form.

Progressive Consequentialism (PC):

An action is right if and only if:

i) *It improves the world against the baseline for improvement, and;*
ii) *There is no other action of the same effort level which improves the world more*

Where the baseline for improvement is the value of the world at T2 (after the act) on the counterfactual assumption that the agent does not exist at T1 (time of act).

I will now present a new problem for PC, that of undemandingness.

§4. THE PROBLEM OF UNDEMANDINGNESS

PC considerably reduces the demands of MC and answers several key objections to SC. However, we expect ethical views to be adequately demanding, and PC reduces demandingness too much; it is implausibly undemanding. Specifically, in cases where a small increase in effort yields a huge increase in the value of consequences, PC does not require agents to expend the additional effort. For example, I present another variation on Mulgan's "Magic Game":

Dionysus: Dionysus is in a room with two buttons; a higher and a lower one. Button A is within reach and clears the debts of 1 person. Button B is out of reach and clears the debts of N people. Pressing button B will require Dionysus to make a running leap, expending a small amount of effort. Dionysus is a progressive consequentialist who knows all this, and who believes the number of people N to be

1000. Dionysus presses button A, arguing that he has improved the world and so acted rightly.

Dionysus improves the world because if he didn't exist neither button could have been pressed. The efficiency requirement is also satisfied, for pressing button B requires additional effort. Therefore, under PC, Dionysus acts rightly in pressing A.

This is a deeply unintuitive verdict. If we can bring about vastly better consequences at only a slight effort increase, then it seems we ought to do so. An ethical view that does not demand this of us is implausibly undemanding. The intuition being violated is the following:

Intuition 1: If an agent can bring about vastly better consequences at only a slight effort increase, then they ought to do so.

However, it should be noted that this intuition is not universal, for it does not seem to apply in some cases. Firstly, if the lower button (A) produces a very high level of good, then it may be argued that the agent is not required to bring about the better consequences. Secondly, if the value of the world is sufficiently high, then the agent may not be required to bring about very high value consequences.

As a response, it will suffice to point out that the intuition does still certainly apply in cases where the lower button brings about consequences of moderate value or less, and that it is very implausible to suggest that the actual world has such a high value. In fact, it is generally agreed that we are very far away from such a morally perfect world. Therefore, this intuition still applies in many cases, and thus PC is implausibly undemanding in many cases.

There is also a second intuition which PC violates. In the example, regardless of how high the value of N is, PC will not require Dionysus to press the button requiring more effort. This is because an increase in the value of N does not change the fact that pressing the lower button

still improves the world and satisfies the ER. We can formalise this intuition as follows:

> **Intuition 2**: If exerting some small extra effort E produces consequences of value N, then there is some value of N for which we ought to exert extra effort E.

Even if the world has a very high value, or the lower button produces a very good outcome, it seems there must be some amount of good for which we are required to exert a small amount of extra effort to bring about. This intuition seems to apply more generally than intuition one, providing additional evidence that PC is implausibly undemanding in many cases. We will return to a version of this second intuition in the next section (§5) when discussing possible responses to undemandingness. It should be noted for later reference, that this intuition also seems strong when the extra effort is larger, for we think people ought to make large sacrifices for *some* level of good.

At this point, having presented an initial example of undemandingness, and isolated the intuitions which are violated, it would be prudent to deflect the initial response that the example is unrealistic and implausible. Perhaps, for the PC advocate it is too abstract and divorced from real life situations, and hence does not form a strong case against PC. While it is understandable to balk at the abstract treatment of the example, we can easily respond with more fleshed out, concrete examples of PC's undemandingness. For instance, take the following:

Choice: Jackson turns a corner onto a mostly empty street, to see two incidents. First, a worker who is in a rush has dropped his papers all over the pavement and clearly requires help. Second, a blind old lady with hearing aids is slowly but surely walking further into a busy main road. Jackson recognises that he only has time to help one person, and that saving the old lady will require slightly more effort, for she is somewhat further away than the worker. However, Jackson is an adherent of PC and bends

over to help the man pick up the papers, while a lorry obliterates the old lady. Jackson is a progressive consequentialist and claims to have acted rightly.

There was no one else on the street to help, so if Jackson did not exist, then the lady would have died, and the worker would have picked up his papers alone. Clearly, he has improved the world, for the wellbeing of the worker is greater than in the baseline. Also, as the example stipulates, saving the lady would have required more effort, so the ER is satisfied. Therefore, under PC, Jackson acts rightly.

This is a realistic example of the undemandingness of PC. Only a small amount of effort would be required to save a human life, the prerequisite for wellbeing, yet PC does not require Jackson to bring about this vastly better outcome. Furthermore, even if we suppose that Jackson could have saved ten or twenty lives by expending this small effort, PC could not have required him to do so, because helping the worker both improves the world and satisfies the ER.

Choice is a realistic example of the implausible undemandingness of PC. Unless the PC adherent can provide a plausible response, PC is implausibly undemanding and should be rejected. I will now consider several possible responses to the problem of undemandingness, arguing that they are all implausible.

§5. Responses to Undemandingness

§5a. The effort differences in the counterexamples are the wrong *kind* of effort.

The effort differences in *Choice* and *Dionysus* are of the physical kind: running as against bending over, jumping rather than standing still. The defender of PC may argue that we are focussing on the wrong *kind* of effort. If this is true, then perhaps the examples given so far do not provide evidence that PC is implausibly undemanding.

To refute this objection, I will consider two plausible accounts of effort suggested by Chappell (253), showing that neither renders PC any less susceptible to undemandingness.

Willpower: The kind of effort that is of normative significance is the exertion of willpower, or the expenditure of mental effort. For Elliot Kipchoge, jogging down the street will require an insignificant amount of willpower, while for someone with a crippling phobia that running will cause them to go into cardiac arrest, it will take enormous mental effort.

This account still leaves PC susceptible to undemandingness. Consider the following example, once again based on Mulgan's "Magic game":

Cronus: Cronus is naturally a nasty and vicious individual, who does not wish for good outcomes. He is in the familiar room with two buttons of equal pressing ease. One button removes the debts of 1 person, the other 1000 people.

Cronus, being a despicable, bad-natured person does not want to produce consequences of higher value. In fact, doing so would require a significant amount of mental exertion. Bringing about the higher value consequences therefore comes at a significant effort cost to Cronus. Nonetheless, it is clear that Cronus ought to bring about the higher value outcome. However, PC will not produce this verdict, as both actions improve the world, and are of differing effort levels. Hence, this account of effort will not solve the problem of undemandingness.

Cost: The kind of effort that is of normative significance is the degree of personal welfare that the agent sacrifices. Running down the street comes at no welfare cost to most people, nor does jumping or bending over.

As a response consider another variation on Mulgan's situation:

Apollo: Apollo can press two buttons, of equal pressing ease: one clearing the debts of 1000 people, but slightly increasing his personal debt; two removing the debts of 1 person and not affecting Apollo's personal debt.

Here, PC permits Apollo to choose the option that does not harm his wellbeing, while a slight harm to it would result in much more valuable consequences, so PC is also too undemanding under this account of effort.

Clearly, these examples are not as fleshed out as *Choice*. However, this is not a problem, for I am just demonstrating that merely changing the *kind* of effort will not be an adequate response to undemandingness. It should be easy to see how such examples could be adapted to be more realistic, so here it is only in the interests of space that I prefer simplicity to realism.

Merely changing the *kind* of effort that is appropriate will do nothing to resolve the underlying problem that PC cannot require us to exert *more* effort when we ought to do so. Therefore, I conclude that this response to undemandingness is implausible.

§5b. The effort differences in the counterexamples are negligible.

Perhaps though, the defender of PC can argue that none of the effort differences in the examples thus far shown are compelling. They are admittedly (and intentionally) slight differences, and maybe for the PC advocate too slight to be significant. Plausibly, the ER considers actions of negligible effort difference to be actions of the same effort level. Thus, if my previous counterexamples present negligible effort differences, then none of them demonstrate PC's undemandingness, provided that the ER requires the agent to exert insignificantly more effort when doing so produces more valuable consequences.

There are two main counter-responses to this response to undemandingness. First, unless the advocate of PC puts forward some principled

way to divide negligible amounts of effort from significant ones, there will be an element of arbitrariness to the response. Of course, Jamieson and Elliot (244) profess arbitrariness as the central problem for SC, so this would be a blow to their hopes for PC. However, they may be wrong. Perhaps, a small amount of arbitrariness is not too unsettling, especially in the case of borderline counterexamples. Therefore, if we wish to thoroughly refute this line of argument, we must pursue the second counter-response: demonstrating the undemandingness of PC through counterexamples with clearly significant effort differences.

Let us suppose for a moment that the appropriate form of effort is personal wellbeing sacrifice. As I have shown, changing the kind of effort does not solve the problem of undemandingness, so it matters little which kind we choose. Recall the *Choice* example and suppose the following is also the case:

Jackson must catch a specific bus to a job interview for a coveted position. Jackson can help the worker and still catch the bus, for the two are near. However, saving the lady – who is far away down the street – will mean missing the interview and losing the job opportunity.

Plausibly, losing the job interview is a significant sacrifice of personal wellbeing. Perhaps Jackson is deeply personally invested in getting the job or is in a desperate financial position. On the other hand, helping the worker will require no such sacrifice, and is only a minor exertion. Thus, it seems that the effort difference is now a clearly significant one. However, this would not mitigate Jackson's obligation to save the lady from certain death. Since he ought to save the lady, and the effort difference is significant, PC is implausibly undemanding.

I suspect the initial response from the PC adherent would be to argue that allowing the lady to die would weigh heavy on Jackson's conscience, and that he would suffer a com-parable wellbeing loss due to guilt. Maybe Jackson is a good enough person that allowing the lady to die costs just as much effort as saving the lady. Then, by the ER, Jackson ought to save the lady, as this would improve the world more, and the two actions are of the same effort level.

However, there is an apt counter. Suppose Jackson is an amoral, capricious and unprincipled individual. Then he will suffer no wellbeing cost from letting the woman die. Therefore, as he still ought to save her, and the ER is satisfied, PC is once again implausibly undemanding.

There are now two lines of further response open to the defender of PC. First, they may still claim that the effort difference is negligible. This seems implausible though, for if the earlier claims of negligibility were bordering on arbitrariness, then this claim (in unqualified form) is almost outlandishly arbitrary. Of course, the advocate of PC may give some principled account of negligibility; but how plausible could such an account be, if it sets its lowest level of effort at the intuitively significant cost of losing a coveted job opportunity? However, perhaps the defender may maintain that the difference is negligible by constructing some principled way to define negligibility as a moving baseline, so that negligibility is somehow indexed to the situation. Here I simply note that this seems to be moving too far away from the notion of negligibility, for such a construction would be a principled specification of which effort levels are good enough in any situation, which is very different to the minor claim that two effort levels are simply insignificantly different. Therefore, I shall consider this as a different response in the subsequent section (§5c).

The second further response would be to argue that Jackson is not required to save the lady, for his own wellbeing cost mitigates his obligation to save her. We really don't have to give up our coveted personal projects to save a life that only we can save. This move recognises that the greater the cost of the higher value action,

the weaker our intuition that we ought to bring about the higher value consequences. When the effort difference is zero, the intuition is practically axiomatic, but in the altered *Choice* there is at least some uncertainty, which gives the PC advocate some leeway to make a stand.

However, we should not forget that PC violates a second intuition that is less affected by changes in effort difference[3]. Once the PC advocate accepts that a certain level of effort difference is significant, non-negligible, they are also committed to the claim that no matter how much more good that agent can bring about by exerting this extra effort, they are not required to do so; provided that the lower effort action also improves the world.

Thus, provided Jackson has a low effort, world-improving action available, he would never be required to sacrifice his job opportunity, even to bring about the highest possible good, provided that the effort difference is non-negligible. Jackson might be able to save 1000 old ladies and yet PC would have nothing more to say on the matter. This is what makes PC so implausible: that once a certain amount of effort difference is conceded to be significant, there is no value that this extra effort can bring about, for which PC would require an agent to exert the extra effort.

Maybe Jackson is not required to sacrifice the job for the lady's life. But he surely is required to sacrifice his job for some level of good; for some number of lives, and it is precisely this which PC cannot require of him. Therefore, in this sense, PC is still implausibly undemanding, and hence this counter-response fails.

Claiming that the effort differences in the counterexamples are negligible is not a plausible response because there are other counterexamples with clearly significant effort differences, and the counter-responses to these new examples fail.

§5c. PC can be altered such that it is no longer implausibly undemanding.

The failure of the previous two responses to undemandingness appears to leave us with just one alternative: PC must be altered such that it no longer suffers from the undemandingness problem.

The natural way to do this[4] would be to place some requirements upon effort, for PC does require agents to exert enough. Clearly though, for our requirements upon effort to be practical, they must somehow be relative to factors that change across situations, as our intuitions about the appropriate level of effort change significantly across situations. Thus, we might define the effort requirement as a necessary condition, attached to PC, and of the following form:

Effort Requirement: An action is right only if it exerts enough effort.

The idea here is that "enough" is a placeholder for a later principled specification of appropriate effort level, similar to how the adherent of SC uses "good enough". That is, "enough" stands in for the moving baseline yet to be specified.

By attaching this additional necessary condition to PC, we create a new ethical view, with three conditions instead of PC's two. This new view will be termed Altered PC (APC), and take the following form:

Altered PC (APC):

An action is right if and only if:

i) It improves the world, and;
ii) It satisfies the ER, and;
iii) It satisfies the Effort Requirement

The motivation for APC is the hope that it can raise the demands of PC to a more plausible level, while still avoiding the problems of SC. Presumably, in cases where PC appears implausibly undemanding APC will produce

more demanding verdicts, as a higher level of effort will be required. Additionally, because APC retains the improvement criterion and the ER from PC, it seems that it will avoid at least two of PC's problems; for as we have already seen (§3), the baseline which Jamieson and Elliot advocate appears to rule out cases of prevention of the good, and the ER cases of blatant satisficing.

However, I will now argue that adopting APC is nonetheless an implausible response to the problem of undemandingness.

The obvious worry for APC is that it appears to be an *ad hoc* construction formulated to increase the demandingness of PC. However, I will now present two more powerful objections to APC, both of which also provide additional support for the worry that APC is *ad hoc*. Together these objections form a strong cumulative case that adopting APC is not a plausible response to the problem of undemandingness.

APC Moves too Far Away from PC and Consequentialism

The first objection to APC concerns whether it is truly a form of PC or even a form of Consequentialism. By introducing the effort requirement to PC, APC moves away from the core motivation of PC: that what is ethically important is to improve the world. Like the improvement criterion, the effort requirement is a fundamental, non-instrumental ethical imperative. Thus, under APC we have two equally important fundamental ethical imperatives: to improve the world, and to exert enough effort. This sets APC significantly apart from PC, for it seems to be more of a hybrid than a true form of PC. If this is the case, then adopting PC would not be a plausible move *for the progressive consequentialist*, as it would fail to respond to undemandingness from within PC. If I solve the problem of rights violations by adopting a deontological approach, then I have not provided a response *for the consequentialist*, regardless of the plausibility of the view adopted.

Jamieson and Elliot can tell a plausible story to explain why our fundamental ethical imperative is to improve the world. However, it is hard to see how such a story could emerge for APC's dual imperatives. This lends support to the worry that APC simply is not fundamental, but instead an *ad hoc* construction designed to raise the demands of PC.

A further worry, which Mulgan (How Satisficers 44) has raised against forms of SC, questions whether consequentialist views which appeal to notions of effort are truly consequentialist. The SEP (Sinnott-Armstrong) defines Consequentialism as the view that: "normative properties depend only on consequences". Clearly though, APC does not fall under this definition, as, under APC, all normative properties depend on both properties of consequences and effort properties of the agent, the latter of which are not properties of consequences. In fact, the effort requirement's satisfaction is entirely dependent upon non-consequences, so half of APC's fundamental criteria are non-consequentialist. If APC isn't Consequentialism, then once again, adopting it as a response to PC's undemandingness will not be a response for the consequentialist, and hence also not for the adherent of PC.

APC is Not Genuinely Explanatory

The second objection to APC is adapted from one raised by Mulgan against certain forms of SC. Mulgan (How Satisficers 44) has argued that ethical views which introduce a notion of *enough* effort, as a fundamental ethical criterion, will fail to be genuinely explanatory. While PC can provide a genuine explanation as to why morality is not unreasonably demanding – we are only required to improve the world, which is not very demanding – APC does not provide a genuine explanation as to why morality does not demand very little of us. Under APC, morality does not

demand very little of us because it requires that we sacrifice enough, that we exert enough effort. As Mulgan (How Satisficers 44) puts it: "This is not an explanation", for we might as well say that morality does not demand very little from us because morality demands enough from us.

This further supports the claim that APC is not a system of fundamental significance, but merely an *ad hoc* construction, for we presumably think that such a system ought to provide genuine, non-circular explanations for its demands. That is, a system of fundamental significance is qualitatively distinct from a mere list of intuitions, or a complex instrument which produces verdicts as close to actual intuitions as possible.

Section 5c. Conclusion

Thus, APC appears to be an *ad hoc* construction; its status as both a form of PC and of consequentialism is in doubt; and it fails to provide a genuine explanation for its own demands, the latter two of which further strengthen the claim that it is *ad hoc*.Therefore, we have good reasons to reject the adoption of APC as a plausible response to undemandingness.

Section 5. Conclusion

In this section (§5) I have shown that none of the responses considered (§5{a, b, c}) provide a plausible response to the problem of undemandingness.

§6. CONCLUSION

I first showed (§4) that Progressive Consequentialism faces the problem of undemandingness, which renders it implausibly undemanding in the absence of a plausible response.

I then showed (§5) that none of the responses are plausible, so I conclude that PC is implausibly undemanding, and hence should be rejected. This conclusion is correct only for a human wellbeing account of intrinsic value, but perhaps could be shown for other accounts by employing similar arguments and examples to those developed here.

ENDNOTES

1. They do not specify this exact baseline, but the baseline I give is in the spirit of the vaguer formulation they give (Jamieson and Elliot 247).

2. The ER must also require that agents ensure no action of a lower effort level improves the world more. This seems to be implicitly assumed by Jamieson and Elliot, and I also make this assumption.

3. Strictly, the second intuition applied to 'small' effort differences. However, as noted earlier (§4) the intuition is also strong for significant effort differences.

4. The other way would be to change the baseline for improvement. However, Jamieson and Elliot (246-8) provide strong reasons for the rejection of other baselines, and it is unclear how changing the baseline could solve the problem of undemandingness.

Works Cited

Bradley, Ben. "Against Satisficing Consequentialism." *Utilitas*, vol. 18, no. 2, 2006, pp. 97–108.

