

Claremont Colleges Scholarship @ Claremont

CMC Senior Theses

CMC Student Scholarship

2016

A Preference for Freedom: Kantian Implications for an Incompatibilist Will and Practical Accountability

Maggie Miller

Claremont McKenna College

Recommended Citation

Miller, Maggie, "A Preference for Freedom: Kantian Implications for an Incompatibilist Will and Practical Accountability" (2016). *CMC Senior Theses*. Paper 1228.
http://scholarship.claremont.edu/cmc_theses/1228

This Open Access Senior Thesis is brought to you by Scholarship@Claremont. It has been accepted for inclusion in this collection by an authorized administrator. For more information, please contact scholarship@cuc.claremont.edu.

Claremont McKenna College

**A Preference for Freedom:
Kantian Implications for an Incompatibilist Will and Practical Accountability**

submitted to

Professor Paul Hurley
and
Dean Nicolas Warner

by

Maggie Miller

for

Senior Thesis
Fall 2015
November 30th, 2015

Abstract

This thesis aims to provide a coherent account of free will and practical grounds to prefer it. Its goal is to develop a pragmatic understanding of agency by which to hold individuals morally accountable. The paper begins with a critique of P.F. Strawson, whose seminal paper “Freedom and Resentment” bypasses the question of free will altogether in its claims about morality. Subsequently, it proceeds to a defense of incompatibilism that traces an argument through the existing literature. From this position, it claims that neither Strawson nor traditional compatibilists can provide an account of morality that is reliable or well enough defined to play the role required of it.

Instead of being left with hard determinism, however, Kant opens the door to a metaphysics that exists outside of our epistemological limits. Rather than derive an account based on this metaphysics, the necessary characteristics of a free will are derived from an account of morality and proven to be possible using Kantian epistemology. The paper concludes by positing three distinct reasons to prefer a free will framework to a deterministic framework, provided our inability to answer the question empirically. These draw on Pascal’s Wager, William James’ “The Will to Believe,” and inference to the best explanation.

Acknowledgements

While the process of drafting this thesis began just months ago, the ideas behind it have their roots in an existential crisis sparked by my very first philosophy class. This is a subject I care very deeply about and which has affected my own life in profound ways. For this reason, I am especially grateful to everyone who has helped me along this process.

First and foremost, thank you to my tremendous reader, Professor Hurley, whose white board comments were an integral part of structuring this project. Due to your thoughtful guidance, this thesis developed into a much more meaningful and pragmatic work. Your time, patience, and candid feedback have been invaluable.

Thank you to Professor Kreines, whose guidance navigating the depths of Kant's transcendental idealism helped me stay afloat. Classes with you are always a privilege.

Thank you to Professor Davis, whose expertly taught Introduction to Philosophy course redirected me down this unexpected path two years ago.

Thank you to my mom, dad, and brother for encouraging me through all of this philosophical nonsense and indulging me as I alternately disappeared into the depths of writing and reappeared to spew arguments at you. Thank you as well to Caroline Hays for your daily poetry, steamed milk, and support. Your kindness and energy astound me regularly.

Finally, thank you to Aviv Caspi, who has devoted untold hours to this exploration. You are [morally] responsible for grounding me through the logical conundrums and the emotional quandaries and for that, I hold you accountable. I am deeply grateful.

Table of Contents

Introduction.....	5
The Determinism Thesis: A Necessary Question.....	6
Objection 1: The Monk.....	9
Objection 2: The Spectrum of Accountability.....	11
Objection 3: The Illusions.....	12
Ought Implies Can: Refuting Compatibilism.....	15
Returning to Practical Compatibilism.....	17
Kantian Indeterminacy.....	20
Free Will from the Other End.....	23
1. Ought implies can.....	24
2. We can be both influenced and free.....	25
3. We have choice when we think we do.....	28
From Possibility to Preference.....	31
1. Pascal’s Wager.....	32
2. The Will to Believe.....	33
3. Inference to the Best Explanation.....	34
Clarifications and Delineations.....	36
Conclusion.....	38
Bibliography.....	39

Introduction

To the non-philosopher, free will is usually taken as given. When a parent scolds their teenager for lying about their whereabouts or a person leaves their spouse after boredom prompts them to be unfaithful, rarely do considerations of whether they had a choice in the matter come to mind. Alternatively, the parent would likely spare no praise when that child's hard work pays off in an excellent report card, while the spouse will almost certainly feel gratitude for an unexpected meal prepared at the end of a long day. The child and spouse were wrong, or right, in each scenario and deserve to be treated as such. Our intuitions about free will, whether implicit or explicit, play into our judgments about accountability. The people in our lives deserve praise or blame because they seem to have some sort of ownership over their actions.

Among philosophers, however, the discussion about free will has lasted as long as it has because, intuitively, it seems so necessary and yet, logically, so contradictory. Furthermore, upon closer examination, the contradictions seem necessary. Thus, whole branches of philosophy have disregarded the concept as impossible and attempted to make do without it. With the incredible advances of the physical sciences and neuropsychology, it has become quite popular to accept a form of determinism. This has spawned countless theories by compatibilists, aiming to reconcile determinism and our moral intuitions, and hard determinists, who have abandoned morality altogether. Compatibilists, however, run into equally as many contradictions in compromising free will as they did trying to resolve it and hard determinists paint a picture of a world dramatically out of line with our experience. Both seem to lead to dead ends and many have given up the question altogether.