Chappell, Richard Yetter. "Willpower Satisficing." *Noûs*, vol. 53, no. 2, 2019, pp. 251–265.

Jamieson, Dale, and Robert Elliot. "Progressive Consequentialism." *Philosophical Perspectives*, vol. 23, 2009, pp. 241–251.

Mulgan, Tim. "Slote's Satisficing Consequentialism." *Ratio*, vol. 6, no. 2, 1993, pp. 121-134.
---. "How Satisficers Get Away with Murder." *International Journal of Philosophical Studies*, vol. 9, no. 1, 2001, pp. 41–46.

Sinnott-Armstrong, Walter. "Consequentialism." *The Stanford Encyclopedia of Philosophy*, Summer Edition, 2019, Edward N. Zalta (ed.), https://plato.stanford.edu/archives/sum2019/entries/consequentialism/. Accessed 30 December 2019.

Slote, Michael, and Philip Pettit. "Satisficing Consequentialism." *Aristotelian Society Supplementary Volume*, vol. 58, no. 1, 1984, pp. 139–176.

Slote, Michael. *Morals from Motives*. Oxford University Press, Oxford, 2001.

Painting and Relation, in Their Existential Significance

ALICIA BADEA

JOHNS HOPKINS UNIVERSITY

Abstract: This paper places Maurice Merleau-Ponty's account of embodiment and painting in dialogue with Martin Buber's account of the existential structure of relation. Drawing on the existential significance of both, I provide an interpretation of Merleau-Ponty's argument in "Eye and Mind" that embodiment serves as a necessary condition to the very possibility of painting's existence, and highlight the key points of Buber's notion of the I-You relation as developed in *I and Thou*. Proceeding from the overlap between the two thinkers' accounts, namely the disavowal of the subject-object structure as the fundamental form of relation and reciprocity as constitutive of subjectivity, I argue that painting expresses the painter's relation to something in the world in a moment of encounter. I conclude with a brief consideration of the import for the spectator, particularly in how painting can illuminate a different mode of being in the world.

What is it that the painter expresses in painting? In "Eye and Mind," Maurice Merleau-Ponty develops an account of the art in which painting expresses and realizes the individual painter's relationship to the world, and thus reveals the human subject's unique position in it more generally. For Merleau-Ponty, this world is, crucially, a visible one. What the painter accomplishes is not divided from his existence as a visible being in a visible world—a fundamental existential structure he shares with all other human beings—but is in fact made possible by and founded upon that very intersection of subjectivity and visibility. Painting is the pure expression and exploration of this reality, interrogates and inverts, as it were, ordinary perception; the painter lives in a mode in which the world shows itself to him *as* visible, not merely as seen. This reciprocal relationship is meant in its deepest sense: it is interpretable and understandable as the encounter of the I-You, or standing in relation, in Martin Buber's *I and Thou.* Buber's con-ception of subjectivity as pri-mordially founded upon and sustained by a mode of being *other* than that of subject-object will illuminate Merleau-Ponty's own account, and will allow us to see that the moment of encounter is what the painter expresses in his work. Once actualized, the painting retains the potential to reveal to the spectator another possible mode of being.

Movement, reciprocity, embodiment, and visibility lie at the core of Merleau-Ponty's understanding of painting. "The eye is an instrument that moves itself, a means which invents its own ends; it is that which has been moved by some impact of the world, which it then restores to the visible through the traces of a hand" (Merleau-Ponty 127). This quotation implicitly conveys, what Merleau-Ponty elsewhere states explicitly, what painting is not. For the *impact* must precede the tracing, in fact, instigates or compels it. Thus a painting is not representation, mere image, mimesis, or a "projection" similar to the one (according to some philosophers) that things in the world project to our mind; and the act of painting is not a simple intellectual and technical exercise in copying the forms and contours of things *out there*, separate from me and self-enclosed, which I come to possess an idea of through my perceptual and cognitive faculties (Merleau-Ponty 133). Against the notions of Descartes, Merleau-Ponty aims to undermine the idea that the lines of this drawing or painting merely happen to 'look like' a tree, because they serve as a clue or hint, a way to trigger the image of a ('real') tree to pop into our minds (131). Paintings do not just "represent" objects. They bear a real relation of resemblance to the things of the world (Merleau-Ponty 131). Resemblance, as Merleau-Ponty uses it here, means more than representation: it suggests both a deep connection to that which the painting resembles *and* accords the painting itself the power of real

presence. Not classified as a "mode or variant of thinking" i.e. of self-sufficient intellectual mastery, painting, for Merleau-Ponty, is "a central operation contributing to the definition of our access to Being" (132). As the visible is integral to Being, in order to understand his conception of painting, we must consider the role and importance he ascribes to vision.

Embodiment is central to vision. "The visible world and the world of my motor projects are both total parts of the same Being" (Merleau-Ponty 124). The space of my agency and the world I see overlap, seemingly with no guarantee that this be the case—for what within the visible necessarily entails or secures that it can be moved or affected by me? In this would reside something almost miraculous, Merleau-Ponty suggests, if not for the fact that "vision is attached to movement," in two fundamental ways: the eye only sees because it constantly moves; and bodily movement prefigures vision (124). Vision does not issue from the workings of a separate mind which merely 'sits in' and controls a mechanical body, as a Cartesian might have it. I see because my eyes are part of my body, *because* I am embodied. In a similar way, the mind is not responsible for movement as a conscious choice then executed in space, as if it thinks to fire neurons and contract or relax the required muscles (Merleau-Ponty 124). My hand reaches toward a cup, a book, or a flower, and it simply *happens,* without a decree or deliberation issuing from the depths of internality. Of course, we can deliberate on a certain gesture or action's ethical and practical considerations, and perhaps the *manner* of doing it (to press softly or firmly, caress comfortingly or seductively), but movement's 'how,' its fundamental coming about, is always steeped in our embodiment in a visible world. Movement is as if a response not simply to what I see, but the fact *that* I see at all—that there is a world to move in, with things to move towards. In this way, vision and movement come together in the body (124). With this in mind, we begin to grasp what it means for Merleau-

Ponty to say that "vision is caught or comes to be in things" (125). Yet the other side of this requires undermining the second aspect of the legacy of Cartesian mind/body dualism—the primacy of the traditional subject-object view of the world.

Reestablishing the centrality of embodiment already gives way to reimagining the human being's position in the world and toward things in it by illuminating the paradox of see-er and seen. For once I no longer conceive of myself as a mind in possession of a body, I am a self *as* body and mind, inseparably, as one. I see things; I also see myself. Suddenly, I realize I am not 'in here,' and the world is 'out there'—I am *part* of the world, "my body is a thing among things" and "the world is made of the very stuff of the body" (Merleau-Ponty 125). This point entails: 1) that as an embodied being, I am visible to others, and so am *seen* by the world, in addition to by myself; and 2) vision "must somehow come about" in things, that there vibrates a resonance of visibility and invisibility across the body and the world (Merleau-Ponty 125). All of this disallows us to recast the human being's position as merely *sometimes* an object for others, and always a transparent subject for himself. For this dual cross-visibility marks a continuity with the world that resists the traditional classification into subject-object. I am not so self-contained and separated from the world that I can possess and discard it at will. I already overflow into it, reaching beyond the contours of my body with vision and with movement, and find that I understand myself not as master of the visual field, but only in relation to things in it. The bird which spots me from the window-sill is not a subject, and I its object. Rather we are held together in visibility and embodiment. Neither am I transparent to myself. For my body is not an object, but a "place the soul inhabits," not only as its most intimate space, but as that with which the soul thinks and exists (Merleau-Ponty 136). For Merleau-Ponty, it is just this consciousness of being at once see-er/seen and sensing/sensed

that constitutes selfhood and establishes a relation of reciprocity with the world—an idea which we will return to shortly.

We are now in a position to understand Merleau-Ponty's quip that "[i]ndeed, we cannot imagine how a *mind* could paint" (123). His remark points, firstly, to the almost startlingly obvious fact that the painter paints with his body: his hands hold the brushes and the palette, adjust the easel; the movements of his arms, wrists, and fingers (among other body parts) *produce* the marks on the canvas; he sits or stands, hunches over, moves closer to or farther away from the canvas. A disembodied cognitive existence does not birth a painting whole. Not to say, of course, that this is how most imagine it; but the exaggeration serves to throw into relief the crucial lack latent in the common idea that painting begins with a mental picture. This conception, depending on its inflection, erases or subdues the significance of painting being fundamentally *also* a physical act. The painter does not arrive to the canvas with a preformed image in his mind which he must find the material means to portray, a conception which would thereby render all models (the fruit of still-lives, the people of portraiture, the natural elements of landscapes) merely *references* rather than the partners of a living engagement. The painter's body moves through space. The work results from the interaction, struggle or cooperation as the case may be, between painter and canvas, and between painter and that which touches his eye. Painting is inseparably material. On the painter's end, it is inseparable from his embodiment.

This inseparability does not remain on the purely contingent level of requirement. It is not the case that it just so happens that we, as physical beings, need our bodies to paint. Rather, painting is inseparable from human embodiment because it is only due to our position as embodied *visible beings in a visible world* that we paint at all. The intersection of subjectivity and visibility gives rise to a reciprocal relationship with the world:

Quality, light, color, depth, which are there before us, are there only because they awaken an echo in our bodies and because the body welcomes them. Things have an internal equivalent in me; they arouse in me a carnal formula of their presence. Why shouldn't these correspondences in turn give rise to some tracing rendered visible again, in which the eyes of others could find an underlying motif to sustain their inspection of the world? Thus there appears a "visible" to the second power, a carnal essence or icon of the first. Merleau-Ponty 125-126

Merleau-Ponty finds himself here at the limit of language. What his thought suggests, however, is a resonance between body and world. The world touches me, and I am receptive to it; I touch the world, and find *it* is receptive to me. Merleau-Ponty reestablishes the centrality of *betweenness*: Being is not only '*I* act, *I* receive,' the crowning and entrenchment of the 'I' perspective (which would be the subject-dominating-object view of the world), but fundamentally also that which acts upon me, can open to me. The possibility of mutual openness and receptivity is constitutive of our being in the world.

In this passage, Merleau-Ponty specifically draws attention to the movement of this communication, as it were, with things in the world. The painter stands before a visible thing: it moves him; its impact and what it arouses in him is invisible. It is just this invisible in-betweenness that the painter then seeks to "restore to the visible through the traces of a hand," to render visible *again* the world's invisible trace in him. This will not be a copy of the thing, not its trivial likeness or imitation (if it has any likeness at all), for there is a change which occurs in the passage. The world, in a sense, passes through the painter; but neither is he just a filter through which the world expresses itself with the unique coloration of his perspective. Perhaps we can draw an analogy, if limited, to a dialogue.

The thing expresses itself to the painter. Feeling that expression awaken something in him, he holds that and mulls on it. He speaks back into the world *in* painting, not the thing itself nor a repetition of what the thing said to him, but precisely (what he can of) [1] this invisible interaction, in the language of the visual. Thus the painting contains and conveys the "carnal essence," which is neither direct transcription nor mere modification of the thing's first expression, but rather a kind of echo and incorporation of what is most fundamental of it into the painter's experience of the world. In using the word "icon," Merleau-Ponty emphasizes that the painting is not a compilation of traces of the visible-invisible-visible movement, but is its realization, and has a real presence in itself. We can then understand "carnal formula" in relation to what Merleau-Ponty later calls the "system of equivalences, a Logos of lines, of lighting of colors […]—a nonconceptual presentation of universal Being" (142). In this silent language of the visual the painter attempts to make sense of the ineffable confrontation with Being, the otherwise inexpressible relation to things in the world. All of this underlies what Merleau-Ponty means when he says, "It is by lending his body to the world that the artist changes the world into paintings" (123). The intertwining of embodiment, subjectivity, and visibility renders possible the restoration of the invisible to the visible in painting.

It is at this juncture that we can turn to Martin Buber's conception of relation and encounter. Merleau-Ponty's account of reciprocity and the reconfiguration of the traditional subject-object structure is unique in its details, but not its broad contours. In *I and Thou*, Buber

reconceptualizes the role of just this structure, and seeks to reveal a human being's originary existence in the world as fundamentally one of relation. For Buber, "The world is twofold for man in accordance with his twofold attitude" (53). Essentially, a human being adopts one of two fundamental existential orientations towards a world: in the saying of I-You or of I-It, which Buber terms basic word pairs (53). This 'speaking' is not a literal verbalization but refers to the fundamentally existential *manner* in which it orients the human being. The alternation of the two characterizes human life.

Buber's point is structural: each basic word pair "establish[es] a mode of existence" (53). In the case of I-It, this structural mode is precisely that of subject-object, where the "I" is the self-contained, knowing subject and the "It" is the classifiable, analyzable 'object,' whether that be an inanimate physical entity, an organism in nature, or another human being. This I knows, thinks, feels, imagines, perceives, and senses some*thing*; the I always mediates the object, breaks it down into its properties or parts so as to be processed instead of purely lived—the world of experience,[2] as opposed to encounter (Buber 55-56). In limiting 'experience' to the It-world, Buber emphasizes the activity-driven and goal-oriented character of so much of human life and interaction in and with the world: the It is an object of experience subordinated to the I, assimilated and absorbed by the I, into memory, perception, and mental or emotional life, never simply acknowledged. Because this I "appears as an ego [3] and becomes conscious of itself as a subject (*of* experience and use)" (Buber 111-112; italics mine), it derives knowledge or a 'product'

---

[1] This parenthetical is intended to qualify *what* he 'speaks': for I do not think Merleau-Ponty implies it is possible to 'transcribe' this invisible interaction, as if it is a simple matter of perfectly turning what is unseen into a visible graphic rendition. The painter does not serve as a scribe, but attempts to convey the realness of this relation to

[2] Buber uses "experience" to refer to any occurrence that has an object, whether that be an emotion, perception, sensation, etc. I always experience *something*, whereas in encounter, I do not "experience" the other, but *am* with them.

[3] This usage bears no relation to the Freudian ego, as translator Walter Kaufmann explains in footnote 7 (111-112).

from experience—there is always something it can say about what it has just experienced. The It-world "permits itself to be taken by you, but it does not give itself to you" (Buber 83). The I as ego exists adjacently to things, separate from them in such a way as to be able to possess or to scrutinize them, but they do not open themselves to the I, just as the ego of the I-It is only ever impartial, withdrawn from the *wholeness* of being (Buber 54).

The mode of the I-You, on the other hand, is relation. The I does not experience the You; the You encounters[4] the I (Buber 55, 62). The You is unsubsumable, unbounded, and unanalyzable. Whereas the things of the It-world appear as if "constructed of their qualities" and possible to be broken down into parts and pieces (Buber 81), the I of relation encounters the You in its wholeness. This does not mean that, to use Buber's example, I do not see the leaves, roots, and greenness of the tree before me, or forget the existence of its chemical processes. Rather, I no longer *pick out* an aspect on which to concentrate, no longer perceive it as an amalgamation of aspects. All its particularities are "included and inseparably fused" (Buber 58). This indivisible wholeness of the You mirrors and, Buber suggests, is only possible by the I's "essential deed": the entering into relationship with one's "whole being." "The concentration and fusion into a whole being can never be accomplished by me, can never be accomplished without me. I require a You to become; becoming I, I say You" (Buber 62).

Thus we arrive at two of Buber's key points: "Relation is reciprocity" (67); and person-hood is constituted by relation. Buber conceives of reciprocity as fundamental to the encounter because standing in relation occurs only when the I gives itself over entirely, withholds no part of its being, and is met by a You that opens itself in the same way. Neither partner can force the encounter to happen, and yet each depends upon the other's entering into relation to be acknowledged as a whole being. The I cannot seek the You, for it cannot find, possess, or take hold of a You as it does an object, as a means rather than an end in itself (Buber 62). It is this very absence of an object that allows Buber to situate the emergence of personhood within the relation: "The I of the basic I-You appears as a person and becomes conscious of itself as subjectivity"—*not* as subject (112). For while an I never exists outside one of the two modes of existence, the I, in a primordial sense, *first* realizes its own existence as a distinct being in encountering a You: the I as person is irreducible, just as the You is. Only after the emergence of the I in this way does its detachment from the You, its removal from relation, and thus the mode of I-It become possible. Once split from the reciprocity of the You, the I becomes conscious of itself as subject, as "the carrier of sensations and the environment as their object" (Buber 73-74). We return to the basic word pairs' fundamental difference: experience is 'in' the I as ego, while relation is *between* the I as person and the You of the world (Buber 56).

Singularity pertains to the encounter with a You. [5] Buber distinguishes three different "spheres in which the world of relation arises": with nature, with human beings, and with "spiritual beings," which he conceives of as invoking the creation of art (56-57). It lies outside the scope of this essay to determine the merits and shortcomings of these distinctions, but, while each has its peculiarities, they share an

---

[4] Buber at times also uses (what Kaufmann translates as) "confrontation" or "confronts" to speak of what occurs in the I-You. However, I will consistently use "encounter" in order to emphasize the unforced or unexpected nature of the I-You relation, and to deemphasize connotations of conflict or aggression.

[5] There is an explicitly theological dimension to the relation with the You Buber develops with the idea of the eternal You in the third part of *I and Thou*. For the purposes of this essay, I will bracket this part of his account, as the salient points remain applicable without it.

essential structure that is of primary interest here. [6] In an encounter, the "power of exclusiveness has seized me" (Buber 58). Not only does a being's specific features fuse into a wholeness, but, in that moment, this being has primacy over all. It ceases to be a thing among things. "You has no borders" (Buber 55); it shapes the contours of the world. The You is "neighborless and seamless […] Not as if there were nothing but he; but everything else lives in *his* light" (Buber 59). Other beings and material objects continue to exist; but so intensely and intimately does this being, having become my You, call upon me, that it compels me not merely to concentrate my whole attention, but to *give* my being in its entirety to this communion. This constitutes the foundation of reciprocity—"My You acts on me as I act on it" (Buber 67). Buber lists such seemingly evident examples as students teaching teachers and creative works forming their artists, but trite they are not (67). For underlying each, we find a mutual influence and affirmation of existence which extends from me to the structurally privileged being before me, and from this being to me.