Kant's transcendental idealism has reopened the door for an incompatibilist account of free will and real morality. Through this lens, free will may not be confounding because the concept is absurd, but because the human epistemological standpoint is limited. Kant's Antinomies identify other unresolvable questions that lead us to consider the existence of a reality beyond the scope of what we can know. Agency, perhaps, also has its roots here.

This paper has a huge scope, traversing ideas surrounding compatibilism, idealism, and ethics of belief. I make the case that the question of whether or not we have free will is a critical one and that we can know how it must work practically for us to regain morality. Furthermore, I argue that Kant's revelations show it may be impossible to know with certainty whether or not this incompatibilist form of the will exists, but that we have legitimate reason to prefer it.¹

The Determinism Thesis: A Necessary Question

There are some in the discussion of free will who would dismiss the issue of determinism altogether. This line of thinking has drawn its prevalence from P. F. Strawson, whose "Freedom and Resentment" has fundamentally altered the conversation. He writes:

¹ I note before beginning that, by incompatibilist, I mean incompatible with the thesis of determinism, that any current or future state is the necessary consequence of some past state plus the laws of nature. I recognize that there are significant scientific advancements that have found evidence counter to this thesis (notably, quantum physics), so it bears noting that while I argue against classical determinism, the argument extends to any non-deterministic worldview that does not allow for metaphysical support for alternative possibilities whose outcome is determined by an agent, e.g. a world with real randomness.

“The central commonplace that I want to insist on is the very great importance that we attach to the attitudes and intentions towards us of other human beings, and the great extent to which our personal feelings and reactions depend upon, or involve, our beliefs about these attitudes and intentions” (Strawson 5).

By his thinking, the question of the truth of determinism is irrelevant because, regardless of the answer, we cannot but experience these “reactive attitudes”. These include feelings tied to accountability, such as blame and approbation or gratitude and resentment. They are so intrinsic to our nature that the idea of discarding them to function in a world we believe to be deterministic is impossible. Strawson argues that to engage on a personal level with other human beings – to hold them accountable and feel positive or negative emotions based on their actions in relation to this standard – is ineliminable from our experience. Thus, these attitudes, regardless of their source, demonstrate the foundational nature of a practical morality, drawn from the uniformities of attitudes that we cannot but accept as true. If I am to assert a theory of incompatible moral accountability, I must first refute Strawson to put the question of determinism back on the table. Specifically, I must show that the difference between a deterministic world and a world of free will has practical implications for our behavior.

There are some for whom my critique may be unnecessary. Most basically, if you believe that there is intrinsic value to truth itself, you will be underwhelmed by Strawson’s argument. Regardless of the practical implications, you will continue to participate in the debate, in search of knowing what really is. It seems that moral truth, in particular, has special importance and is especially worth determining.

But for those who are less concerned with these theoretical realities, Strawson has a compelling case. He would allow non-philosophers to continue about their lives without worry that they are acting in fundamentally illegitimate ways. For those with pressing practical matters (e.g. how to create policy, how to treat those who do not act in accordance with laws), there is no need to wait on philosophers to come up with answers. It is enough that we have the reactive attitudes we do and that we cannot function otherwise.

For those who draw on this easy comfort, I wish to resist your inclination for reasons other than the intrinsic value of truth. Strawson has simplified the problem to the point of error and made logical leaps to which he is not entitled. There are, in fact, practical questions that Strawson leaves unanswered and ignored implications of determinism, both of which can only be settled by a coherent conception of free will.

Below is a summary of Strawson's argument:

Premise 1: Regardless of determinism, we cannot but occupy the standpoint of reactive attitudes.²

Premise 2: Reactive attitudes presuppose accountability, which is constitutive of agency.³

Premise 3: Accountability presupposes moral standards.⁴

² "A sustained objectivity of inter-personal attitude, and the human isolation which that would entail, does not seem to be something of which human beings would be capable, even if some general truth were a theoretical ground for it" (Strawson 10).

³ "The personal reactive attitudes rest on, and reflect, an expectation of, and demand for, the manifestation of a certain degree of goodwill or regard on the part of other human beings towards ourselves; or at least on the expectation of, and demand for, an absence of the manifestation of active ill will or indifferent disregard" (Strawson 13)

Conclusion: The Determinism thesis is irrelevant to the practical derivation and implementation of moral standards.⁵

I would object to this argument at three points. The first critiques Premise 1. The second critiques Premise 3, revealing some of the practical difficulties of Strawson's claim. Finally, the third regards Premise 2 and highlights the major flaw in both Strawson's theoretical case and those of more traditional compatibilists.