The elements of exclusiveness and mutual acting-upon emerge as features of the structural basis of the encounter: the ideas of the present, presence, and actuality. Buber remarks that the You "appears in space" and "in time," but unlike the It which is confined or caught up in the fabric of these dimensions (81). Rather, as noted, the You is spatially 'unbounded,' in that its presence and exclusiveness turns all else into a background. Even more essentially,

the You appears in time, but in that of a process that is fulfilled in itself—a process lived through not as a piece that is a part of a constant and organized sequence but in a 'duration' whose purely intensive

dimension can be determined only by starting from the You. It appears simultaneously as acting on and as acted upon, as, in its reciprocity with the I, the beginning and end of the event. (Buber 81)

The You rounds the limits of the moment from the inside. For human beings, the encounter gives time its shape and tenor. Buber's conception fundamentally opposes the notion of a linear time composed of and able to be broken down into infinitely small points. The present is not a brief dot on a calculable time line. 'Present' names what is most *real*, what is most intensely lived. It exists only *because* encounter and relation do (Buber 63). The presence of the You establishes the present (Buber 63). The two are not identical, but they are inseparable. For Buber, presence is not fleeting, nor simply standing still, "but what confronts us, waiting and enduring" (64). In this way, the You liberates the notion of the present from its sense of ephemerality. When I stand in relation, the eternal marks the present of this moment. "What is essential is lived in the present, objects in the past," for they "consist in having been" (Buber 64). The It-world resides solely in the past, for an object never addresses me; the I as ego has already subsumed and processed it. It retains no presence, and thus no actuality. [7] "Whoever stands in relation, part-icipates in an actuality; that is, in a being that is neither merely a part of him nor merely outside him" (Buber 113). Actuality draws together the present, presence, and the essential act of entering-into-relation: it is life as lived in the moment of encounter, when the I as ego does not subjugate or brush up against something in the world, but the I as person exists *with* what it encounters, not as idea or abstraction, but *in itself*—the You "confronts me bodily" (Buber 58).

---

[6] Buber himself does not seem to intend for his brief elaboration of each to establish a divide in *kind*, but rather to suggest the possible variations of how an encounter may be lived, which can be cross-referenced to illuminate what is singular about the relation with the You.

[7] Kaufmann notes in the prologue: "Buber's persistent association of *Wirklichkeit* with *wirken* can be carried over into English to some extent by using 'actuality' for the former (saving 'reality' for the rare instances when he uses *Realität*) and 'act,' in a variety of ways, for the verb" (45-46).

Reciprocity as central to subjectivity emerges as the key points of overlap between Buber and Merleau-Ponty. There exists a reciprocity between human being and world, that involves a mutual acting upon such that it is both active and passive at once (touching/touched, seeing/seen for Merleau-Ponty; entering into relation as an action of choosing and being chosen for Buber [62]). This reciprocal relation constitutes subjectivity: for Buber, it specifically based in the I-You encounter; for Merleau-Ponty, the human being as embodied understands himself only in relation to things in the world. And although Buber does not explore embodiment itself, he underscores the I-You encounter as one of bodily confrontation. The notion is not physical, but is steeped in the idea of actuality, with its emphasis on presence. "What I encounter is neither the soul of a tree nor a dryad, but the tree itself" (Buber 59). I encounter an *actual world*—the visible, sensible world, in the parlance of Merleau-Ponty—not a Cartesian or idealist world of images or appearances in the mind.

The import for the painter is this: what he restores to the visible in painting *is* the relation in the moment of encounter with the You. The relation is the invisible, for it is always what is *between* him and something in the world which in the encounter appears as wholly *other* than object. For, as Buber says, in this moment, I *have* nothing, but I stand in relation (55); the painter does not possess an object, mental or physical, but encounters and is encountered by that which he cannot subsume, which he must live in its wholeness and fullness. "Immersed in the visible by his body, itself visible, the see-er does not appropriate what he sees; he merely approaches it by looking, he opens onto the world," a line from Merleau-Ponty which echoes Buber (Merleau-Ponty 124): "All actuality is an activity in which I participate without being able to appropriate it" (Buber 113). In the moment when something has become a You, the question of mastery disappears. Art is neither a graphic

rendering of space and objects, nor an overflow of 'pure subjectivity' into materiality. What the painter does is express this relation to the You in painting.[8] No one line, color, figure, or contrast of the painting conveys this, and neither do any sets of compositional techniques or practices. It is not schematic. We cannot divide the relation into parts; in the same way, the painter does not perform a "translation" of it, as if the lived relation were a foreign tongue to be translated into the language of the visual. The painter does not translate. He *responds*. Merleau-Ponty's "system of equivalences" as a "nonconceptual presentation of universal Being" then refers to just this language of the visual, as the mode through which we as human beings make sense of Being, and the way by which we can enter and access the realm of the You again. Relation resists any on-to-one correspondence; rather, the painting as a *whole*—its brushstrokes, its color, its vibrancy, its lightness or darkness, together— conveys relation. In painting, the way the world is seen reveals the way it is related to.

Thus embodiment, vision, and the painter's individuality are central to the creative act. Buber writes: "He listens to that which grows, to the way of Being in the world, not in order to be carried along by it but rather in order to actualize it in the manner in which it, needing him, wants to be actualized by him—with human

---

[8] To me it seems that both (1) the instigating moment and (2) the creation of the work can be I-You encounters. By (1), I refer to a moment before the painter begins a project, which afterwards compels him to paint, or to which he returns to in memory at a later date. That is, a moment of relation need not be immediately followed by painting, but painting always seeks to express a relation. By (2), I refer to each time a painter returns to work on a painting, as that duration itself being an encounter. The specificities of (2), and the connection between (1) and (2) would require further elaboration; however, I do not propose here a full structural account. My main point is to establish the painter's work as expressing a relation to the something in the world in Buber's sense, whether that be the painter's meeting of Being in the visible akin to (1), or a combination of (1) and (2).

spirit and human deed, with human life and human death" (109). Although Buber is referring more generally to the "free man," his description applies equally well to the painter. Buber accords a primary significance to our humanity. We as human beings hold a special relationship to Being, which we live particularly in the moments when something ceases to be part of the It-world. In the context of Merleau-Ponty's account, this is inseparable from our embodiment. We return to the idea that the painter paints *as* a human being, because embodied. What this means, specifically, is that "the painter's gaze asks them [light, lighting, shadows, reflections, color] what they do to suddenly cause something to be and to be this thing, what they do to compose this talisman of a world, to make us see the visible" (Merleau-Ponty 128). Merleau-Ponty in fact uses "vision" to mean more than basic perceptual seeing (which he calls "profane vision"), as this passage already attests to: for these 'objects' (light, etc.) exist only in the visual field, and they are not "ordinarily seen" (128). He extends this idea of the not-ordinarily-seen to include the invisible: "This voracious vision, reaching beyond the "visual givens," opens upon a texture of Being of which the discrete sensorial messages are only the punctuations or the caesurae" (127). Thus the painter's vision, learned and not immediately given, remains sensitive to what is most fundamental about the human condition. If one of the deepest mysteries is *that* there is a world at all, the painter's interrogation of the visible, and particularly of those intangible aspects which make the visible world a world for us (light, color, etc.), reflects at the same time an exploration of our very access to Being, in all its dimensions. These we never experience discretely, but find an opening to in encounter. The painter's particular individuality, as a person, informs *his* relation to the world, and how *he* will live each specific encounter with a You, and thus his painting will contain and convey this singularity.

What does this entail for the spectator? If vision is a way of seeing the world, meant in its profoundest sense as already a way of standing in relation, then to see a painting is to open onto a new mode of being. The painting itself, once actualized, becomes a part of the It-world: "That which confronts me is fulfilled through the encounter through which it enters into the world of things in order to remain incessantly effective, incessantly It—but also infinitely able to become again a You, enchanting and inspiring," Buber says specifically about artworks (66). Thus the painting does not necessarily remain an object, in the structural sense. I approach it, stand before it. Suddenly, it opens itself up to me and I am drawn into the world of the painter. "Rather than seeing it, I see according to, or with it" (Merleau-Ponty 126). It ceases to be a subject-object experience; it is, in one way, a revelation. It opens for me new possibilities. It gifts me the painter's particular vision. If this vision resonates with me, I take on this new mode of being in the world, which can increase my access to Being; and perhaps I find myself more in the existential orientation of the I-You, and thus more in encounter. None of this is to suggest, however, that the painting itself functions merely as a lens into this, the painter's, mode of being, that it rests solely as a means to be discarded once the end is attained. *Its* particularity and material existence remain important. For it, too, can become a You for me. The painting addresses me, makes a claim upon me; it seems that I stand in relation to the meanings imbued in it. Perhaps it is just this encounter with the painting itself which renders available the second, deeper layer—the artist's vision, the expression of his relation. Thus, for the spectator, a painting both brings to light a new way of seeing, and so a different way of being, and can be the You in a moment of encounter.

Ultimately, Merleau-Ponty's account of painting in "Eye and Mind" aims to consider how art reveals and brings us closer to Being. He undermines the notion that art is artifice, copy,

or imitation. The painter seeks to convey in his work nothing less than his way of seeing the world, which is to say: his way of being in the world. Painting expresses the moment of relation, when the painter encounters something as a You: ineffably, exclusively, in its wholeness, and as that which inflects time. If, as Buber holds, the I-You relation constitutes the foundation of personhood, in expressing this, painting returns us to the very roots of subjectivity—and so, for Merleau-Ponty, to our embodiment. Embodied being, the painter roves his eye over the world until something latches, and he is caught in a relation in which the world touches him, and he touches it. The world having traced itself invisibly in him, his own hands restore for us the unseen possibility of a real relation to Being in the language of the visual.

Works Cited

Buber, Martin. *I and Thou,* translated by Walter Kaufmann, Simon & Schuster Inc., 1970.

Merleau-Ponty, Maurice. "Eye and Mind." *The Merleau-Ponty Aesthetics Reader: Philosophy and Painting*, edited by Galen Johnson and Michael Smith, Northwestern University Press, 1993.

Three Visions of an Ideal World: Kant Against Rival Idealisms

JOE BROWNSBERGER

YALE UNIVERSITY

Abstract: This paper seeks to show that Kant's Refutation of Idealism succeeds in proving the existence of a world apart from the self, while noting that this conclusion does not disprove Berkeleyan idealism, which posits an illusory world with an external source. In making my way to this conclusion, I present two step-by-step interpretations of the Refutation of Idealism. The first relies on the necessity of a persistent ground for our representations of the world, and while a more literal reading of the text it succumbs to objections. The second interpretation brings to light a vicious circularity caused by reliance on self-caused representations, and though a less literal reading it succeeds in proving the existence of some sort of external world.

Immanuel Kant's Refutation of Idealism within the *Critique of Pure Reason* is generally seen as his attempt to distance his Transcendental Idealism from traditional forms of what he calls "empirical idealism," such as that of Berkeley or of Cartesian philosophy. To do this, he aims to prove that the very fact of our interior experience, undoubted by the most thorough of skeptics, would be impossible were it not for the existence of a world of external objects which we perceive as in space. In this paper I examine the force that Kant's argument has against its two opponents: the solipsist, who denies that we can know of the existence of an external world or even rejects its possibility, and the Berkeleyan idealist, who admits the existence of such a world but denies that what we perceive corresponds to it in any way, attributing the world of our perception to a great illusion on the part of God or another external thing. In this paper I provide two interpretations of the Refutation of Idealism. The first, though a more literal reading of the text, ultimately fails to prove its point against either opponent, while the second, I argue,

succeeds against the solipsist but remains powerless against the Berkeleyan idealist. After discussing both of these reconstructions of the argument along with their strengths and failings, I will propose that Kant limits his scope, in the introduction to the Refutation of Idealism, to a refutation of *solipsistic* idealism. In this way his argument retains full force against its opponent and achieves a significant result, even if it cannot effect the hoped-for separation between him and Berkeley.

First Reconstruction: The Argument

One of the ways in which it is possible to reconstruct Kant's argument is as follows:

1. I am determined in time.
2. Determination in time requires something persistent in perception.
3. The two forms of perception are space and time.
4. Time cannot be perceived in itself and is only determined through a spatial analogue.
5. Therefore, time-determination must occur through the perception of persisting things in space (from 2, 3, and 4).
6. Any perception of persisting things in space that allows me to determine myself in time would be a representation.
7. All representations require another persistent thing apart from them to be determined.
8. Therefore, there is a persistent thing prior to perception that allows the determination of (i.e., grounds) my representations of the spatial world (from 1, 5, 6, and 7).
9. This prior-to-perception persistent thing is either outside me or inside me.
10. It cannot be in me.
11. Therefore, there is a persistent thing outside me which enables the determination of my representations of space and

therefore the determination of myself in time (from 8, 9, and 10).

The argument starts from the basic claim that I am determined in time. Of this fact I am immediately aware; it is what enables me to say of myself that I existed in some way in the past and exist in another in the present. Premises 2 through 5 come from the Analogies, and rest on the concept of a persisting substance. Though I will not here rehearse the entire argument for the premise's claim, one line of reasoning that supports it is found in the First Analogy. In a world in which everything is always changing, and nothing persists, each point at each moment is completely disconnected from every other point-moment (though in reality to speak of a "moment" here is a misuse of terminology). We will call a point-moment an "event." In this world of total and eternal change, instead of a continuum of connected events we would have a sequence, in which "existence is always disappearing and beginning, and never has the least magnitude" (Kant A183/B226).[9] It is only when we introduce something that persists through change that we can relate events to each other, forming a continuum and allowing for the determination of time. Thus, "that which persists is the substratum of the empirical representation of time itself, by which alone all time-determination is possible," (premise 2) (A183/B226).

Now, "time cannot be perceived in itself" and thus by itself it cannot provide this persisting thing (A181/B225). We must look elsewhere: "Consequently it is in the objects of perception, i.e., the appearances, that the substratum must be encountered that represents time in general and in which all change or simultaneity can be perceived through the relation of the appearances to it" (premise 4) (A181/B225). Since we perceive all appearances in space, "appearances"

here means spatial intuitions. The self cannot be the requisite persistent appearance because "[t]he consciousness of myself in the representation *I* is no intuition at all, but a merely *intellectual* representation of the self-activity of a thinking subject" (B278). In sum, we need not only a persisting thing, but a persisting thing in space in order to determine ourselves in time (premise 5).

The idealist could agree with everything said thus far. Neither the Berkeleyan nor the solipsist denies that we perceive what we perceive; rather, he or she merely posits a different source for those perceptions. We may need a persisting thing in our perceptions to make sense of ourselves and the world, but those perceptions could very well be contained within myself, as they are in dreams. To eliminate this possibility, we move to the second stage of the argument.

Premise 6 states that any perception of ours, like the perception we use to notice the necessary persisting spatial objects, are representations. This is nothing more than a definitional claim: on B376 Kant says, "The genus is representation in general *(repraesentatio)*. Under it stands the representation with consciousness *(perceptio)*." Kant goes on to say that representations "require something persistent that is distinct even from them, in relation to which their change, thus my existence in the time in which they change, can be determined" (premise 7) (Bxxxix). This seems like a reasonable claim—take each frame of a representation individually, and there is no necessary relation to every other frame; in other words, there is no order inherent in representations alone. To take one of Kant's examples, given an image of a ship upstream and a ship downstream there is no way of knowing, solely on the basis of what is contained in the images, which comes first. There must then be a second persistent thing that grounds the perceptions and puts them in the correct relationships to each other (premise 8). This persisting thing is not to be confused with the persistent thing in

[9] All following in-text citations refer to Kant's Critique of Pure Reason, as cited below.

perception—Kant says that this is "distinct even from [the representations]" and so cannot be something within them. Rather, this persistent thing is "behind" the perception, so to speak. Like a table on which are arranged photographs, this persistent thing underlies representations; it is prior to them and enables them. Unlike the table, it also determines the order in which the representations appear.

Next Kant must show that the prior-to-perception persistent thing is outside me and not inside me (premise 10). This claim rests mainly on arguments only fully worked out in the Paralogisms later in the *Critique* (B406–B432). There Kant works his way to the conclusion that since we cannot have an intuition of the self as it is, and therefore also can have no cognition of it, we cannot know the self to be a substance, that is, a persisting thing. Therefore, if the required persisting thing prior to perception is not myself or something within me, then it must be outside me (premise 11). Note, however, the gap between our lack of knowledge about the substantiality of the self proved in the Paralogisms and the positive claim that the self is not a substance—this will be returned to later on. Disregarding this worry for the time being, we have proved that there is something outside me that enables the determination of my perceptions, i.e., that grounds them. And by grounding my persistent spatial perceptions, this thing allows me to determine myself in time. In the empirical realist model to which Kant ascribes, it is external things that ground my perceptions by giving them reference to something. Our perceptions of the sun, for example, are put into proper relation with each other because there is an actual sun (in some form) to which they correspond. A perception of the sun rising precedes that of the sun setting because that is the way it is in the external world. On this view, without actual external objects we would not have any basis for our representations being ordered and therefore would not perceive

anything as persistent. Consequently, we could not determine ourselves in time.

Objections to the First Reconstruction

It is in the aforementioned gap between the arguments of the Paralogisms and the claim of premise 10 that primary problem of the argument lies. The argument in the Paralogisms only proves that we cannot know whether or not the self is a substance. It does not, however, definitively prove that it is in fact not a substance. If it is possible for the self to be a substance, then premise 10 fails, and so does the argument. Kant might attempt to plug the gap by saying that if the self is dependent on the representations, and the representations are dependent on the self, then we run into circularity. But it does not seem like this is necessarily vicious circularity—it could be the case that the unperceived self grounds the representations which are perceived. Kant himself draws a hard line between inner sense, by which we perceive the self as it appears, and apperception (see B153). Thus, it would seem the argument of the Refutation of Idealism, at least in this form, does not definitively prove anything against the solipsist.

The argument does not fare much better against an idealist like Berkeley, who posits God as the ground of perception. In Berkeley's view God immediately gives us each one of our perceptions. Most idealists hold some variant of this position, claiming that there is something external to us that gives us our perceptions, which nevertheless do not correspond to the true nature of the external world—this is true from Descartes' evil demon to the computer program in the 1999 film *The Matrix*. Now the conclusion of Kant's argument—that there is an external persistent thing that grounds our representations—is perfectly compatible with this idealist picture. Kant's view is, of course, noticeably distinct from Berkeley's in that Kant thinks that when we perceive things, those perceptions are

actually perceptions of external things in some way; appearances are actually apparitions of something. However, they are perceptions that are filtered through our human sensibility, thus taking on the forms of space and time which do not pertain to them in themselves. Kant explains this in the Transcendental Aesthetic:

> If I say: in space and time intuition represents both outer objects as well as well as the self-intuition of the mind as each affects our senses, i.e., as it *appears*, that is not to say that these objects would be a mere *illusion*. For in the appearance the objects, indeed even properties that we attribute to them, are always regarded as something really given, only insofar as this property depends only on the kind of intuition of the subject in relation of the given object to it then this object as *appearance* is to be distinguished from itself as object *in itself*. (B69)

Thus, Kant asserts that our perceptions in some way correspond to actual external things, that those things are the cause of order among our perceptions. Meanwhile, for Berkeley the representations that we perceive have no necessary correlation to things as they are in transcendental reality, God being the persistent ground of all perceptions directly. However, both positions are equally valid if the only criterion is that there is a persistent thing outside me that grounds perception—God can act as the persistent thing that produces and orders all representations in us just as well as an external world of objects can.