Objection 1: The Monk

Imagine a man who in most ways is perfectly ordinary, with the exception of an anger problem, which flares up on occasion. With encouragement from his friends and co-workers, he enrolls in an anger-management class in an attempt to curb his outbursts. The instructor coaches him to, upon feeling his rage build, remove himself from the situation and rationally intervene in the emotional process. He learns to tell himself that his feelings are unfounded and imagine possible explanations for the situation that makes him so overwhelmingly angry. With practice, he is able to identify his anger early and stop himself from acting out in these sorts of situations.

Over time, as he gets better and better, he finds that he is able to more instinctively diffuse his anger through rational explanations of the situation. Eventually, this anger dims to the point of near elimination. He no longer feels controlled by his instinctual emotions.

⁴ "The generalized and non-generalized forms of demand, and the vicarious and personal reactive attitudes which rest upon, and reflect, them are connected not merely logically. They are connected humanly." (Strawson 14)

⁵ "No such sense of 'determined' as would be required for a general thesis of determinism is ever relevant to our actual suspensions of moral reactive attitudes." (Strawson 17)

Imagine now that this man, having overcome his anger issues, decides that *all* of his reactive attitudes are unfounded. Because he rejects things that he thinks do not have value, he quits his job, bids farewell to his friends, and moves to an isolated temple. There, he works day and night to have rational interventions in all of these reactive attitudes. Instead, he finds value in the appreciation of natural beauty and the development of his new hobby of cooking.

If you accept the plausibility of this final situation, it denies Strawson's first premise that our reactive attitudes are ineliminable from our experience. There may be practical objections to this thought experiment, but you can hold many of them constant as you please. If worried about his ability to achieve this in a lifetime, grant him an extremely long life. If worried about his long upbringing in modern society, assume he begins this process very early in childhood. If you grant theoretical possibility of the case, however, it puts this lifestyle within the range of goals of practical morality. Historically, there seems to be precedent for dramatic shifts of this sort, even across cultures. Dodds suggested that Homeric Greek society may have been dominated by shame and lacked strong feelings of guilt until much later in the Archaic age. This calls into question the inconceivability of lacking these attitudes.

If you do not grant the full case, even acceptance of the partial case will go a long way to support further objections. The situation, up to his departure for monkhood, is not an unusual or implausible one. There are innumerable recipients of this sort of emotional treatment who would affirm its aid in their ability to control, first, the effect that their emotions have upon their actions and ultimately, their emotions. This is not limited to anger. The critical point here is that we have the ability to move substantially

in the direction of reducing the influence of these reactive attitudes in our daily lives, and even the reactive attitudes themselves, should we want to. This will come back to help us.

Objection 2: The Spectrum of Accountability

This second objection applies to Premise 3, that accountability presupposes moral standards, and its implications for the conclusion. While our feelings of blame and gratitude imply that we have an intuitive moral standard, many of us seem to have differing moral intuitions. An intuition that there is a moral standard generally is not enough to know how to practically apply it. Without a theory of free will, we can know nothing about whose intuitions are wrong or right and must just assume that our “common-place” and unconsidered instincts about morality are correct. Let me elaborate.

Use of reactive attitudes to support a moral standard might not be a problem if we all had the same moral intuitions. While Strawson points out many commonalities between us (e.g. feeling resentment towards those who intentionally mistreat us, while making excuses for those who do not know any better or who are not themselves at the time), there are those whose intuitions differ greatly. Some claim, for instance, that you have an obligation to know that physically harming another person is wrong regardless of your upbringing, while others will morally excuse, for a variety of reasons, the vast majority of those who have committed crimes. Even those whose intuitions align generally may differ in their applications. Of two people with very middling moral intuitions, one may resent and honk at the car that cuts them off on the on-ramp, while the other may

recognize that the driver may be in an abnormal rush or have been disabled by their blind spot.

Debates among any of these groups (forgiving versus condemning) will come down to the question of accountability. To resolve them, it is not enough to abstract from our intuitions about accountability; they seem to differ so greatly (recall also the case of the ancient Greeks) that they cannot underwrite moral accountability. It matters in these cases whether the person in question had sufficient control over their behavior and could have done otherwise. Did the person have reasonable access to information or experiences that would have allowed them to make a different decision?

This is a metaphysical question with practical implications. Answers will support the intuitions of some, while denying *and causing practical adaptations to* those of others (e.g. forgiving a driver you may not have or holding accountable a culture you might have excused). We have a spectrum of reactive attitudes available to us and where we fall will depend on our framework of accountability. We cannot rely on our emotional reactions for answers; they are the consequences, rather than the source. We need outside grounds. While I have not yet shown that free will can have these sorts of answers, it seems a difficult task to undertake without it.

Objection 3: The Illusions

With the first two objections as support, we reach the final and most fundamental objection: if determinism, all Strawson has done is show that morality is an illusion that is impossible, or at least very difficult (if you accept the full monk objection), to escape. It is a flaw faced by all philosophers who attempt to give us a practical way out of the

question of determinism. Hilary Bok is one of these philosophers and, if determinism, Strawson is making a claim that parallels the one outlined in her Pocket Oracle thought experiment. Strawson claims that, regardless of determinism, we cannot but have reactive attitudes and Bok claims that, regardless of determinism, we cannot but take ourselves to be agents. Both of these are insufficient for morality. Because they reduce to the same flaw, I will refute Bok, who lays out the case explicitly, and then extend the criticism to Strawson.

The Pocket Oracle argument posits a mini oracle in your pocket who knows everything that has ever existed or happened and all the laws of nature. Assuming determinism, it can predict what you will do in any given moment. Bok's claim is that, even if you could consult the Pocket Oracle before making a decision, its knowledge of what you would do does nothing to undermine or usurp your practical need to decide. Thus, we still have agency.

Our intuition inclines us to agree with Bok. We like to think that practical reasoning is something that is "ours", that we can claim credit for, because it is such an intimately personal part of our experience. The Pocket Oracle is incapable, however, of saving agency (and with it, morality), only the illusion of it. As Bok recognizes, it excludes a critical influence: whatever it tells you. An oracle that is also a cause in your decision-making process compromises the integrity of its prediction. For a spiteful person, it would always be wrong because the person would always do the opposite. Bok's argument rests on this inability to predict. It does not prove the power of the case though, as Bok would like it to, but rather the flaw in the experiment.

The Pocket Oracle is not incompatible with a separate, 3rd-person oracle that simply observes and silently predicts. If determinism, then this 3rd-person oracle always knows what will happen because it can account for the Pocket Oracle's prediction. It would also be able to predict infinitely far out into the future because it could generate further states based states it has already predicted. It is this oracle that has moral significance. Specifically, it tells us that we can have no action-guiding morality because we do not guide our actions. Imagine that you are walking through your neighborhood and come to a fork in the road. Out of curiosity, you ask the Pocket Oracle which direction you will go. Based on all information to this point, it tells you that you will go left. Because you are generally obstinate, you opt to go right. Throughout this interaction, the 3rd-person oracle would know not only that the Pocket Oracle would say left, but also that you would then go right. It would be absurd for you to then turn to the 3rd-person oracle and say, "I know you knew that the Pocket Oracle would say 'Go left' and also that I would turn right. I know you have known this for years, extending back before I was even born. But I *still* have agency over my decision." For this is what follows from Bok and many compatibilists' theories. From the moment you were conceived, you never could have done any differently than you have. Thus, no one could say you ought to have done anything otherwise. There is no "ought", because there is only one "can".

To make this claim more fully, we must take a detour to contextualize the claim of "ought implies can". In doing so, I will narrow the landscape to exclude not only the practical compatibilists, like Bok and Strawson, but also the theoretical compatibilists.

Ought Implies Can: Refuting Compatibilism

The interpretation of “ought implies can” is often a primary differentiator between incompatibilism and compatibilism. As such, there is an extensive literature on the subject. I will not recreate an entire argument on its behalf here, but rather trace my line of thinking through existing authors who have done credible and much more extensive work to prove its necessity.

Free will is most commonly defined as the “ability of persons to exercise control over their conduct in the manner necessary for moral responsibility” (McKenna and Coates, SEP). While most compatibilists, including Strawson and Bok, will claim that this control is reconcilable with determinism, incompatibilists will counter that there are key aspects of agency missing from their accounts. There are two traditional models of incompatibilism: the principle of alternative possibilities (PAP) and sourcehood (sometimes known as ultimate responsibility). Some incompatibilists focus on one or the other, though many, myself included, argue that you need both for free will. This is because agency is possible if and only if you have more than one option available (PAP) and if the option that is played out is traceable to reasons that originate in you, the agent (sourcehood).

Most compatibilists try to weaken or deny the PAP. Perhaps the most common way to weaken the case is to add to the end of “I could have done otherwise” the clause “had I wished to do otherwise.” Authors like Hume and Hobbes have famously made the case that a truly free agent is an unintelligible concept and that the freedom to do as one

pleases is sufficient for moral responsibility.⁶ Many have later gone on to counter this claim as empty since, if determinism, what you wish to do is equally as determined.⁷ The most famous attempt to deny PAP altogether is by Peter Frankfurt, who aimed to show that we hold people accountable even when the outcome is inevitable.⁸ His thought experiment fails, however, as Ginet identifies, because, by either available interpretation, it generalizes distinct states of affairs or begs the question of determinism. In accordance with these authors, I maintain that a libertarian reading of PAP is necessary for agency.

With regard to the sourcehood model, it seems that no compatibilist can accept any version of a source or cause that originates in us as persons. They need, though, some sort of agency. Thus, in a similar vein to the efforts to weaken PAP, the most common counterarguments tie sourcehood to our character, aligning our actions with first or second order desires, rather than a cause that begins uniquely in us.⁹ Pereboom's manipulation argument is a sound rebuttal of these claims, revealing that no amount of abstracting away from sources obviously outside of ourselves can give us accountability.

Unfortunately, Pereboom himself goes on in the even more questionable direction of trying to salvage some sort of morality with no free will at all. He and many compatibilist philosophers have given up hope of any real agency and compromised on its logical necessities because they do not see a way to resolve the issues at hand. I

⁶ For reference, see Hume, "An Inquiry Concerning Human Understanding" or Hobbes, "Of Liberty and Necessity".