The Revised Reconstruction
The Argument

If this interpretation of Kant's argument in the Refutation of Idealism fails to prove itself against either of its opponents, solipsism or idealism, we must try to find another way to interpret it. One way to do this is to take a different tack in answering the central question of why representations alone cannot allow my time-determinations. Instead of positing two types of persistent things—the perceived persistent thing and the prior-to-perception persisting thing—we could do away with the second entirely, along with the argument following from it. We replace it with the premise that if we are to cause the change and persistence of our representations, we must already be determined in time. This new reconstruction, enumerated below, is the same as the first reconstruction through premise 5.

1. I am determined in time.
2. Determination in time requires something persistent in perception.
3. The two forms of perception are space and time.
4. Time cannot be perceived in itself and is only determined through a spatial analogue.
5. Therefore, time-determination must occur through the perception of persisting things in space. (from 2, 3, and 4)
6. These perceived spatial objects are either outside me or inside me.
7. If the perceived spatial objects were something in me, they would be representations contained entirely in me, and I would cause the change and persistence that I perceive.
8. For me to cause change and persistence in my representations presupposes the determination of myself in time.
9. Therefore, the perceived spatial objects which enable me to determine myself in time would presuppose the determination of myself in time—a vicious circle. (from 5, 7, and 8)
10. Therefore, the persistent spatially perceived object cannot be in me and must instead be outside of me. (from 6 and 9)

11. Therefore, there exist objects outside me that we perceive as persisting in space. (from 1, 5, and 10)

In the revised reconstruction of the argument, Kant sets up a dichotomy after premise 5: the objects that we perceive as outside us in space could be actually outside us, or they could be inside of us (premise 6). Now if the perceived spatial objects were inside me, they would exist as representations produced and perceived entirely within myself (premise 7) (Bxxxix). On this view we would create a representation of a persisting substance (and its alteration) for ourselves; perceiving this we would be able to determine ourselves in our self-created world of space and time. But to do this we would have to make our representations change in an ordered, connected way. This is necessary not only from an *a posteriori* perspective (since we do in fact perceive such change), but from an *a priori* one as well. Change is required for us to "perceive" time and therefore determine ourselves within it—if nothing at all ever changed, either inside us or outside us, we could not even form a concept of time. And of course, for things to change in the way that enables our time-determination, there must also be something persistent, as described above.

But if our representations are entirely contained within ourselves, then *we* are causing our representations to persist and change in this ordered manner, making them move from one state to another and situating them with reference to each other. In short, we are determining them in time (premise 8). And if the representations are depending on me to be determined in time, and I am dependent on them to determine myself in time, we have vicious circularity (premise 9). Thus, the entire line of argument on this side of the dichotomy, which stems from the supposition that the spatially perceived objects are in me, is faulty and must be rejected. This leaves us with the conclusion that the objects we perceive in space are actually outside of us.

The Necessity of Matter

This view, naturally, has its own weaknesses. For example, if we take premise 8 to mean that we cannot impose a framework of time on our representations, then we would be led to reject Kant's thesis that things in themselves are outside of space and time and that we impose those forms on them as we perceive and cognize them. Rather, premise 8 must mean that I cannot provide the *matter* of change and persistence for myself; I cannot represent space and time if I have no external source of information to work with. There must be something external to us which we translate as change and persistence; in this way we are not the ones causing the persistence, but rather persistence in space is how we perceive the external world. In this manner we can reconcile the claim that we cannot cause change and persistence with the claim that space is not a property of things in themselves.

This concept is further elucidated in what Kant says about the imagination. He distinguishes the "transcendental synthesis of the imagination" from the "empirical synthesis of the imagination," noting that the latter is only reproductive and subject to empirical laws of mental association, while the first is a central aspect of our cognition and is, in a sense, productive (B152). However, this "productive imagination" is only productive in that it is actively caused by me and not passively perceived: "insofar as the imagination is spontaneity, I also occasionally call it the productive imagination, and thereby distinguish it from the reproductive imagination" (B152). This productive imagination, or "transcendental synthesis of the imagination," is "an effect of the understanding on the sensibility" and it exercises its spontaneity by being the faculty that "connects the manifold of intuition" (B152, B164). This is a vital role: the imagination gives an intuition to inner sense

(which by itself has no objects of intuition) by connecting it and analogizing it to the manifold given in space, thus making time-determination possible. Kant says, "inner sense, on the contrary, contains the mere form of intuition, but without combination of the manifold in it, and thus it does not yet contain any determinate intuition at all, which is possible only through the consciousness of the determination of the manifold through the transcendental action of the imagination" (B154). Not only does such a concept of imagination play a crucial role in the Analogies, on which the Refutation of Idealism largely rests, but it also means that neither the transcendental synthesis of the imagination nor the empirical is able to form images out of thin air: the former requires something to connect, and the latter something to reproduce. They both require some matter to work with, a matter which must be given in spatial intuition. And, given Kant's theory of the imagination, the argument that we imagine an external world without any matter being given to the imagination becomes self-refuting: by taking away the matter with which the imagination works, the imagination would cease to be able to function.

Keeping in mind this necessity of matter for the imagination, we can explain those cases in which it seems that the imagination creates an illusory outer experience for us. Lifelike dreams and imaginings, on which much of idealism's appeal rests, are accounted for as reproductions of previous intuitions—our memories provide the matter for the imagination. Moreover, these representations are made possible by the fact that we are already determined in time. They depend on my time-determination as premise 8 prescribes, but I have already been able to determine myself in time on the basis of spatially perceived persistent images caused by an external world independent of me. Only after already having these intuitions stored in the memory, and after being previously determined in time,

can my empirical synthesis of the imagination create illusions.

Berkeley's Objection Returns

Thus, it seems that this reconstruction of Kant's argument fares better than the first. By pointing out the impossibility of determining oneself in time on the basis of representations alone, it is successful at least against the solipsist—a significant accomplishment. Yet the objection of Berkeley's God still looms large. Once again, Berkeley could accept that we perceive objects outside us in space and that they are caused by an external source while still maintaining his claim that it is God who creates the perceptions of the objects, that it is solely God who is this external source. Since these representations are not dependent on our determination in time for their persistence and change, they fulfill premise 8 and render idealism perfectly compatible with Kant's Refutation. It would seem that the only way to refute Berkeleyan idealism while maintaining trans-cendental idealism would be for Kant to prove that there must be some correspondence of our representations to the things represented—a proof that does not seem to appear in the text and likely would violate the assertion that we can know nothing about things in themselves.

Another option for Kant, and the one that I argue he takes, is for him to shift the burden of proof and say that there is simply no reason to believe that our perceptions do not correspond in some way to reality. In his introduction of the Refutation of Idealism, found on B274–B275, Kant notes that the motivation behind Berkeleyan idealism was that transcendental realism, the supposition that space is a property of things in themselves, is logically inconsistent. Since Kant has removed that motivation with his new transcendental idealism, there is no philosophically rigorous reason to believe Berkeley. Kant says:

Dogmatic idealism is unavoidable if one regards space as a property that is to pertain to the things in themselves; for then it, along with everything for which it serves as a condition, is a non-entity. The ground for this idealism, however, has been undercut by us in the Transcendental Aesthetic. (B274–275)

Instead, Kant takes as his main opponent the "problematic idealist" who "professes only our incapacity for proving an existence outside us from our own by means of immediate experience," and he intends with his argument to give such a person a reason for believing in external things (B275). In this light, Kant's Refutation of Idealism never took Berkeley as its main opponent to begin with. Though this approach deprives Kant of a full refutation of Berkeleyan idealism, i.e., a proof that Berkeley cannot be right, it leaves Kant's argument, in its limited scope, intact.

Conclusion

In this paper I first laid out what seems to be the most literal reading of Kant's Refutation of Idealism, consequently showing that it failed to prove its point in any definitive manner against both the solipsist, who claims that there exists nothing external to me, and the Berkeleyan idealist, who suggests that our perceptions do have their source outside me, but that they do not correspond with reality. I then put forward a modified reconstruction of the argument, founded on the premise that persistence and change in my representations, if those representations are entirely contained within me, is dependent on my prior determination in time. This argument was able to refute the solipsist, while remaining ineffective against the Berkeleyan idealist. Finally, I suggested that Kant, in the preamble to the Refutation of Idealism, in fact limits his scope to a refutation of solipsism, asserting that there is simply no good reason to agree with Berkeley. The compatibility of Berkeleyan idealism and transcendental idealism remains troubling, however. Since we cannot have knowledge of things in themselves, it seems entirely possible that behind the veil of our perceptions is a puppeteer God foisting representations—albeit persistent ones—on our sensibility. In fact, it seems Kant and Berkeley would agree on this statement from the Transcendental Aesthetic, that "[o]ur expositions accordingly teach the reality (i.e., objective validity) of space in regard to everything that can come before us externally as an object, but at the same time the ideality of space in regard to things when they are considered in themselves through reason" (A27–A28/B44). In this light, Berkeley's true fault would be that he makes unjustified claims about things in themselves, apart from the world of possible experience. So, if Kant's larger aim in the Refutation of Idealism was to separate himself from Berkeleyan idealism, that hope has been disappointed. However, against the solipsist or lonely Cartesian meditator who issues a skeptical denial of the possibility of knowledge that there is an external world, Kant's argument, at least in its second reconstruction, retains full force.

Works Cited

Kant, Immanuel. *The Critique of Pure Reason*. Edited by Paul Guyer and Allen W. Wood, Cambridge University Press, 2009.

The *Phainomena* in Aristotle's *De Anima*

AMMAR PLUMBER

UNIVERSITY OF PENNSYLVANIA

Introduction

Aristotle's *De Anima*, a treatise on the soul, exhibits many of the same methods seen in his other texts. Relatively little literature has examined the extent to which Aristotle privileges observable data in formulating theories. In the philosophy of science community, the term "saving the phenomena" is often used to denote this empirical focus. Many recognize the term from the writings of Bas van Fraassen and Pierre Duhem, who was inspired by the Copernican revolution era debates and the writings of Thomas Aquinas. Substantial evidence suggests that the concept pre-dates Aristotle, particularly in the field of astronomy.

The aims of this paper are twofold. First, I hope to characterize Aristotle's approach to scientific explanation in ontological, epistemological, and logical terms. Second, I examine how Aristotle saves the phenomena in his study of the soul.

I begin with a brief discussion of the history of "saving the phenomena" and Aristotle's ties to this astronomical tradition. Then, I provide an overview of Aristotelian methodology and further develop ideas put forth by Aryeh Kosman. After a brief comment about the roles of inference and direct observation, I proceed to a thorough examination of the methodological positions Aristotle takes in the three books of *De Anima*. Of particular importance are Aristotle's approach to defining the soul, his treatment of his predecessors' views, and the potential-actual distinction as a lens for studying the soul's faculties.

The Astronomical Origins of "Saving the Phenomena"

According to John Cleary, who cites Simplicius, the notion of "saving the phenomena" as it has come to be understood originates in the Platonic tradition of astronomy. These astronomers — Callippus and Eudoxus, for example — aimed to mathematically reproduce the observed movements of celestial bodies. A crucial assumption, inspired by Plato, was that celestial motion must be perfectly circular and uniform to reflect the divinity of heavenly bodies. Given this prior logical commitment or first principle, the crucial task was to apply this assumption in a manner that would be consistent with appearances. To this end, the observation of retrograde motion posed the greatest challenge for Platonic astronomers, as uniform circular motion seemed to do a poor job of accommodating retrogradation. Despite its quantitative indeterminacies, Eudoxus's theory of homocentric circles seemed, for a time, the most plausible solution to this problem, and both Callippus and Aristotle undertook to formalize and refine it. Also noteworthy is that Eudoxus wrote a book that detailed his observations of the heavens, which he titled *Phainomena*. The book described with great specificity the rise and fall of constellations, among other celestial bodies. This context broadly conveys what the method of "saving the phenomena" has been subsequently taken to mean: theory, insofar as it is apt, must fit with what is observed—that which is to be explained.

Aristotle and the Astronomical Tradition

Aristotle's awareness of the methods of astronomers is evidenced in *Metaphysics* XII, 8. When examining the question of how many unmoved movers must exist, Aristotle notes:

> …when we come to the number of these spatial movements, we must investigate it on the basis of the mathematical science that is most akin to philosophy, namely, astronomy. For it is about substance that is perceptible but eternal that this produces theoretical knowledge, whereas the others are not concerned with any substance at all—for example, the one concerned with numbers and geometry.
> (*Metaphysics* XII, 1073b3-8, trans. Reeve)

Here, Aristotle distinguishes astronomy, a science concerned with perceptible substances, from those that are only concerned with abstract intelligible objects—arithmetic and geometry, for instance. Astronomers, unlike arithmeticians, cannot ignore sensory data. Likewise, when attempting to answer the question of how many unmoved movers there must be, Aristotle cannot ignore the relevant perceptible substances: those things that are moved. Further demonstrating Aristotle's familiarity with the methods of astronomical study, Aristotle later cites Eudoxus's theory of homocentric circles and the subsequent work by Callippus (*Metaphysics* XII 1073b17-35).

In *Prior Analytics*, it is further possible to discern Aristotle's understanding of how first principles are obtained given astronomical observations. In Book I, he says, "…it is the business of experience to give principles which belong to each subject. I mean for example that astronomical experience supplies principles of astronomical science; for once phenomena are adequately apprehended, demonstrations were discovered. Similarly with any other art or science" (*Prior Analytics* I, 46a18-26, trans. Jenkinson). Aristotle takes it as a methodological rule that first one must collect appearances and then obtain the first principles. How explanatory principles and observables relate to one another demands further elaboration, which I undertake to do forthwith.

An Overview of Aristotelian Methodology

In examining Aristotle's methodological and epistemological views, it is perhaps useful to evaluate how his rhetoric compares to a variety of contemporary philosophical positions. Aryeh Kosman aims to do so in Chapter 8 of his book *Virtues of Thought: Essays on Plato and Aristotle* by comparing Aristotle's position to scientific realism and instrumentalism. Realism reflects an attitude that scientific theory aims to approximate truth. Instrumentalists, on the other hand, contend that the success of a theory consists only its predictive accuracy and that its exact semantic content—truth-like or wholly unrealistic—is unimportant. It is interesting that Kosman chooses only to discuss these two positions on scientific explanation, as there exist many others that potentially resemble Aristotle's views. For instance, empiricism, logical positivism, structural realism, and constructive empiricism are all differentiable from the two positions Kosman considers. However, for the sake of concision, I omit these comparisons, as they are not necessary to make good sense of Aristotle's position.

Interpreting remarks offered in Book I Chapter 2 of *Posterior Analytics*, Kosman discusses how Aristotle conceives of the relationship between understanding and explanation. For Aristotle, something is understood when its cause—what is responsible for the phenomenon's being the way it is—is known. Given an applicable cause, the explanandum would come about as a matter of necessity. I will later describe in greater depth how Kosman understands Aristotle's conception of causality.

Now, I discuss how explanatory demonstration relates to understanding. Between the two, there exists a bidirectional connection.

Explanatory demonstration is a crucial step in achieving understanding. However, understanding also has a dispositional sense; one who understands something is better disposed to explain it. Explanation proceeds from what is "true… and better known than and prior to the cause of the conclusion" (*Posterior Analytics* 1.2, 71b21-23). Later, Aristotle says, "…we understand something only when we know its cause, prior, insofar as they are causes, and known before it, not only in the other sense of being aware [of what they are] but knowing as well that they are the case" (71b31-34). This final requirement is important. For an explanation to be successful in producing genuine understanding, its premises must be true. A fictionalism, in other words, cannot produce understanding in the Aristotelian sense, as understanding requires that the true causes of a phenomenon are known. As such, an instrumentalist reading of Aristotle appears implausible. Moreover, what Aristotle means by referring to a cause as "prior" is ambiguous. At least two notions of causal priority may be operative. The first is logical causation—the causal relation between a set of premises and its conclusion. The truth of the premises logically entail the truth of the conclusion. The second notion of causation is an ontological one; some events physically necessitate the emergence of other events. Kosman points out, "…Aristotle moves comfortably from one sense to the other, as though here logic and ontology were easy bedfellows" (Kosman, p 141). A third sense of priority is suggested by Aristotle's requirement that a cause be "better known than" the conclusion. This phrasing implies an epistemological sense of priority. I proceed now to a discussion of each of these three notions of priority.

Elaborating on the aforementioned criteria of understanding, Aristotle says, "…'prior' and 'better known' may be understood in two senses, for what is prior by nature is not the same as what is prior to us, nor what is better known

without qualification the same as what is better known to us" (*Posterior Analytics* 1.2, 71b33-72a1). He continues, "…things prior and better known without qualification are furthest from sense. Now things most universal are furthest from sense, and particulars nearest to sense, and they are thus exactly opposed to one another" (72a1-6). Through the lens of logic, particulars and universals stand at opposite ends of a two-way street. One uses induction to proceed from observation of particulars to statements about universals. Deduction, on the other hand, proceeds from universal premises to conclusions about particulars. The logical entailment of particulars from universals is true independently of our order of learning, which explains why Aristotle refers to these pro-positions as "prior without qualification", distinguishable from something that is "prior to us". Sense, too, is central to Aristotelian epistemology. Because universals are distant from sense, they can only be evaluated with reference to particulars. For instance, one cannot directly observe that all human beings are mortal. However, one can observe a particular individual's death, which accords with this universal claim about human mortality.

Given this notion of logical priority, one may wonder whether Aristotle thinks understanding is achieved whenever the explanandum are deductively entailed by the explanans—something akin to Hempel's deductive-nomological model of scientific explanation. In other words, is the successful identification of what is logically prior a sufficient condition for scientific explanation? Recall that Aristotle says that in order to understand something, we must not only know what its causes are, but we must know "as well that they are the case" (71b34). Thus, it is not sufficient to conjure up universal premises that logically necessitate the explanandum. Epistemological priority is also crucial to explanatory demonstration.

A final criticism of this interpretation concerns its neglect of Aristotle's apparent

emphasis on physical causation. One can deduce the height of a flagpole from the angle of the sun and the length of the flagpole's shadow. But, to say that the sun and the shadow are the *cause* of the flagpole's height does not seem productive of understanding because neither the sun nor the shadow explain why the flagpole possesses the height that it does as a matter of physical necessity. After all, the flagpole is prior in being to its shadow, and the length of the shadow is understandable to us a consequence of the flagpole's height.

As I will demonstrate further, Aristotle in *De Anima* proceeds in a similar way. The nature of visual sense data (color or light) determines how the eye must be in order to perceive it. Therefore, successful causal explanation of sight must "prioritize" light in a manner that is not reducible to logical (deductive) entailment. Ontological priority is relevant as well. But, the task of identifying what is ontologically prior poses an epistemological puzzle: how does one differentiate what is ontologically prior from other features from which one can deduce the explanandum? For instance, how does one learn that the height of a flagpole is prior in existence to its shadow? This is a question that I will set aside for now and revisit later when discussing the importance of collecting a large quantity of appearances.

Expounding upon this notion of ontological priority, Kosman says, "The prior is that which exists without the posterior, but without which the posterior does not exist" (p 143). Generative cause and effect is one example of something that reflects this priority in being. For example, beer requires the preexistence of yeast and the process of fermentation, but neither yeast nor fermentation depend upon the existence of beer. Similarly, light is ontologically prior to sight. Without light, the faculty of sight could not exist, but the existence of light is not preconditioned on a creature's ability to perceive it.

An alternative division of priority is evident in Aristotle's *Metaphysics*. He remarks, "…things that are prior in the order of rational discourse are different from those prior in relation to perception. For in the order of rational discourse universals are prior, whereas in relation to perception individuals are" (*Metaphysics* 5.11, 1018b30-33). Rational discourse is the medium of explanatory demonstration. Moreover, Aristotle thought of science as the power of rational discourse to render nature intelligible. As such, scientific explanation must involve the prioritization of universals over particulars. A further wrinkle is that explanation involves more than identification of a proximate cause and an effect. Rather, explanandum are intelligible in the context of a broader explanatory framework. Something that is prior in rational discourse is "so constituted as to provide the ground of intelligibility for other elements in the system" (Kosman, p 146). Priority in this sense implies that "one thing is intelligible in terms of another, but not vice versa, or not to the same degree" (p 147). This conception of priority is both logical *and* ontological. One thing is intelligible in terms of another if there exists an explanation that successfully proceeds from universals to what is less known by nature (logic). The directionality of this explanation in part depends on what exists prior and what is posterior (ontology). Together, these three facets of Aristotelian priority constrain the directionality of explanation and what are identifiable as causes as opposed to effects. In sum, understanding is not an isolated piece of explanatory demonstration (i.e. identification of a proximate cause). It is dependent on the phenomenon being intelligible in light of the entire body of rational discourse, traceable to primary conditions and first principles (*Physics* 1.1, 184a13). First principles, according to Aristotle, are apprehensible through *nous*, the faculty of intelligence that explains our understanding of imperceptible propositions.

Furthermore, *nous* confers the ability to produce logically connected discourse.

How does epistemological priority factor into Aristotelian explanation? Given that universals are distant from sense, how can one come to know the first principles with which one can explain observed particulars? The answer to this question, in part, depends on Aristotle's account of *nous*, a topic to be discussed later. In the meantime, it is worth examining what else Aristotle says about the apprehension of first principles. In Book I Chapter 1 of *Posterior Analytics*, Aristotle holds that the discovery of first principles proceeds simultaneously alongside the task of explaining particulars—"revealing the general by making clear the particular" (71a8). Kosman characterizes this process as "grasping the phenomena clearly by explaining them, and, in the process, grasping the principles of explanation" (p 149). One who seeks first principles is accountable, first and foremost, to the *phainomena*. How appearances facilitate the discovery of principles is less clear, but one helpful suggestion, defended by Jean De Groot, is that the quantity or breadth of observations are important for this task (De Groot, p 86). Perceptual data is to be used in an inductive manner in producing generalizations to be explained. For inductive science to be dependable, one must have multiple appearances in hand (*Prior Analytics* 1.30). Furthermore, opinions that are widely held reflect a greater diversity of experience, and those opinions held by experts in a certain field are also more trustworthy. Thus, starting with the *phainomena*, for Aristotle, involves two priorities that are important for successful induction: quantity of observation and commonality or authoritativeness of opinion.

Terence Irwin in *Aristotle's First Principles* further discusses how Aristotle proceeds from what is better known to us (appearances) to what is better known without qualification (first principles). In Chapter 2 Section 12, Irwin interprets Aristotle's approach to science as something similar to the puzzle-solving enterprise that Thomas Kuhn puts forth in *The Structure of Scientific Revolutions*. One begins by surmising general laws from observation. Then, using these laws (first principles), one attempts to solve the puzzles generated by the *phainomena*: those appearances and common opinions that have yet to be explained. If the difficulties are resolved and the *phainomena* are successfully demonstrated, the first principles are confirmed.

Returning to Kosman's original question, it is clear that Aristotle's views cannot be characterized as purely realist and certainly not as instrumentalist. Instrumentalists would disagree with Aristotle that successful explanations must have true premises. However, Aristotle diverges from realism insofar as he believes that first principles are inaccessible to sense and therefore are not necessarily veridical. Only pure appearances can be infallibly known, but truthful explanations can be reliably approached through diligent observation and a persistent effort to synthesize appearances and first principles.

The Objectivity of Sensible *Phainomena*

Another possibility that must be considered is that observation or predication itself may be fallible. Pure sensory data can be distinguished from propositional knowledge, which requires cognitive judgment. For example, the color red is directly perceived and, therefore, veridical, but the judgment of a red object as a cardinal is not. A fundamental question, then, is which qualities are sensible and reliably true and which are intuited via something other than pure sense.

In Book II, Aristotle admits of the possibility of being deceived by one's perception. In Chapter 6, he delineates three categories of sense-objects: proper objects, common objects, and incidental objects. This first category, proper objects, can only be perceived by one sense and not by others. For instance, hearing is connected with sound, sight with color, and taste with flavor. Aristotle notes, furthermore, that it is

impossible to be deceived in one's direct perception of a proper object. However, one can be mistaken about the identity or location of substances that transmit sense data (418a11-13). Common objects of perception are those qualities that not particular to one sense: movement, number, size, and shape. The size of a substance, for example, can be sensed via either touch or sight. A final category is of those objects that are only incidental to their substances—not inherent to their physical composition. The example given is of a white thing that happens to be the son of Diares (the incidental object). That the white thing is the son of Diares is merely incidental to what is sensed: its color (418a20-24). In other words, having a white color does not logically entail that it must be the son of Diares. Like with common objects, one can be incorrect in judging incidental characteristics of a thing, but the proper objects of perception (i.e. whiteness) are undeceiving.

A methodological question follows: how does this categorical scheme influence what *phainomena* are considered veridical as opposed to inference-dependent? After all, only proper objects of perception are reliable and directly knowable. Aristotle's investigative approach to the study of homoiomerous bodies (bodies that are uniform in structure) in *Meteorology* sheds light on this question. As previously mentioned, Aristotelian epistemology tends to proceed from what is better known to us to what is ontologically prior. Even his efforts to interpret perceptible objects proceeds in his way. In Book IV Chapter 8 of *Meteorology*, Aristotle says:

> All these bodies differ from each other, firstly, in the particular ways in which they can act on the senses (for a thing is white, fragrant,… hot or cold in virtue of the way it acts on sensation), and, secondly, in other more intrinsic qualities commonly classified as passive—I mean solubility, solidification, flexibility, and the like…. It is by these passive qualities that bone, flesh,… stone and all the other natural homoiomerous bodies are differentiated. (385a1-11)

This passage suggests that truths about investigative objects are not confined to the realm of sensible qualities but also include passive qualities that do not act on the senses, such as solubility or flexibility. However, only that which is directly sensed can be reliably known, and passive qualities—while no less real—must be inferred. The definition of a natural body, however, is dependent on both sensible and passive qualities, both reflective of the body's *dunameis* or characteristic function, which cannot be reduced to the *dunameis* of its material constituents. In order to interpret what something is, therefore, one must infer its additional qualities from its sensible features.

The need to interpret sensory data helps to explain, therefore, the importance of common opinion to Aristotle's *phainomena*. Only direct sensation of proper objects is veridical. This category of objects is bleak as an investigative starting point; colors, tactile qualities, smell, and taste alone yield only very crude investigative queries. For this reason, interpretation must take place in if meaningful insights are to be gleaned. Accordingly, the *phainomena* to be explained must include not only the proper objects of perception but also the functional and physical interpretations that all can agree on—those that are beyond doubt. When examining Book III of *De Anima*, I will offer an additional explanation of this interpretative process by examining its parallel in the soul.

The Declared Methodology of *De Anima* in Book I

Aristotle in Book I offers a general overview how he will approach his investigation of the soul. At numerous points, his rhetoric is consistent with Kosman's account of Aristotelian explanatory priority.

In Chapter 1, Aristotle considers the appropriate starting point of an investigation of the soul. He notes that different fields rely on different principles (402a21-22). The soul's

genus—the kind of thing it is—determines what domain of rational discourse is appropriate to explain it; if the soul is one kind of thing, it falls within the proper domain of the natural philosopher, and if it is another, a metaphysician is better suited to study it. Each field entails a different web of ontological, logical, and epistemological starting points, and different universal premises would be required to explain the soul's activities.

To choose an appropriate starting point and realm of inquiry, Aristotle must first determine what the soul is — its definition. Beginning with appearances (*phainomena*) first, Aristotle endeavors to discern which attributes are essential from those that are merely accidental to the soul. To illustrate, consider how one might define a human. While humans may have hair of a certain color, the possession of hair of a particular color is not essential to one's humanity but only accidental; simple observation reveals that humans can be bald or have red rather than brown hair (Hahmann, p 8). In the same way, by examining appearances of ensouled and lifeless bodies, Aristotle hopes to clarify the soul's substance (402b25-26). This task of defining the soul does not quite characterize Aristotle's approach to scientific explanation, as it is not an explanatory step but a descriptive one. Successful identification of the soul's substance is not a manifestation of demonstrative knowledge, as it does not serve to explain the faculties of the soul in causal terms. At this stage, Aristotle aims only to identify, given certain observable facts, some attributes that are essential to the soul. However, it is later evident that the definition of the soul ultimately guides the search for explanatory first principles and provides grounds for inferring other essential qualities. In this sense, one might say that the *phainomena* determine which explanations are adequate and which are not.

How Aristotle evaluates the views of his predecessors (*endoxa*) in Book I reveals his methodological commitments. As such, I offer Aristotle's treatment of Democritus's views as an illustration of how Aristotle views the relationship between appearances and explanation. Aristotle begins by summarizing Democritus's position. Unlike bodies that are not ensouled, ensouled bodies are capable of motion. An adequate account of the soul should explain how it initiates motion in the body. Democritus reasons that the soul must be in motion, as that which is not in motion cannot move something else. The soul, he supposes, must be a sort of fire or heat—composed of infinitely small, round particles. The ever-moving soul draws the body along with it and, in doing so, sets the body in motion. Aristotle dismisses Democritus's view on account of the following observation: ensouled bodies both move and rest. If a soul is in motion by its internal principle, then how could ensouled beings rest? The explanation given by Democritus, therefore, is empirically inadequate (406b20-26). It does not save the phenomena. Throughout the remainder of the text, Aristotle proceeds in similar fashion. He consistently dismisses views that contradict appearances and uses the appearances to justify alternate theories.

The Soul in Terms of Potentiality and Actuality

In discussing Book II, I focus on those parts that are relevant to understanding Aristotle's approach to analyzing the faculties of the soul and offer two examples of this approach manifested: nutrition and sight.

Leaving behind the views of his predecessors, Aristotle begins Chapter 1 of Book II by offering a conceptual scheme through which the soul can be defined. He briefly considers the category of substances, which contains three subcategories: matter, form—both simple substances—and composites of matter and form. Every object in the world is a composite of matter and form. Aristotle gives the example of a wax candle. The object is composed of material that has the function of a candle. But, the matter can be conceptually distinguished from form. Imagine if the characteristic features

of a candle—presumably, its function as something to carry a flame and provide light— were separated from the matter. While conceptually distinct, form cannot exist in the physical world but in matter, and matter cannot be formless, but the distinction is still meaningful. Because Aristotle concludes that the soul *is* a substance—the form of a living body—the notions of mater and form and, relatedly, potentiality and actuality will be helpful in understanding his approach. Evidently, Aristotle's search for a definition is entangled with his quest for explanatory first principles.

The soul's definition, according to Aristotle, is that which explains the difference between living bodies, which are ensouled, and non-living ones, which are soulless. The soul, therefore, is the actuality of life in a body that is potentially alive. It is necessary to pay particular attention to this principle that guides much of the analysis in *De Anima*: to know what is potential, one must first examine what is actual. When discussing Chapter 4, I will return to this topic and explain it in terms of explanatory priority.

Bodies that are potentially alive are those that possess organs. This is an inductive claim supported by the following observations, among many others: "…even the parts of plants are organs, although altogether simple ones. For example, the leaf is a shelter of the outer covering, and the outer covering of the fruit; and the roots are analogous to the mouth, since both draw in nourishment" (412b1-5). Here, Aristotle demonstrates a commitment to the *phainomena*. An organ-focused explanation of the soul must account for plants, which are also living. Because what distinguishes an ensouled from a non-ensouled body is life, in characterizing the soul, one must first answer the question of what it means to be living—the focus of Chapter 2.

In Chapter 2, Aristotle identifies a number of activities characteristic of living bodies. Again, in doing so, he begins with the *phainomena*. One can observe that plants, if

nourished, grow (413a26-31). In addition to this faculty, animals also have perception and movement. Perception can be subdivided into touch, sight, hearing, and taste. Finally, humans alone have the capacity for intellection or thought. The question then arises of whether the soul is divisible according to these various functions. Because the soul is defined as the first actuality of these faculties characteristic of living bodies, it *is* conceptually partitioned. The vegetative soul consists only in a nutritive part. Animals additionally have perception: sight, touch, etc. Humans must have an intellectual soul in addition to their perceptual, nutritive, and appetitive souls. Clearly, Aristotle has no *a priori* commitment to a monopartite soul. Rather, his conceptual divisions are rooted in *observable* divisions; a creature can have the capacity for sight without the capacity for hearing. A creature's faculties are divisible, so the soul, too, must be divisible (413a11-28). Before analyzing the senses, Aristotle notes that many of these faculties are only actualities of a potentially living body. Perception cannot exist without a body that can potentially perceive. Therefore, most of the soul's parts can only inhere in the body and must be investigated with this relationship in mind.

Aristotle in Chapter 3 admits of a hierarchy of souls, rooted in the observable hierarchies of living beings. Because plants have the least living functions, they are at the bottom of the hierarchy, and principal soul of plant life must also be at the bottom—namely, the nutritive soul. Next, the animal soul additionally has locomotive, perceptive, and appetitive parts. The reason animals must have desire is that touch and other forms of perception involve pleasure and pain, which, by necessity, are tied to appetite or desire. Animals are drawn to what is pleasurable and repelled by what is painful. Again, the *phainomena* (observed behavior of animals) require this appetitive faculty to be existent. Finally, the contemplative soul *nous* is only present in the beings at the top of this

hierarchy. Because thought does not observably or necessarily inhere in any anatomical part, a non-organic account may be necessary, as we see in Book III.

Chapter 4 offers the most penetrating insights into how Aristotle's potential-actual distinction is bound up with three different kinds of priority: epistemological, ontological, and logical priority. In explaining his approach, I intend to provide a lens through which the rest of Book II can be understood.

The actualization of a faculty (i.e. seeing something, digesting something, or cognizing something) are epistemological inroads to the nature of the faculty. For instance, only in studying *what* is seen and *how* these objects can be seen can one properly characterize the faculty of vision. In other words, first, we observe the actuality of the sense. Then, we know what potentially senses. Moreover, the objects of a faculty are ontologically prior to the faculty in that the objects determine the faculty. If colors were by nature different than they are, vision, too, would have to be different. Some notion of logical priority is also operative in Aristotle's framework. The nature of faculties or potentialities are inferable from the nature of their objects. This multifaceted notion of conceptual priority is substantiated in Chapter 4 before Aristotle discusses the nutritive soul: "…if one ought to say what each of these [faculties] is…, then one should first say what reasoning is and what perceiving is, since actualities and actions are prior in account to potentialities. But… it would for the same reason be necessary to make some determinations about… nourishment and the objects of perception and reasoning" (415a15-25). Here, "prior in account" presumably refers to what is prior in rational discourse—that which is logically prior. One can gather from the objects how the faculty must be in potentiality.

Aristotle later discusses the soul in causal terms. He says, "The soul is the cause and principle of the living body… in the three of the

ways delineated; for the soul is a cause as the source of motion, as that for the sake of which, and as the substance of ensouled bodies" (415b5-15). The first way refers to the soul as the "efficient cause" of motion—that which generates motion. Efficient causality reflects ontological priority, as the soul is prior in being to motion; the soul must exist first for motion to be generated. The second way the soul is the principle of the living body is as the body's "final cause". The soul represents the telos or purposive end of the body in that the body exists in order to carry out the soul's functional purposes. In observing the body's behavior, therefore, one can understand what purposes it is designed for and thus gain insights about the soul. Finally, the soul as the substance of ensouled bodies refers to "formal causality"; the living body's form depends on the kind of soul it has. An ensouled body with a vegetative soul would have the form of a plant. This, too, implies a kind of epistemological priority that is related to telos. Form, for Aristotle, has a functional character. The formal substance of an axe consists in its aptitude as a tool for cutting. If this property were separated from the object, "it would no longer be an axe, aside from homonymously" (412b10-15). Thus, the functions of the body characterize its soul—the form of the body. As noted, a body has the form it does because of its telos. These notions of causality and priority are seen throughout Aristotle's investigations of the various souls.

I now offer Aristotle's accounts of nutrition and sight as illustrations of his signature approach to studying the soul's functions. Beginning with nutrition in Chapter 4, Aristotle divides the process of nourishment into three: "what is nourished, that by which it is nourished, and what nourishes—that which nourishes is the primary soul; that which is nourished is the body which has the primary soul; and the nourishment is that by which it is nourished" (416b20-26). Noting that the purposive end of nourishment is growth, he begins with the *endoxa*—how his predecessors accounted for the growth of living

beings through nutrition. Empedocles does not treat the soul as the efficient cause of growth. Rather, he attributes growth to the elements that inhere in living beings. For plants, the presence of earth causes downward growth of the roots, and fire promotes upward growth of the rest of the organism. Aristotle rejects this account on the grounds that fire's limitless upward motion and earth's limitless downward motion would cause plants to tear apart (416a5-9). The *phainomena* contradict this implication. Thus, something else must explain plant growth. The other position that Aristotle rejects is that fire alone is responsible for growth. Unlike the body, however, fire can grow in an unlimited and unsystematic manner. Aristotle concludes that the soul is needed to explain the organized growth process of the body. From his treatment of the *endoxa*, it is clear that the *phainomena* come first. Subsequently, Aristotle turns his attention to food—that which nourishes—and discusses two positions. The first is that like nourishes like, and the other is that unlike nourishes unlike (416a29-32). Can both of these views be accommodated? Aristotle responds affirmatively. While food is unlike the body pre-digestion, it becomes like the body post-digestion. Both views can be maintained. Heat is the process by which food is digested, and the soul is both the formal and efficient cause of the structured nutriment and growth that ensues. The presence of heat also accounts for the fact that all ensouled bodies are warm (416b28-30). Throughout this inquiry, Aristotle focuses first on the actuality of nourishment: food intake and growth. By observing this process, one can infer what qualities must exist in the soul—what nourishes. Two aspects of the actuality of nourishment stand out to Aristotle. First, he notes the organized growth process of the body and concludes that the soul must account for this structure. Second, he observes that undigested food is unlike the body, and, heat, therefore, must make it alike. In characteristic fashion,

Aristotle proceeds from the object and activity to the faculty itself—potential to actual.