⁷ Authors include Chisholm and van Inwagen.

⁸ In Frankfurt's thought experiment, Jones has set out to kill Smith and Black has rigged a device in Jones' brain such that, should Jones fail in his resolve, he will be forced to do so anyways. In this case, should Jones *not* decide to spare Smith, we will hold him morally accountable, regardless of his inability to do otherwise.

⁹ See again Hume and Hobbes.

would like to spare him and others that dead-end direction by opening a new door that would allow us to maintain incompatibilism, while leaving room for free will.

Returning to Practical Compatibilism

Given this context, determinism is not a Pocket Oracle, but rather a 3rd person oracle that knows absolutely what you will do, even if you do not, and can explain even the most unconscionable acts, even if you must feel resentment. All Bok has done is proven the necessity of 1st person indeterminacy; no information you can have can tell *you* what *you* will do. All Strawson has done is suggest that we cannot eliminate the reactive attitudes (which is true if you do not grant the monk). They are both just necessary illusions.

Illusions, however, do not need to affect our behavior. In fact, we live with a number of illusions. The most classic example is a straw that appears bent in water. As with agency and reactive attitudes, we cannot help but perceive the straw as it seems to us. But we can know that it is not bent. We do not need to treat it as bent. To add implications, replace the straw with a fishing spear. If we did not adjust for the spear's bent appearance in the river, we might skewer our foot instead of a fish.

So what does it mean to account for determinism and thus, lack of morality? Here, Strawson would protest that this is not a coherent option, thus grounds for dismissing the question. In his own words,

“Our question reduces to this: could, or should, the acceptance of the determinist thesis lead us always to look on everyone exclusively in this way? For this is the only condition worth considering under which the acceptance of determinism

could lead to the decay or repudiation of participant reactive attitudes” (Strawson 10).

He clearly answers no to both. But to “could...the acceptance...”, a critic could draw upon the monk objection. To “should the acceptance...”, the critic has an even stronger response. Strawson has misconstrued the entire question. The determinism thesis does not tell us to adopt the objective attitude in all circumstances. The logic he is implicitly using to get us there is that, as is, we employ the objective attitude when we encounter people or situations that are excused from our moral framework. Thus, since determinism excuses everyone all the time, we should always adopt the objective attitude.

The problem with this argument is that determinism does not excuse all situations from a moral framework; it disregards morality entirely. It is not, as Strawson suggests, that no one is ever responsible, but rather that there is no responsibility. Any framework from which all situations are excused is trivial. It is incoherent to ask what we “should” do in case of determinism because there is no should. To exist in a deterministic world is not to eliminate reactive attitudes, but to discredit their moral significance. Some people may try to adopt the objective perspective and some may, to spite the illusion, do some thing radically opposite. Many may be paralyzed by the inability to decide. To quell this indecision, they would almost certainly be forced to adopt the non-reflective state, in the footsteps of many philosophers before them. Hume, for example, wrote of his encounter with this problem, saying:

“I am confounded with all these questions, and begin to fancy myself in the most deplorable condition imaginable... [So,] I dine, I play a game of backgammon, I

converse, and am merry with my friends. And when, after three or four hours' amusement, I would return to these speculations, they appear so cold, and strained, and ridiculous, that I cannot find in my heart to enter into them any farther” (Hume, *Treatise* 97)

From a practical standpoint, we must continue to perceive ourselves as agents, just as we must see the straw in the water as bent, but it has no compelling moral force on what we do. This perception may play a physical causal role, but outside of the non-reflective state, we will find no grounds for it. In times of sincere moral quandary, we can derive no answers from it. You will do what you will do, but it does not matter, because all decision-making is arbitrary.

Furthermore, against Strawson, the world in which we perceive ourselves as being determined is substantially different than one in which we perceive ourselves to have free will. In the former, those who adopt the non-reflective standpoint are granted a largely comfortable existence. They can succumb to their illusions of agency and morality, so long as it is in their best interest; as soon as it becomes difficult or uncomfortable, they have an easy out. They have no particularly grounded claim to make against the spitefully “immoral”, other than that their actions are bad, based on values questionably obtained.¹⁰

¹⁰ In a world without moral accountability, it becomes an interesting question to derive value. It is much more difficult to place intrinsic value on qualities or experiences that people enjoy, because we are simply “cogs in the system” and seemingly no more important than any other non-human cog. The easiest means is to grant a god who has superimposed objective value or to arbitrarily value common human interests. There is much more to say here, but does not look to be a promising or compelling force in the current debate.

In the latter, however, we have reason to put faith in a moral standard that compels us to action whenever we have choice. It is a standard that we can apply to ourselves and to those around us. We do not get freebies when it gets hard and can have legitimately grounded expectations of each other.

Any real form of morality must have a compelling claim on our *actions*, not just our emotional instincts. Since they are not necessarily tied, Strawson's and Bok's do not make the cut. Any legitimate form of free will must enable one to act in accordance with this free will *or not*. That is the only means of saving real accountability.

If determinism, then we lose this account. Thankfully, we have an alternative.