Aristotle's approach to sight is similar. He begins with the object of sight: color. The transparent is the medium of color, and it inheres in air and water. Light activates the transparent as a medium, and, when illuminated, color can be perceived through it. The transparent then affects the eye and allows color to be perceived. Aristotle explains why the transparent is necessary as a medium for color to be perceived: "if someone should place what has colour upon the eye itself, it will not be seen. Rather, colour moves the transparent… and the sensory organ is moved by this" (419a12-14). What the eye receives is the form of the perceived color, not the actual composite being observed (424a17-20). But, the eye alone is not sufficient to explain sight because an eye in a non-living creature does not perceive color. The soul, therefore, is the first actuality of eyesight for an eye that can pot-entially see, and the second actuality is realized when color affects the eye. Again, Aristotle begins with the actuality of vision and the object of eyesight—color—and proceeds to an account of the faculty of vision. Aristotle's approach to each sense throughout Book II reflects the aforementioned notions of priority, which are built into his potentiality-actuality framework.

Aristotelian Epistemology and *Nous*

Because much of Book III is not concerned with external objects and appearances, it is best to focus on that part of Book III that speaks to a previous question: how does one acquire first principles? Chapter 8 discusses the relationship between imagination (*phantasia*), reason (*nous*), and perception. What is said on this matter directly parallels Aristotle's stance regarding the determination of first principles. Aristotle suggests that the only forms that can be contemplated are those that inhere in the objects of perception (432a5-10). A person who does not perceive anything could not learn or contemplate

anything. Accordingly, the objects of contemplation involve images, apprehended by *nous* as forms of perceptible objects without the matter. However, Aristotle then ponders what distinguishes images from first thoughts (432a10-15). Forms cannot be contemplated unless one has prior assumptions that allow the apprehension of form from images. Without these prior assumptions or first thoughts, there is nothing other than images and *phantasia*, and contemplation cannot take place. First thoughts, therefore, are not images even though they are actualized only with images; these objects are proper to *nous*. Aristotle does not elaborate about how these first thoughts originate, but it is clear that they cannot be learned through observation. They are prior to the interpretation of sense data.

The relationship between *nous* and *phantasia* is instructive with regards to a question discussed earlier: how can *phainomena* as pure sense data translate into propositional knowledge to be explained in terms of first principles? Recall the earlier example of a red cardinal. In order to issue any explanations of the behavior or biological traits of a cardinal, one must first interpret the colors received by the senses as those of a cardinal. This step requires a first thought—an interpretative assumption about what a cardinal is and what a cardinal is not. Whether this preconceived notion of a cardinal is a kind of "first principle" is unclear because Aristotle does not explicitly equate first thoughts and first principles in *De Anima*. However, it is clear that the two are similar.

This analogy between first principles in Aristotelian philosophy of science and first thoughts in the soul is also helpful in clarifying the nature of first principles. Like first thoughts, first principles, too, must be postulated before explanation can take place. In other words, first principles cannot be discovered in the *phainomena*. As such, first principles are prior without qualification although they are not prior to us. However, some first principles may be rejected because they fail to explain any

phainomena of interest. If one assumes that all celestial bodies must move triangularly but never observes a celestial body that does so, the first principle fails to explain any of the relevant *phainomena*. It is an open question whether Aristotle believes that first thoughts can have no viable objects in the same sense. In any case, it is clear that first principles cannot be proven or deduced from observation.

Conclusion

"Saving the phenomena" is an apt characterization of Aristotle's approach to scientific explanation for several reasons. A valid explanation for Aristotle begins with what is "prior and better known without qualification". I purport to show that this sense of priority is both ontological and logical. Furthermore, he consistently demonstrates a concern for theory's consistency with appearances. In fact, throughout *De Anima*, Aristotle starts with what can be known through observation. In studying the senses, he begins with the objects of perception, better known to us, and from these infers how these faculties must exist in potentiality in order to be actualized. Aristotle holds the *endoxa* to the same standards; he dismisses those views that are contradicted by the phenomena and retains those that are empirically adequate. Finally, a suitable analogy can be made between first thoughts in the *nous* and first principles in scientific explanation. In drawing this analogy, it becomes clear why first principles cannot be abstracted from the phenomena. Rather, they must be formulated first before explanation can be attempted.

Bibliography

Aristotle. (2016). *De Anima*. trans. Christopher Shields. Oxford UP, Oxford.

Aristotle. (2016). *Metaphysics*. trans. C. D. C. Reeve. Hackett, London. Aristotle. (1931). *Meteorology*. trans. E. W. Webster.

Aristotle. (1928). *Posterior Analytics*. trans. G. R. G. Mure

Aristotle. (1928). *Prior Analytics*. trans. A. J. Jenkinson.

Cleary, J. (2013). "*Phainomena* in Aristotle's Methodology". from *Studies on Plato, Aristotle, and Proclus: The Collected Essays on Ancient Philosophy of John Cleary*. BRILL, Leiden.

De Groot, J. (2014). *Aristotle's Empiricism: Experience and Mechanics in the Fourth Century BC*. Parmenides Publishing, Las Vegas.

Hahmann, A. (2016). "Draft – Session 1: § 1 Prooemium (I.1)".

Irwin, T. (1988). *Aristotle's First Principles*. Clarendon UP, Oxford.

Kosman, A. (2014). "Saving the Phenomena: Realism and Instrumentalism in Aristotle's Theory of Science". from *Virtues of Thought: Essays on Plato and Aristotle*. Harvard UP, Cambridge.

Nussbaum, M., and Rortie, A. (1992). *Essays on Aristotle's De Anima*. Clarendon UP, Oxford.

Derrida and Habermas on Modernity and the Production of Meaning

MAXIMILIAN PAUL JONATHAN FORSTER

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

Abstract: In this paper, I will contrast the way in which Derrida and Habermas conceptualize both meaning and modernity. In part one, regarding meaning, I identify a diametrical opposition between the concepts of Derridean *différance* and the Habermasian communicative action. In part two, I contrast their conceptions of modernity. It will turn out that Derrida takes a *totalizing* approach while Habermas offers a *differentiated* concept of modernity. I argue that their conceptions of meaning, based on the concepts of différance and synthesis, inform their approaches to modernity. Additionally, I will argue that the Habermasian theory of modernity is preferable due to its sociological explanatory power. I recognize the possibility that this criterion and the focus on modernity itself might do a certain violence to Derrida. Nonetheless, this paper will conclude with a critique of Derrida's theory of meaning, arguing that its political implications are implausible and undermine its credibility.

Jürgen Habermas and Jacques Derrida are arguably the most prominent recent intellectuals of Germany and France. They face off in what Andrew Bowie calls the "German-French debate" (Knellwolf and Norris 121). Representatives of the movements of critical theory and deconstruction, Habermas and Derrida discuss literary theory, social theory and philosophy. This paper will weigh in on this debate. It will focus on the Habermasian and Derridean conceptions of meaning and modernity. Part one discusses the disagreements about meaning and how it is produced, and part two will focus on the debate around concept of modernity. Finally, it will be argued that we ought to adopt a Habermasian theory of modernity and that the Derridean theory of meaning is rendered implausible by its political implications.

1. Meaning

To begin with an analysis of Derrida's understanding of how meaning is produced, it will be helpful to recall the semiotic theory of Ferdinand de Saussure (1857-1913), which was undoubtedly highly influential for Derrida. Saussure breaks down the sign into two parts: the signifier (e.g. the word or sound-image 'tree') and the signified (e.g. the concept of a tree). He characterizes the "bond between the signifier and the signified" as "arbitrary" (Saussure 67). Due to this arbitrariness, the signified (e.g. the concept of a tree) cannot explain why its signifier (e.g. the word 'tree') functions as such, i.e. has its particular meaning. Thus, Saussure formulates an alternative explanation of the 'production of meaning' grounded on the situation of words in relation to each other. The *value* of a word is provided by its opposition to all the other words and their values.

Two years before Derrida's pivotal *Structure, Sign and Play*, Barthes recognized the 'prophetic' character of this idea. He concisely sums up the essence of the Saussurean notion of meaning, which holds that meaning results from a division:

> For Saussure imagines that at the (entirely theoretical) origin of meaning, ideas and sounds form two floating, labile, continuous and parallel masses of substances; meaning intervenes when one cuts at the same time and at a single stroke into these two masses. The signs (thus produced) are therefore *articuli*; meaning is therefore an order with chaos on either side, but this order is essentially a *division*. (Barthes 18)

Derrida, too, emphasizes the importance of value over signification. Like Saussure, Derrida does not only present his own conception of meaning. First, he presents the 'history of meaning' that has been prevalent until an event occurs—or something, as Derrida reprimands us, only a naïve mind would call an event—and then proposes his own alternative. This requires some disambiguation.

On the one hand, we find "A history of meaning [sens]—that is, in a word, a history—whose origin may always be reawakened or whose end may always be anticipated in the form of presence" (Derrida, *Writing and Difference* 353). In this conception of meaning, for Derrida, the notion of signification is at work. According to this conception, meaning is falsely assumed to be provided by a transcendental signified. A false sense of the 'presence' of the signified in the signifier seems to provide the signifier with meaning.

Derrida criticizes this understanding as the metaphysics of presence and as a "reduction of the structurality of structure" (*Writing and Difference* 353). Presumably, the structure that is 'reduced' by this false assumption of the presence of the signified is the structure of differences that, in Saussurean terms, provide value.

Analogously to the *signified* in linguistics, in the history of metaphysics, a *center* has always been assumed as present. Historically, these centers are substituted and receive different names, such as "essence, existence, substance, subject (…) transcendentality, consciousness, God, man" (Derrida, *Writing and Difference* 353)

However, with the emergence of structural semiotics, a rupture occurred, and the presence of the metaphysical center was substituted not by a new center, but a function or a 'non-locus'—language. The resulting absence of a present center or the "absence of the transcendental signified" extended "the domain

and the play of signification infinitely" (Derrida, *Writing and Difference* 354). Here, the Saussurean influence becomes clear. What he calls the undivided—and therefore *meaningless*—'chaos' is provided with meaning by division. Likewise, for Derrida, after the theory of signification and the metaphysics of presence have been uncovered as naïve and as insufficient to explain the production of meaning, it is clear that "the original or transcendental signified, is never absolutely present *outside a system of differences*" (*Writing and Difference* 354, emphasis mine). This system of differences, amongst other things, is what Derrida calls 'différance'. As constitutive of meaning, différance, is "the possibility of conceptuality, of a conceptual process and system in general" (Derrida, Margins of Philosophy 11).

Similar to Derrida and Saussure, Habermas gives an account of an outdated, traditional understanding of language. He tells us that traditionally in European philosophy, language was regarded as an instrument, allowing perceiving subjects to communicate pre-linguistically perceived objects through the means of designation (Cf. Habermas, "Liberating Power" 12f.). According to Habermas, Cassirer was the first to oppose this conception and precipitate the linguistic turn.[10] Cassirer claimed that language contains a productive and generative power primary and constitutive to the objects:

The object is not that which is given, but that which must first be attained, not that which is determined in itself, but that which must be determined. Because, linguistically speaking, this fundamental

---

[10] "Cassirer was the first to perceive the paradigmatic significance of Humboldt's philosophy of language; and he thus prepared the way for my generation, the post-war generation, to take up the 'linguistic turn' in analytical philosophy and integrate it with the native tradition of hermeneutic philosophy" (Habermas, Liberating Power 12)

determination takes place within the sentence, Humboldt's language philosophy asserts the primacy of the sentence over the word, just as Kant's transcendental logic asserted the primacy of the judgement over the concept. (Cassirer, 2003, 122)

This resonates with Saussure, who holds that "[t]here are no pre-existing ideas, and nothing is distinct before the appearance of language" (Saussure, 112).

However, the departure from this concordance consists in Cassirer's emphasis of the primacy of the sentence over the word. If meaning in language is not derived from designated objects, but rather language constitutes and therefore precedes the objects, what can alternatively provide the meaning that is necessary to constitute the object? For Cassirer, the answer is provided by the Kantian notion of '*judgement*':

> All judgements are accordingly functions of unity among our representations, since instead of an immediate representation a higher one, which comprehends this and other representations under itself, is used for the cognition of the object, and many possible cognitions are thereby drawn together into one. (Kant A69/B94)

Essentially, a judgement is a synthesizing function of thought. It synthesizes its components according to a formal structure found in the transcendental subject. Linguistically speaking, a judgement itself, its unity and its making in the transcendental structures or modes of thought, provides its components with meaning. The functionality of judgements, for Kant, has its root in the transcendental subject. The basic functionalities of judgement include judgements of quantity, quality, modality and relation, each of which is again comprised of other, more particular judgements.

Of course, Habermas does not subscribe to a Kantian transcendental conception of meaning. However, what remains in Habermas is the essential idea that meaning is produced through an act of synthesis. But rather than the transcendental subject, for Habermas, it is the hermeneutics of communicative action that constitutes meaning. In an inter-subjective attempt of mutual understanding, meaning is developed. Its ingredients are provided by two pre-understanding subjects who shift as a result of their exposition to each other. Thus, essentially, the production of meaning is a synthetic process.

Having arrived at the center of the disagreement, we can identify Derridean meaning as resulting from différance, from a necessary play of differences, while for Habermas, meaning is produced through syntheses.

Derrida must believe in the hermeneutic incommensurability of distinct systems of meaning (e.g. the occident, an idiom etc.) in so far as they are, as such, differentiated. A full translation *between* such systems is rendered impossible (Cf. Derrida, "Relevant Translation" 176) due to the fact that meaning is constituted by an infinite play occurring *inside* such a system. For Derrida, this incommensurability leaves available as a starting point for a discussion of a system of a particular meaning only the inside. Furthermore, the infinite relativity constitutive of meaning renders it impossible to isolate specific elements to critique[11]. While Derrida does seem to consider e.g. single concepts, the overall methodology remains the same. Thus, as we will see, when criticizing any particular system of meaning, Derrida is only able to consider the functionality of the system as a whole.

---

[11] Defending the claim that Derrida *has* a certain agenda or critique–and, thus, a certain normativity–would exceed the limits of this essay.

In contrast, Habermas assumes hermeneutic commensurability of systems of meaning — after all, their having a meaning suggests that they each have been synthetically generated in communicative action once before. In an interview in the aftermath of September 11th, Habermas remarks:

> The constant deconstructivist suspicion of our Eurocentric prejudices [i.e. intercultural incommensurability] raises a counter-question: why should the hermeneutic model of understanding, which functions in everyday conversations and which since Humboldt has been methodologically developed from the practice of interpreting texts, suddenly break down beyond the boundaries of our own culture, of our own way of life and tradition? An interpretation must in each case bridge the gap between the hermeneutic preunderstanding of both sides-whether the cultural and spatiotemporal distances are shorter or longer, or the semantic differences smaller or larger. (qtd. in Borradori 36)

Contrary to Derrida's genealogical, retrospective analysis of the production of meaning, Habermas' orientation may strike us as particularly prospective. This results from the irreversibility of the hermeneutic synthesis. Once two parties engaged in communication achieve a mutual understanding, they cannot simply 'unsee'. Meaning and understanding can thus be altered and improved only by further communication. Habermasian critique thus starts from the *status quo* of a flawed system of meaning and insists on a corrective process of producing, through mutual attempts of understanding and communication, a reformed meaning.

2. Modernity

In the following, I will try, to the extent possible, to extrapolate from *Structure, Sign and Play in the Discourse of the Human Sciences*, Derrida's critique of modernity. This formulation is, in a way,

misleading because Derrida, rather than focusing his critique on modernity alone, offers a total critique of the western episteme, culture and history. The controversy over *modernity* might present itself as such only from a Habermasian perspective[12]. As we shall see, modernity, if at all, deserves special attention only to the extent that it is a climax of the totality that is the history of metaphysics.

This history of metaphysics, for Derrida, is "a series of substitutions of center for center" (*Writing and Difference* 353). This series of centers includes "essence, existence, substance, subject (…) transcendentality, consciousness, God, man" (Derrida, *Writing and Difference* 353.). Against "what has always been thought" (Derrida, *Writing and Difference* 352), the center is not self-identical. Rather, it is constituted through opposition to an elusive, intangible alterity.

Despite an exceptional, because *historically specific* passage in which Derrida credits Nietzsche, Freud, and Heidegger with the "decentering, the thinking the structurality of structure" (*Writing and Difference* 354), this decentering "has always (…) already begun to work" (*Writing and Difference* 354). Thus, although decentering thought finds its most radical expression in those authors, it is perpetual and ahistorical.

However, for Derrida, a genuine decentralization is unfathomable and unattainable. This becomes evident in several passages: "We have no language—no syntax and no lexicon—which is foreign to this history [of metaphysics]; we can pronounce not a single destructive proposition which has not already had to slip into the form,

the logic, and the implicit postulations of precisely what it seeks to contest" (Derrida, *Writing and Difference* 354) and "the notion of a structure lacking any center represents the unthinkable itself" (Derrida, *Writing and Difference* 353).

The opposition of such a necessary center and an equally perpetual, decentralizing momentum thus pervades all of history. Then what makes the moment of Heidegger, Nietzsche, and Freud historically salient to Derrida?

There is some *prima facie* evidence for the fact that Derrida ascribes some 'destructive' capacity to these writers: metaphysical "concepts are not elements or atoms, and since they are taken from a syntax and a system, every particular borrowing brings along with it the whole of metaphysics." Every metaphysical concept, i.e. every concept, is monadically permeated by the whole of metaphysics. "This is what allows the destroyers [Heidegger, Nietzsche and Freud] to destroy each other reciprocally" (Derrida, *Writing and Difference* 355 f.).

However, it is important to remark that Derrida does not stipulate the possibility of destruction *simpliciter*. Rather, destruction is possible only to the extent that it is both reciprocal and has as its object specific authorial vocabulary and not metaphysics as a whole. While Heidegger, Freud and Nietzsche can destroy *each other*, they cannot destroy meta-physical complicity in their own text, much less metaphysics as a whole. Any critical or destructive effort is dependent on metaphysical complicity, which "we cannot give up (…) without also giving up the critique we are directing against this complicity" (Derrida, *Writing and Difference* 355). The linguistic subject is thus necessarily *subject to* metaphysical language.

In line with this, Derrida argues that "the writer writes *in* a language and *in* a logic whose

proper system, laws and life his discourse cannot dominate absolutely" (*Of Grammatology* 158). This conception of language as the determinant for a writers' discourse echoes one aspect of the relationship between *langue* and *parole* that Barthes points out in his 1964 *Elements of Semiology*. In it, *langue* is defined as "essentially a collective contract which one must accept in its entirety if one wishes to communicate" (Barthes, 3). However, *langue* not only determines *parole*, but is also constituted by it. Without instances of speech, there would be no such an institution as language. Barthes therefore calls the *langue-parole* relationship 'genuinely dialectic'.

In slight tension with the idea of the subject as *subject to* language, Derrida, too, ascribes at least some degree of autonomy to the writing subject: A writer "uses them [language and logic] only by *letting himself*, after a fashion and up to a point, be governed by the system" (Derrida, *Of Grammatology* 158, emphasis mine). A writer "commands and (…) he does not command of the patterns of language" (Derrida, *Of Grammatology* 158). However, in articulating any such 'command', a writer necessarily adopts language and its commands.

In tension with the ahistoricity of Derrida's analysis, when proposing a new, eccentric methodology that does not seek 'full presence,' Derrida particularly contrasts it with interpretations centered around the (historical) "humanism and man" (Derrida, *Writing and Difference* 370) and defines this new interpretation *ex negativo* as one that is not concerned with the "inspiration of a new humanism" (Derrida, *Writing and Difference* 370). Beyond Derrida's totalizing critique, we can therefore identify a specific animosity against humanism. However, in light of his otherwise decisively ahistorical conception of the perpetual structure of center and periphery, we are justified to understand Derridean modernity as one of many elements of the totality that is the history of metaphysics.

One may ask many questions about the political content and implications of Derrida's take on the discrimination of concepts in the history of metaphysics. Is the history of metaphysics a history of 'mere' intra-linguistic discrimination? How does this justify the radical totalizing critique of western culture and language? Is Derrida's animosity to the history of metaphysics based on its connection to the "most original and powerful ethnocentrism (…) imposing itself on the world"— the logocentrism (Derrida, *Of Grammatology* 3)?

I would like to focus on one of these many question in particular. One may wonder about the emancipatory rhetoric of Derrida. This rhetoric poses the question of the normative standpoint of his own project of dismantling humanism as a constituent of the logocentrism. Is emancipation (e.g. from intellectual authority) not a characteristically humanistic ideal? If this is so, what alternative rhetoric remains available for Derrida after his totalizing critique? Formulated in this way, the problem of legitimacy arises for Derrida, precisely because he performs a totalizing, as opposed to a differentiated, critique of modernity. This challenge is mitigated, however, by the fact that Derrida does not critique modernity and history from *the outside*— in fact, he does not believe that taking the outside as a starting point is possible. Therefore, he must not provide an alternative to the object of his critique. Instead, he embraces a 'rupture *from within*' that takes place as structuralism investigates the structurality of structure in the eccentric non-locus that is language. This avoids an incoherence between the impossibility of a critical standpoint on the 'outside' and a radical and totalizing critique of the inside. The problem of legitimacy, however, is not only a problem of formal coherence. Even the articulation of an enterprise such as a rupture from within requires resources whose necessary origin is the inside, again raising the problem of legitimacy.

Habermas strongly opposes the idea of a subjective rationality and forwards the idea of an *intersubjective*, communicative rationality, which can only be actualized in communicative processes *between* subjects. For Habermas, society remains suboptimal and prone to crises not due to the individual subject's failure to achieve rationality's full potential (e.g. due an instrumental rationality caught in a dialectic functionality of emancipation and oppression (Cf. Horkheimer, Adorno, Dialectic of Enlightenment 36). Rather, Habermas considers mutual communicative efforts on both sides of the communicative process as necessary to the realization of rationality and societal progression.

In the essay *Modernity versus Post-modernity*, Habermas invokes a conception of modernity first developed by Max Weber. Modernity is understood as a process of differentiation of religion and metaphysics into three autonomous spheres: science, morality and art. Those spheres are equipped with their own criteria of validity, namely, truth, rightness, and authenticity, or beauty. This differentiation is the condition which makes possible a process of institutionalization of the spheres. As their own institutions, they could then develop their own types of rationalities, the "cognitive-instrumental, moral-practical, and aesthetic-expressive," respectively (Habermas, "Modernity versus Postmodernity" 8).

According to Habermas, the Enlightenment had two goals. The first goal was to achieve, by means of developing the spheres on the basis of their immanent logic, "objective science, universal morality and law, and autonomous art" (Habermas, "Modernity versus Postmodernity" 9). Secondly, from rationalization, enlightenment also hoped to benefit everyday life, or, in Habermas' terms, the "lifeworld." This goal, however, was counteracted by modern professionalization and institutionalization of the different spheres which, as a consequence of their professionalization, with-

drew themselves from the "hermeneutics of everyday communication" (Habermas, "Modernity versus Postmodernity" 9).

This differentiation and withdrawal, for Habermas, ground the problem of modernity. This historical characteristic of modernity — rationality's differentiation and withdrawal from 'hermeneutics of everyday communication'—seriously undermines the efficacy of rationality (which is always *communicative*) and thus prepares the grounds for social maladies.

Understanding modernity as an integral element of the history of metaphysics, Derrida's critique of it must be equally as totalizing as his critique of the history of metaphysics as a whole. As we have seen, the comprehensive scope of Derrida's considerations results directly from of his understanding of meaning. Consequently, Derrida's does not, unlike Habermas, uphold certain modern ideals (e.g. rationality), while critically reflecting upon other aspects of modernity (e.g. the differentiation of rationality and its withdrawal from 'hermeneutics of everyday communication'). Instead, Derrida "takes into consideration 'the Occident in its entirety" (Habermas, "Philosophical Discourse" 161). Derrida conceptualizes the object of his criticism as necessarily pervading history. Thus, he seeks a break, a rupture, or an opening from within.

Habermas proceeds in a more 'conservative' manner in the strictly literal sense that he appreciates and seeks to strengthen certain aspects of modernity. He seeks to utilize the emancipatory, productive power of modernity and to preserve certain ideals, while, in an ongoing effort, criticizing the enigmatic regressive elements (e.g. any obstructions in communicative processes). Habermas' approach thereby circumvents the challenge of the normative legitimacy, as his differentiated critique allows him to spare and occupy certain ideals as his own normative standpoint.

We owe credit to Habermas for offering a theory that regards tangible social circumstances. For example, I hold that, with Habermas, we can give a plausible explanation of a phenomenon of contemporary populism: its blatant denial of scientific facts. Recall that Habermas is concerned with the modern professionalization and, as a result, detachment of scientific, cognitive-instrumental rationality from "the hermeneutics of everyday communication." Fortunately, we need not speculate when it comes to this prediction. In *The Growing Inaccessibility of Science*, sociologist Donald P. Hayes measures a rising incomprehensibility of the vocabulary of prominent scientific journals to non-specialist between 1930-1990. In this article, he identifies both a differentiation and a professionalization of the scientific sphere. While recognizing the possibility that this is conductive to scientific progress, he holds that, at the same time, it "must surely diminish science itself. Above all, it is a threat to an essential characteristic of the endeavour — its openness to outside examination and appraisal" (Hayes 739 f.).

From a Habermasian perspective, Derrida, "[d]espite his transformed gestures, in the end … too, promotes only a mystification of palpable social pathologies; … he, too, lands at an empty, formulalike avowal of some indeterminate authority [of text as opposed to presence]" (Habermas, "Philosophical Discourse" 181).

Both theories present valuable conceptual tools to discern aspects of the history of the occident in general and modernity specifically. The contentious point seems to be the scope of analysis: Habermas presents a more tangible theory. This has the benefit that it can instruct a political strategy. However, Derrida does not seem primarily concerned with the usefulness of his theory nor does he seem to focus on the specific characteristics of modernity as a particular historical phase. Thus, perhaps, Derrida would simply disregard this kind of

criticism as inapplicable and based on a wrong measure. He would presumably accuse Habermas of misidentifying certain ideals as desirable, which are, instead, operative in an oppressive cultural economy.

However, by entertaining suspicions about every present aspect of culture, Derrida is guilty of an appropriation. As soon as something does as little as presenting itself, it is declared as part of a cultural economy, and thus as cooperative with a system of oppression against an unfathomable absent alterity. His theory does not conceptually allow for this Other to partake in presence, i.e. for autonomous representation.

Conclusion

As we have seen, both Derrida and Habermas present their own conceptions of meaning, replacing the earlier, naïve model of signification. For Derrida, meaning emerges in a play of differences, or what he calls the différance. From this understanding of meaning, we have explained that the Derridean approach has to start from within (due to the income-mensurability of systems of meaning) and why it has to criticize the whole (due to the infinite play that is the différance). For Habermas, meaning emerges in a synthetic process of communicative action. He takes a prospective approach due to the one-directedness of synthetic processes, and is able to offer a differentiated critique of modernity due to the possibility of systematic improvement guaranteed by the possibility of effective communication. Thus their respective conceptions of meaning inform their approaches to modernity. Additionally, I have argued that we ought to adopt a Habermasian theory of modernity on the basis of its tangible orientation. Derridean theory of meaning has been criticized for its implausible political implications such as the insistence on the impossibility of trans-cultural communication.

Works Cited

Adorno, Theodor W.; Horkheimer, Max. *Dialectic of Enlightenment*. Translated by Edmund Jephcott, Stanford University Press, 2002.

Barthes, Roland. *Elements of Semiology*. 1964. Hill and Wang, 1968.

Borradori, Giovanna; Jacques Derrida; Jürgen Habermas. *Dialogues with Jürgen Habermas and Jacques Derrida*. Translated by Luis Guzman, The University of Chicago Press, 2003.

Cassirer, Ernst. *Die Kantischen Elemente in Wilhelm von Humboldts Sprachphilosophie* in *Aufsätze und kleine Schriften. Gesammelte Werke*. Translated by author, Volume 16, Meiner, 2003, Hamburg.

Derrida, Jacques. *Margins of Philosophy*. Translated by Alan Bass, The Harveer Press, 1982, Sussex.

Derrida, Jacques. *Of Grammatology*. Translated by Gayatri Chakravorty Spivak, John Hopkins University Press, 1997, Baltimore and London.

Derrida, Jacques. "What Is a "Relevant" Translation?", Critical Inquiry, Vol. 27, No. 2 (Winter, 2001), pp. 174-200, Translated by Lawrence Venuti, The University of Chicago Press, 2001.

Derrida, Jacques. *Writing and Difference*. Translated by Alan Bass, Taylor & Francis e-Library, 2005.

Fraser, Nancy. *Unruly Practices: Power, Discourse, and Gender in Contemporary Social Theory*. University of Minnesota Press, 1998.

Habermas, Jürgen. *Beyond a Temporalized Philosophy of Origins: Derrida*. in *The Philosophical Discourse of Modernity*. Translated by Frederick Lawrence, Polity Press, 1987, Cambridge.

Habermas, Jürgen. "Modernity versus Postmodernity", *New German Critique, No. 22, Special Issue on Modernism* (Winter 1981), pp. 3-14, Duke University Press, 1981.

Habermas, Jürgen. *The Liberating Power of Symbols. Philosophical Essays*. Translated by Peter Dews, The MIT Press, 2001, Cambridge, Massachusetts.

Hayes, Donald P. "The Growing Inaccessibility of Science", *Nature*, Vol. 356 (1992), pp. 739-740, 1992.

Kant, Immanuel. *Critique of Pure Reason.* 1787, Second edition, Translated and edited by Paul Guyer, Allen W. Wood, Cambridge University Press, 1998.

Knellwolf, Crista; Christopher Norris. *Literary Criticism. Twentieth-Century Historical, Philosophical and Psychological Perspectives.* Cambridge University Press, 2001.

Saussure, Ferdinand de. *Course in General Linguistics.* Translated by Wade Baskin, Columbia University Press, 1959.

Heidegger's Neglect of and Dependence on the Body in Being and Time

JACOB JORAM

JOHNS HOPKINS UNIVERSITY

Abstract: Martin Heidegger explicitly refuses to discuss the ontological nature of the body in *Being and Time.* For him, to discuss the body is to mistake Dasein for something objectively present, an unbecoming description of the being with a relationship in its being to its being. Herein, I endeavor to show that Heidegger not only missteps when he does not discuss the body, but he actually needs it and presupposes it in his ontological treatment of Dasein as care. The body is what allows for our particular kind of being-in-the-world as beings that take care. I will show how Heidegger relies on the body in his ontology, and attempt to give an account of the body that remains in the framework of *Being and Time.*

Heidegger's lacking conception of the body in *Being and Time* (hereafter cited as BT) is both astonishing and vexing. He either presupposes it in the structure of Dasein, and so does not feel the need to give it a separate treatment, or he finds discussion of the body confounding in the overall discussion of Dasein. Perhaps it is actually because of the former that the latter is true; that is, to focus on the body at all would confound the analysis of the being of Dasein as something other than what it actually is, and the body is thus presupposed or ignored as obviously being part of Dasein. Even as Heidegger presupposes body in his descriptions of being-in-the-world, he does not ever give the body the full ontological treatment it deserves as a part of Dasein's being. Heidegger takes discussion of the body to be misguided because to think of the physical nature of Dasein is to consider it as an objectively present object (BT, 117). Further, an analysis of the physical nature of Dasein would perhaps miss the point and reduce Dasein to a merely biological being, rather than a being with a particular relationship to being. Dasein, which is a being oriented toward a future of possibility, cannot just be an object that is present-at-hand, or a biological system determined by its physical characteristics. This would be to make the mistake of all previous ontologies in considering the human being as situated first and foremost in the present, between past and future. To put the human being in this position misses the fundamental structure of Dasein's being.

This need to situate Dasein in possibility unfortunately leaves the body in shadow, an obviously-there but nonetheless unarticulated part of Dasein's being. And yet one might wonder how Dasein as being-in-the-world relates to the world at all (this point is one Merleau-Ponty discusses at length in *The Phenomenology of Perception*.) For Heidegger, the world is disclosed to us, most primordially as ready-to-hand, and then more deficiently as present-at-hand. By our dealings-in, or our projects, we come to know the world by way of tools. Beings in the world that are not Dasein are first disclosed in their uses for our projects. But how can we discuss readiness-to-hand without a discussion of the hand? Is the hand also disclosed to us in the way other beings are, or is the hand taken for granted as our way to disclose the world – or is it even disclosed at all? It cannot be that we come to know our own bodies as objects separate from us. We must not revert to a phenomenological dualism of self and body by refusing to discuss the nature of the body in our overall being. Yet Heidegger seems to do something like this when he neglects the body in his ontology.

What I intend to show is that Heidegger indeed presupposes the body in his ontology, and he needs the body in order for his ontology to work. While Heidegger specifically expresses his intention not to discuss the body in *Being and Time* (BT, 108), he is mistaken in not doing so.

If one of the purposes of *Being and Time* is to elucidate the fundamental ontological structures of Dasein, to leave out the body is a gross misstep. I will show that it is only by body that we encounter the world as beings that relate to the world through care, proving that Heidegger must assume the structure of the body in order to develop his existential analysis of being-in-the-world. What I will show is that body is not only presupposed in Dasein's ontological structure, but also that it is an unstated and neglected structure of Dasein. I will then attempt to give an account of the body that is true to the ontology of *Being and Time*, staying as true to Heidegger's vocabulary as possible.

Body as Heidegger understands it is inessential to Dasein because to discuss it reverts back to traditional ontologies and misses the question of being altogether. The body is certainly a part of Dasein, it is true, but this does not tell us anything about the *being* of Dasein: "But in the question of the being of human being, this cannot be summarily calculated in terms of the kinds of being of body, soul, and spirit which have yet first to be defined. And even for an ontological attempt which is to proceed in this way, some idea of the being of the whole would have to be presupposed" (BT, 48). The being of Dasein is not the sum of the being of its various parts. Rather, some idea of the being of Dasein must already be thought in order to understand the being of its parts. Certainly this is true, and yet, as we will see, the body is what makes some of the ontological structures of Dasein possible. One of these ontological structures is care: "[T]he expression xz is used in this inquiry as an ontological term (an existential) to designate the being of a possible being-in-the-world…because Dasein itself is to be made visible as *care* [Sorge]" (BT, 57). Care is a fundamental structure of Dasein's being-in-the-world. It is how Dasein manifests in the world, not as an object or a regular being among other beings, but as a being that is always first and foremost involved in the world. But how is this care made possible? It is certainly one of the ontological structures of Dasein, but we can trace our ability to care by way of our bodiliness in the world.

Dasein has world ontologically, and Dasein also always has a body, at the very least as an ontic feature. More precisely, Dasein *is* its world. Here I will embark on an analysis of being-in-the -world that will conclude that Dasein *is* its body. It might immediately be objected that body is not a separate structure of Dasein. Insofar as Dasein is its world, its body is also part of world. Body is thus not separate from the structure of world, and to think of it as so is to separate Dasein too much from world. The point I will show here is that although body and world might be encountered at the same time and possibly in similar ways, they are still separate existentials of Dasein. I will first give an exposition of the ways in which we encounter beings in the world. Beings in the world are encountered most primordially in terms of their utility, not as merely things that are objectively present. To be sure, beings may be encountered as objectively present as well, but this is only when the network of interconnected involvements of things breaks down in some way. We thus encounter them in a deficient way. Most fundamentally, however, we encounter other beings in the world as tools that are at hand for us. We do not think of them as things per se, but as ways for us to take care in the world, and complete various goals and projects. Even a simple projection like picking up a cup of tea still encounters the cup not as a thing but as a way of taking care in the world (i.e. we want tea, we pick the cup up without really considering the cup itself, and we drink the tea). Our involvements in the world disclose world for us in the most fundamental way.

Further, no being in the world stands in isolation. Each being is part of a network of interconnected associations, such that one being discloses a totality of beings. Heidegger gives the example of the pen, which discloses paper and

the desk and the room as a whole (BT, 68). Dasein thus never encounters just the pen, but all the other beings that are connected to the pen. We do not necessarily notice this network, but we encounter it most primordially in this way. Not to notice it can still be an encountering. We also never encounter just one thing at a time. We encounter many things all at once; a tree never stands as a single thing that we encounter, thus unlocking a gateway to many other connections which form a landscape. We encounter the tree, the grass, the bird, and the sky all at once.