Kantian Indeterminacy

It is unclear whether we can understand the realities of the world and, even if we can, whether we have any way to confirm that we have reached this sort of truth. As much effort as we invest in trying to understand the world, we have no means of determining whether it all sits within the bounds of our possible knowledge. On one hand, it seems egotistical to suggest that all we possibly could know is all there is to know. On the other, it is hard to imagine what the limits of that scope are and even harder (perhaps definitionally) to imagine what could exist beyond them.

The two primary approaches to collecting new knowledge are rationalism and empiricism. While some philosophers, particularly famous ones, tend to fall into one camp or the other, the bulk of the population falls somewhere in between. The bounds of our knowledge seem to be those things that we can observe and deduce. If the world is reduced to what can be learned through just those two means, determinism seems the

likely outcome. To reach the theoretical extremes of both would allow us observe everything there is to know about a given state in time and to deduce all the laws of nature. After that, there is nothing else to serve as input or mechanism, so we can explain everything that ever has happened and predict everything that ever will.

Kant, however, has given us reason to think there may be more. For while it is very difficult to hypothesize that there may be further existence outside of what we can know if everything observable fits into this framework, when there are things left unexplained, then we have reason think there may be something more. Kant identifies some of these unexplained aspects of the world in his Antinomies. The Antinomies seem to be not only unexplained, but also unexplainable. They are cases that expose the areas in which reason and experience break down, where we can seemingly prove the truth of two incompatible things. There is some controversy over individual cases, but if you accept any one of the four examples as a valid dilemma, then you have reason to think that there exist things beyond what we are able to know.

Kant believes the reason we cannot answer these questions is that our knowledge is constructed from, and thus limited by, our concepts and our intuitions. Our intuitions are our frameworks of space and time, through which we construct the phenomenal world. Reality – “things in themselves” – however, exists outside of these intuitions, in the noumenal realm.

Free will must lie in this latter realm. If it exists, it is necessarily incomprehensible; this necessity is rooted in how we determine truths about the world. Our means of understanding is through laws that predict states and behaviors. But if we could create laws that explained the metaphysics of free will, it would be predictable and thus un-free.

Many Kantians will accept such a statement. Problems arise, however, because free will needs to get us to morality. Moral accountability is even written into the definition. While free will must be unknowable, morality must be knowable. This latter statement is intuitive. If morality is a standard that outlines the behavior of the ideal person and one that we can apply as a guide to our own actions, then we must have a full picture of that standard. We cannot hold someone accountable for acting in or out of accordance with a standard of which they can have no idea. The question becomes: how do we derive a standard of morality from an unknowable account of free will?

This is where Kantian metaphysics cannot help us. After accepting a fundamentally unknowable theory, many Kantians go on to try to know at least as much as they can about it. While accepting that they may not be able to get all the way to true knowledge, they are tempted to draw an epistemological box around it and say “look, it exists somewhere in there. While we cannot know what it is, we can say something about what it is not.”

This may be a moderately viable tactic in the search for a little more truth, but it is certainly not enough in the search for morality. In this effort, there are substantial and, I will argue, insurmountable problems that prevent us from getting to a substantive account of free will. One major stumbling block is timelessness. Recall that free will must lie outside of our intuitions, which include time. Thus, for one trying to get a general metaphysical idea of free will, it seems at first glance reasonable to hypothesize that our self-initiated decisions happen simultaneously at a point either before time or encompassing all of time. We have, though, no idea what it means for the will to exist outside of time. Timelessness is not a stop at a train station. Timelessness is to throw out

the train and the track and the station altogether. We have, necessarily, no idea what could be left.

We run into a further problem when we try to imagine how our will could insert its way into our practical reasoning. Reason can only work with the information that it is provided with and this information seems to come only from the phenomenal world. We need an input in our decision-making of which we have ownership. Thus, our decision-making must involve reason, but not be limited to it. Again, we can have no idea what this noumenal input might be or how it might work.

If we continue to posit hypotheses in answer to these questions and others, we only offer strawmen to be hacked at. Where we cannot prove them wrong, we are left with only vague suggestions at what free will is. We cannot possibly put enough faith in metaphysical speculation about realms outside of our bounds of comprehension to derive a morality we are willing to apply not only to ourselves, but also to others. Morality requires a confidence that this theorizing cannot provide. We need a different tactic.

Free Will from the Other End

The problem that metaphysically-oriented Kantians encounter is that they feel they need to give an account of how the will works, even after stating that they cannot. This is not, however, the only coherent means of supposing morality. Rather than deriving it from metaphysical uncertainties, I propose an altogether different approach to the problem. Let us instead ask: what are the necessary characteristics of a free will that gives us morality? Let us approach it from the other end of the problem, analyzing what the will would need to do to get us morality and then proving that it is possible that this

sort of will exists. If we can come to a coherent conception, then I argue that we have legitimate reason to prefer it to our deterministic alternatives. I will do the first in this section and the latter in the section to follow.

This becomes a surprisingly simple question, particularly when compared to the elaborate attempts to characterize the will through metaphysics. To practically hold people accountable, we (1) need them to in fact have agency and (2) need to know when they are acting as agents. In fact, from these requirements, I will simplify the parameters the will to just three:

1. Ought implies can
2. An agent can be both influenced and free
3. We have choice when we think we do

The first two are derived from (1) and the third follows from (2). From here, I will elaborate upon each and explain both why it is necessary and why we cannot prove it to be impossible.

1. Ought implies can

I have already discussed the need for this first condition and traced my line of thinking to authors such as Chisolm, van Inwagen, and Ginet. To briefly restate, agency is possible if and only if you have more than one option available and if the option that is played out is traceable to reasons that originate in you, the agent. To say that one “ought” to do something implies that there must be at least one other option truly available to them, which is morally inferior. If their decision is determined by factors beyond their control, as determinism suggests, then they cannot be responsible.

We cannot rule out this agency because we cannot know that we do not have multiple options. We may be tempted to think otherwise because, as we consider past experience, there has only ever been one outcome. One outcome, however, does not mean one option. Until we are able to perfectly predict behavior in instances that seem, in the moment, undetermined, we should not assume that what happened was inevitable. Some wholehearted believers in modern psychology might claim that we are well on our way to the ability to make these sorts of predictions, but this is a leap to which they are not entitled. Psychology has done an extraordinary job making sense of many of our instincts – the reasons we are inclined in one way or another or why we react intuitively to a stimulus – but it has not given us reason to determine conclusively that our rational decision-making is reducible to these instincts. As long as perfect prediction is impossible, there is room for a will to have theoretical influence.

2. We can be both influenced and free

When Kant opened the door to a reality in which “things in themselves” exist, a transcendental universe to which we have little or no access, the door to all sorts of multi-world, phenomenalist, and epistemological interpretations burst open. For much of history, including Kant’s own lifetime, readers attributed to his transcendental idealism a two-objects (or two-worlds) view.¹¹ Things in themselves exist in the outside universe, while we perceive only appearances. The two are metaphysically distinct because appearances are simply mental representations. While appearances obey causal laws,

¹¹ Recent proponents include Grant and Strawson, “The Bounds of Sense: An Essay on Kant’s Critique of Pure Reason” and Van Cleve, *Problems from Kant*.

things in themselves exist outside of our minds and are not so limited. Thus, our wills could exist freely in this outside world yet not be incompatible with causality.

This interpretation has seen significant critique, particularly given its implications for free will. It seems that even though we have allowed for things in themselves to be free, if everything we can perceive is virtually distinct from their real counterparts, there is too much of a separation between the two to conflate their accountability. It helps me not at all in any practical sense if I am free only in a world to which I have no access.

More recently, a view of Kant has emerged that classifies these distinctions as “two aspects” of the same object.¹² A variety of distinctions have been made, but the most promising have emerged recently and are based in epistemology rather than metaphysics.¹³ Rather than suggest that appearances and things in themselves are separate, these authors suggest that the human standpoint is inherently limited by our a priori intuitions of space and time and therefore can only sensibly interact with things in themselves through this epistemic filter. Thus, there may (and in some cases, certainly do) still exist aspects and relations between things in themselves that we are incapable of perceiving.

This is perhaps the interpretation of transcendental idealism that leaves the most ambiguity. It gives us next to no metaphysical foothold from which to consider things in themselves. The uncertainty that it builds into our human standpoint, however, does

¹² Questions exist about whether it is fair to attribute this interpretation to Kant himself. I do not take this up here. Whether Kant’s own view or a more recent amendment is an interesting issue, but not relevant to the discussion at hand.

¹³ Authors who make this distinction include Allison, “Transcendental Realism, Empirical Realism, and Transcendental Idealism”, Rosefeldt “Dinge an sich und sekundäre Qualitäten”, and Allais, *Manifest Reality: Kant’s Idealism and His Realism*.

nothing to undermine its possibility of being true. In many ways, it helps achieve a coherent picture of an agent. On the subject, Allais writes:

“We think that deterministic causation threatens freedom of the will when we are in the grip of a certain metaphysical picture... In my view, part of the point of transcendental idealism is to undermine the grip of this metaphysical picture” (306).

Two-aspect theory allows for the laws of nature, as we know them, to still affect us in the ways we traditionally imagine. It also, however, leaves room for them to be contextualized in a world that we cannot fully know. Allais continues:

“Laws of classical mechanics predicting the interactions of billiard balls are not broken when someone picks one of the billiard balls up, because the laws simply have nothing to say about whether that will happen. The laws make predictions about what will happen on the assumption that no other explanatory factors are in play, but it is not a part of the laws of classical mechanics that they are the only things that explain why anything happens” (306).

In this way, it is possible that the laws of nature affect a large portion of what we do and how we behave, as they do the objects around us. In some circumstances, they can be used to predict our behavior. But there are large swaths of time in which they may not be the only factor at play: when we employ our will. Two-aspect theory allows for this other influence. This is the influence that brings in moral accountability. The question becomes, when is this influence a factor?