It is when the pen breaks (or the hammer, to use the usual example) that we suddenly notice the paper, the desk, the lamp, the room, because the web of connections between beings has been broken. Suddenly the pen, no longer ready-to-hand, becomes *conspicuous* (BT, 73). It is unwieldy; it cannot be used to complete the project originally undertaken. It is a kind of unhandiness (BT 73). But we might also lose the pen, and thus it is missing, or *obtrusive* (BT, 73). And finally, we might find that we do not have a pen, but a carrot, but what we really need is a pen. The carrot is *obstinate*, it stands in the way of the care being taken (BT, 74). In each of the three cases mentioned, the involvements and interconnectedness of beings break down, and we view not only the being that is broken, missing, or incorrect, as something unhandy and merely objectively present, but the entire world becomes viewed from this distant and thematic position.

The descriptions used for beings in the world can in a way be applied to body, but with different effects. For example, we encounter the body when we are taking care in the world. Body, like other beings, may not consciously be perceived, and it fades into the background of our dealings in the world. We do not ignore it, just as we do not ignore other beings in the world when we encounter them in taking care. We do not, however, encounter body as a tool. Beings in the world are encountered in terms of their utility. They are ready-to-hand. But is it appropriate to speak of the body as being ready-to-hand? Is the hand ready-to-hand, or do we have some other relationship to our bodies that is not the same as with the world? The term ready-to-hand seems to insinuate it is to the hand that other beings are first disclosed, and this can be generalized to the body as a whole. Insofar as we are taking care, we may do this with any part of our perceptive body. It is through the body, then, that world and other beings are disclosed in their most primordial fashion. It is through the body that we come to know the world. It is also via the body that world is disclosed in the various deficient ways of being present that are described above. Any sort of perception of world that we have access to is through the body.

What does this actually mean as far as our relationship to the body is concerned? We are beings-in-the-world, after all. Does it not follow, then, that our bodies, being a part of us, are in the world as well? Heidegger says that Dasein is its world, and if body is part of Dasein, is body not part of Dasein's world as a sort of innerworldly being? In a sense, yes it is. Insofar as Dasein is its world, body is part of Dasein's world as well. But this does not necessarily mean that we encounter our bodies in the same way that we encounter other beings in the world. We have a much more intimate relationship with our bodies than with other beings in the world. This is shown in the term ready-to-hand, as explained above. The body facilitates our involvements in the world by being that to which world is always disclosed. It is by our bodies that we can take care in the world at all. Heidegger argues that Dasein is only spatial insofar as it takes care in the world (BT, 367). I agree with this claim, but it is through body that we can take care at all. Without body, there is no taking care. Nothing could be handy without the hand. Even as we use tools in the world to take care, it is first through our bodies that we can encounter these tools. Our bodies orient us spatially in the world as care. They could not do this, of course, if they were not in the world as part of Dasein. But it is not the same as world, and it is certainly not

encountered in the world in the same way as other beings in the world.

Dasein encounters its body first and foremost as that which orients us in the world as care. We are spatial beings insofar as we take care in the world. The body, much like the other beings in the world, is not perceived as merely present, although it is not simply ignored, either. The body fades into the background as part of our everyday dealings-in; we do not typically think about our bodies as we go about our daily lives. Body in many ways seems much like part of the network of interconnected involvements that make up world and the relationships between beings in the world. Certainly, insofar as Dasein is a being-in-the-world, body is also part of Dasein's involvements. But it is part of them by being that which allows for these involvements at all. Dasein could not be involved in the world without body, and even if it could, it would be a very different sort of being-in-the-world. It would be a sort of phenomenological idealism, with the mind of Dasein somehow interacting with world psychically and taking care without the modes of perception that it has with body.

Much like beings in the world, we can encounter our bodies in deficient ways. We can repurpose Heidegger's own vocabulary to describe these deficient modes of encounter. For example, the hammer breaks, and it becomes conspicuous, unhandy. Suppose the hand breaks. We can also call the hand conspicuous. The breakdown of the body, as happens in injury and disease, often frustrates our taking care, and it does this even more intimately than when other beings become conspicuous. When the body breaks, it does not disrupt the interconnectedness of beings in the world, yet it prevents us from encountering them in their primordial mode of being. If the hand is broken, it does not matter that the hammer, the nail, and the wood are all intact. Without the working, healthy hand, the hammer cannot disclose itself by its utility. Instead, the hammer becomes something objectively present, because we cannot encounter it as

a tool without the hand to grasp it. Further, the network of involvements the hammer belongs to is disclosed deficiently as a set of objectively present objects. None of these innerworldly beings themselves become conspicuous, but the body can easily be said to be conspicuous here.

If a part of the body is missing, as in amputation or developmental problems where a part of the body was never there to begin with, we could call the body (or a specific part of it) obtrusive. Despite the fact that these situations are not the same, we can still call the body obtrusive in each case. An amputated limb would, of course, be a loss for us, whereas the absence of a limb we never had would not a loss. Still, even in the latter case, the world is disclosed in a way that is different from if we did have the limb. Even if we do not feel a loss in the absence of this limb, we might still find ourselves in situations where to have the limb would inevitably allow the world to be disclosed in ways that it cannot without the limb. Again, the world cannot but be encountered with this different orientation occurring. Nothing has necessarily changed about the beings in the world themselves, even if they all happen to be in their rightful place and not missing for any given dealing-in. The body, on the other hand, will be receptive to disclosure of the other beings in the world in a deficient mode. The mode will reveal other beings as objectively present, insofar as the part of the body missing would typically be needed for the other being to disclose itself in terms of its utility.

It is more difficult, it seems to me, to discuss the ways in which the body might become obstinate, since we are always with our bodies. In what way can the body get in the way of our taking care that is different from when the body is conspicuous or obtrusive. Perhaps one distinct way body might get in the way is if we do not yet have a skill, or are particularly bad at a type of dealing-in. Suppose we are bad carpenters, but need to hammer together a piece of wood. It is not the fault of the hammer being

broken, missing, or the wrong tool that we cannot complete this project. Our body is not properly equipped with the skill needed to take care in this particular way. The body gets in the way by being clumsy, and we are frustrated in our taking care by this clumsiness. This might be one way in which the body might be obstinate. Here again, the body is recognized by us as deficient, as are the other beings in the world. However, just like the other examples, the body and the world are disclosed to us in very different ways. When the body is conspicuous, obtrusive, or obstinate, the world need not necessarily be these things as well. We have proven again that the body and world might be disclosed at the same time, but not necessarily in the same way, and we have shown that the body is indeed separate from world insofar as the body and the world are constituted and rendered meaningful by Dasein in different ways.

What is most profound here is that we have shown why we must consider body and world to be two separate but nevertheless necessary components of Dasein. Body might be given at the same time as world, and thus given *with* world in a sense. But body cannot be the same as world. To return to one's body is precisely to withdraw from the world. Insofar as the body withdraws from the world in pain, the body is clearly distinct from world. Pain makes us, as bodies, withdraw from world, and this pain discloses our spatial limitation insofar as we are bodily.

We have discussed the difference between body and world, with the body as being how we can be-in-the-world as Dasein. Dasein must be bodily, but bodily how? There are not many clues in *Being and Time* about how we should think of body. However, we might be able to use this empty account to our advantage. Perhaps there is more insight in Heidegger's refusal to discuss the body than initially appears. At least we might be justified in expanding on this point while still remaining true to his ontology.

Dasein is *bodily*, and many sensations and experiences for Dasein are also bodily. Pain and suffering are two such experiences that can manifest in a bodily way (joy and pleasure are others). Indeed, it does seem like intense pain and suffering place limitations on the person feeling them. They may disclose to us a sort of limitation of our possibilities. Imagine for instance, a person in pain so intense that she cannot think about (or really be aware of) anything else. Presumably this person is not thinking about her ontological structure of possibility and relating to it in a way that allows her to participate in that possibility. In other words, she is so consumed by a relation to her limitation (her finitude) that she cannot be meaningfully aware that she also has a relationship to her possibility. I do not think that pain of that intensity is necessary to make one particularly aware of the limiting nature of their body. I do think however, that some bodily sensations are more prone to making one aware of the body's limiting nature than others, especially pain and suffering.

Dasein does not just have body. It is not merely embodied. Dasein is *bodily*. "Bodiliness" is always given with Dasein; it is part of the ontological structure of Dasein to be bodily. As such, we can be-toward-body. We may do so authentically or inauthentically. When we are being-toward-*death* authentically, we are confronted with our ownmost possibility. When we are authentic toward our death, the mode of attunement through which possibility, finitude, and world are disclosed is anxiety (BT, 251). Anxiety is the mode through which we authentically view our limits as finite beings, but also our possibilities as beings that are to-be. Similarly, it seems that with regard to the body, pain is the mode through which finitude and possibility are disclosed spatially. I think this finitude is something similar to the finitude we experience when we are authentically being-toward-death. I will speak more about this later.

Heidegger does not want to speak of the body because he believes doing so mistakes Dasein for something objectively present, and I agree that to speak of the body in this way is incorrect. The body is not objectively present. But the body has presence, nonetheless. Pain, which has no object, cannot disclose objective presence. But pain discloses something immediate, and if it is the mode through which the body as limitation is disclosed, this might mean the body is present in some way as well. Pain might then be something analogous to a mode of attunement to the body, a way through which the body is disclosed. In fact it is also a way in which the *possibility* of the body is disclosed. Consider for instance exercise or physical activity. We put our bodies through a certain amount of pain that discloses to us what the limitations of our bodies are, but at the same time the possibilities of our body are disclosed. What is the body capable of? We find out through pain. Pain is not just a sensation, it is an existential mode of disclosure for Dasein, a mode of attunement through which possibility is revealed. I want to term modes of bodily disclosure such as pain as feelings. Feelings reveal the body.

Death represents a limitation on Dasein of a kind, in that it is because of death that we are finite beings. The kind of finitude that death discloses is that we are beings determined by a not (BT, 283). We cannot, as it were, get behind ourselves to discover our origin. Nor can we ground ourselves in anything outside ourselves. We are beings thrown into a world, ungrounded and without reason. In this sense, Dasein is said to be a temporally finite being.

Pain, similarly, represents limitation of Dasein in a spatial sense. How does it do this? Pain, or perhaps more encompassingly, vulnerability (this includes pain, illness, any form of suffering) shows the bounds of our body, its limitations, and its distinctness from world at the very least at the brute existentiell level. We choose from here on out to use the term vulnerability for the fact that everyone, as long as they are bodily, is vulnerable. Death is an existential of Dasein because death is always a possibility for us, and thus anxiety is a mode of attunement available to everyone. But not everyone feels pain; there exist conditions that prevent one from feeling physical pain at all. If we want to make the body a structure of Dasein, we must find a mode of attunement that is available to all Dasein. For the moment, vulnerability is the best we can do, even though I acknowledge it is not a necessary part of Dasein's structure. It is merely certain in an existentiell sense.

Vulnerability, which reveals the limitations of our body, does so by returning us to our body. When we are in pain, or ill, we are faced with our body as a limit of us. But more interestingly, we view ourselves as distinct from world. As Elaine Scarry argues in *The Body in Pain* (hereafter cited as BP), pain has no object, there is no pain "of" or pain "for" anything (BP, 162). There is nothing in the world that is the object of pain. And yet, pain is only possible by virtue of our bodiliness. Pain is bodily. Vulnerability, too, is bodily. But nothing in the world is the object of pain (nor for that matter, is our body an object of pain). It is only that pain is a mode of attunement of our existential body. When we feel most aware of the limitations of our body, we withdraw from the world, or at the very least, we are reoriented in our bodiliness to the world. When the hammer breaks, the world is no longer ready-to-hand, and the relations disclosed in our concernful dealings break down. Tools in the world break, and the world is disclosed as present-at-hand. What if the hand breaks? What is disclosed? World, certainly, as we previously discussed, but also body. Body as present in some way calls out to us, calls us back from the world, even as we are always in the world as bodily. I do not mean to say we come out of world and into body. As bodily we are always in a world. But we certainly can withdraw from the world. We can no longer relate to it

with the same concernfulness, or we might take care in a different way.

Importantly, the breaking of the hand does not distance us from our body in the way the hammer breaking distances us from world. Rather it does the opposite. The failure of the hammer discloses the hammer deficiently as present-at-hand, but the failure of the hand discloses body more intimately as present, immediate body. Even as we are distanced from the involvements the body took part in, we are still brought back to our bodies. Thus the actualization of our vulnerability makes us aware of our bodies rather than allow them to fade into the background of our dealings-in.

What is ontically nearest to us is ontologically often what is furthest from us (BT, 108). Heidegger gives the example of glasses which are sitting on a person's nose, thus being quite literally near, but in terms of de-distancing through care, they are much further away than the pictures on the wall that are ontically quite far. Useful things are often furthest away from us, despite literally being at hand, in the ontological sense, they fade into the background of our projects. The body is something like this. Dasein is bodily, is always bodily, and it is only through body that Dasein can take care in the world. Insofar as Dasein is its body, body nearest to us in one sense. But insofar as the body is part of our taking care, it is perhaps even further away than the useful things it grasps. If we are not thinking of the hammer as a thing when we use it, we are almost certainly thinking less about our hand that grasps the hammer.

Can the body ever be what is nearest to us in an ontological sense? To the extent that the body can be disclosed to us or become a project for us in itself, this is possible. Pain, for example, brings us out of our concernful dealings in the world and back our bodies as that which is in pain. When the body breaks down in some way, we notice it as orienting us toward itself, it becomes that which we project towards, even as it is what orients our projection in the first place.

While it is true that pain and brokenness can lead to other possibilities that we can project towards, in the moment of that pain or brokenness, we are not projecting toward these possibilities.

Does the body in pain orient toward the body itself or toward nothing at all? Scarry would suggest that pain has no object. Heidegger, I'm certain, would say we are always already projecting toward something at all times. We could not be concernful beings in the world if we were not always projecting in some way. Yet we may not be oriented toward anything at all in the case of pain. In a similar way to how anxiety is about nothing, pain is also about nothing. Clearly anxiety is a mode of attunement nonetheless, and pain is a feeling nonetheless. But anxiety, even as it is disclosive, has no object. Anxiety is the mode of attunement that we are in when we are authentically being-toward-death. Being-toward one's ownmost possibility is what discloses all other possibilities.

Pain works in a similar way, but with regard to our spatial orientation in the world. Pain is of nothing at all, and yet it still invites possibility. As Scarry argues, "Any state that was permanently objectless would no doubt begin the process of invention" (BP, 162). Further, Scarry points out that "it is especially appropriate that the very state in which he is utterly objectless is also of all states the one that, by its aversiveness, makes most pressing the urge to move out and away from the body" (BP, 162). The body in pain is very closely linked with the imagination, which is of course full of possibilities. Perhaps what pain shows us is that we are nothing without the body, because we grasp helplessly for an object outside of the body to which we can cling. The world, insofar as we are in pain, becomes quite small; in fact it becomes nothing more than body. We withdraw from the world and our involvements in pain, and possibility, ever a part of Dasein's being, becomes a flailing of trying to escape the body via the imagination. We want to escape the body, but instead, we are confronted immediately with the body. It is true that when

the body is in pain, we are opened up to a whole new set of possibilities for how we might heal the pain. But in the moment of pain itself, and insofar as we are in pain, we are shown nothing more than our finite bodies. The projecting into the imagined world fails to find an object for pain, and thus we find that pain cannot orient us toward anything. It is pure sensation, without meaning.

Even as we find ourselves tethered to body when we are in pain, body is not the object of pain. True, the body hurts, but it does not hurt *toward* or *for* anything. Pain, even as it makes us aware of our bodies, does not take the body for its object. Thus pain remains objectless. The body is not objectively present for Dasein. Its attempt to orient us in pain may fail, but this failure does not reduce the body to an object. The body simply orients us toward nothing. The important conclusion here is that the body can never be objectively present for Dasein, even when its orientation is toward nothing. Further, pain orients us toward a not, in a similar way to how anxiety does this. When we are in the feeling of pain, we are momentarily ungrounded in a spatial sense. The orientation toward the world fails to find an object of pain, and thus ungrounds us from the world in a sense. We are still in the world, of course, but we are without involvement to the extent that we are in pain. Spatially, we become ungrounded; not re-oriented, but *disoriented*. In this sense, we see our spatial finitude via the feeling of pain.

The body, when its vulnerability is actualized, cannot but orient us in a different way toward the world, and when it is in pain, this reorientation fails. Even more intimately, Dasein will feel its body become uncanny, as Svenaeus points out in his analysis of illness (Svenaeus 2010). We will be pulled out of average everydayness by pain and illness, and we are drawn back to a conscious noticing of our bodies. Even as we are drawn back into our bodies, we are distanced from the involvements and projects our bodies were a part of. Insofar as Dasein is its

projections, Dasein becomes distanced from itself in pain and illness. It is a simultaneous drawing into and distancing from oneself, and this feeling will no doubt be quite strange. The uncanniness of our most intimate connection to the world is indeed very striking. Alienation from the body and from Dasein's projects leaves it with a new set of unfamiliar projections that it can undertake. These might include the way in which we dress our wounds or take medicine or go to the hospital. Many projections no doubt remain intact; one can pick up a cup of tea in sickness just as one can do in health. But the tone of the phenomena is different in this case. Picking up the cup in each case signifies something different, and each experience takes on a different mode of attunement to the projection. One can also pick up a cup after having suffered a stroke or a broken hand, and this might be done with ease or difficulty. However, the significance of the projection is different than if a person had not had either of these injuries, and again, the mode of attunement to the projection is something different.

The body is always given as part of Dasein's structure, and it is spatially what orients us in the world as care. Through the hand, beings in the world become handy. For many reasons, the body is never something objectively present for Dasein. This is because bodiliness is how we can be-in-the-world at all, and the body is never to be taken as an object of that orientation toward the world. The body certainly can be a project for Dasein, as in when we exercise, or when the body breaks and we attempt to heal it—but this does not make it an object, since we project through our bodies. This strange paradoxical relationship will not be explored further here, but it goes to the point of how our bodies are more intimately part of us than any object could ever be. Moreover, pain, which is the spatial analog of anxiety, can never take the body for its object. What it does do is disorient us in the world, thus showing our finitude and ungroundedness in a different way. We have

shown the body to be an essential structure that is always given with Dasein, which makes being in the world possible, and which can reveal to us our finitude in a spatial way. We have done so while staying within the Heideggerian system as much as possible, thus showing that Heidegger indeed makes a mistake to neglect the body in his ontology.

Works Cited

Heidegger, Martin. *Being and Time*. Edited by Dennis J. Schmidt. Translated by Joan Stambaugh, State University of New York Press, 2010.

Scarry, Elaine. *The Body in Pain*. Oxford University Press, 1987.

Svenaeus, Fredrik. "Illness as Unhomelike Being-in-the-World: Heidegger and the Phenomenology of Medicine." *Medicine, Health Care and Philosophy*, vol. 14, no. 3, 2010, pp. 333–343., doi:10.1007/s11019-010-9301-0.