3. *We have choice when we think we do*

The first two characteristics of free will we must assume get us to an agent who is able to choose among a range of morally variable actions. But to hold anyone accountable, particularly ourselves, we need to know *when* we have options and the ability to select between them. To get to this conclusion, we must consider what we actually hold accountable. It is not, as some propose, the actual action. It is instead the decision made among perceived options.

To illustrate, consider Frankfurt's famous objection to the PAP, to which I objected earlier. In this thought experiment, Jones has decided to kill Smith. Black has a vested interest in Smith being killed, but does not want to do it himself. To ensure the outcome, he implants a control in Jones' brain; the implant is inert as long as it is off, but once switched on, will plug into Jones' nervous system and ensure that he kills Smith. Black will leave it off as long as Jones intends to kill Smith. If he changes his mind, then he will flip the switch. Jones does not, in fact, waiver in his intentions and does indeed kill Smith.

Frankfurt's argument here is that Jones does not have an alternative possibility, yet we still want to hold him accountable. This argument, however, excludes free will from its framework. Jones *does* have an option: choose to kill Smith or be forced to kill Smith. Though resulting in the same physical outcome, this is a morally significant option! Note that, had Jones opted not to kill Smith, we would *not* be inclined to hold him accountable, because while he had no option not to kill him, he did not want to.

What physically occurs sometimes lies outside of our control. Free will posits that our influence lies in our decision-making, which can then become a causal force in the world (though not the only one). This control pinpoints accountability.

Think back to the two-aspect theory. It allows for laws of nature to play a role, but also for another means of influence that lies outside of our epistemic limits. For moral accountability, this influence must be a force that could not be predicted by the Pocket Oracle and must be a factor when we make considered decisions (as opposed to reactions). Decisions exist where we perceive ourselves to have options. Thus, we have choice where we think we do. This theory is consistent with our current scientific understanding as well. We have little trouble predicting complex controlled experiments regarding objects and physical forces. We have even made incredible progress predicting behaviors of people reacting to stimuli, particularly under stressors that inhibit rational thinking. We struggle though to make much headway at all in explaining or predicting the decisions of people making thoughtful, considered judgments. While some make chalk this up to the limits of current science, it seems plausible to think that it may lie beyond the limits of science altogether. We cannot know until we achieve it, so until that point, it remains uncertain.

We cannot rule this out for reasons similar to the first characteristics, surrounding “ought implies can.” If determinism, then this power of decision-making is an illusion. But if free will, then it is not. The only means by which to suggest this is an illusion is to be able to reliably predict the outcome of these decisions based on previous states and the laws of nature, the rules of determinism.

“Choice exists where we think we have it” may open up a couple lines of practical critique, which I will address briefly here. First, what about people suffering from severe mental health disorders? Many cases can be written out because most of these disorders severely inhibit a person’s rational decision-making power, leaving them to instinctual reactions in which there is no apparent choice.

Some disorders, however, limit apparent options but still leave more than one. What of these cases, or those in which the subject’s apparent options are limited by lack of exposure? Here, the subject is still accountable for deciding on the best of the options available to them. If they sincerely do not know that a morally superior option is possible, then they are not responsible for doing it. Many might exclaim, however, that we are responsible for knowing what our options are and the reasons to do or not to do them. Particularly in the age of the Internet, it seems we have heightened obligations surrounding knowledge, given its availability. I do not deny this. It is difficult to imagine a scenario in which you can hold all variables constant with the advent of smartphones and portable computers, so for the sake of simplicity, imagine a couple in the woods. While you cannot hold them accountable for, at Time C, continuing their hike through a fragile alpine tundra, if they had no idea of the difference or the implications, you certainly can hold them accountable for their actions leading up to Time C, e.g. at Time A, opting not to do due diligence in advance of embarking into unfamiliar public property or, at Time B, choosing to walk by signs that they knew may have important information. When speaking informally, you might say that they should not have walked through the tundra, but when speaking technically, you should be

referring specifically to the other times when options were available and yet they chose the sub-optimal one.

If we can accept these three characteristics of free will as possible, then we have created a possible picture of the will that grants us an accessible standard of morality.

From Possibility to Preference

By invoking Kant and leaving so much of reality outside of our possible knowledge, I have left open a huge number of theoretical realities, of which perhaps only one grants real free will. As long as it is possible, however, we have reason to prefer it. In fact, we have several.

It may seem that to consider whether or not we have free will begs the question. The answer depends on whether the notion of a preference relies on a concept of free will. I propose that it does not. To prefer is simply to want one of two or more apparent options as best aligns with your utility function. (Note: this does not presume utilitarianism. I do not suggest that your preference is in all scenarios the best option.) Regardless of the truth of determinism, preference looks and is the same.

Additionally, by “prefer”, I do not mean “believe free will to be true”, or “believe determinism to be false”. I have given no grounds to think that. I have instead given us reason to refrain from ruling out free will. In this case, in which either is possible, to prefer free will is to prefer the framework to a framework of determinism and to act as though it is true.

To continue, I will propose three separate groundings for preferring free will to the multitude of other realities. Of them, you need to accept only one. They are:

1. Pascal's Wager
2. The Will to Believe
3. Inference to the Best Explanation

I will elaborate upon each.

1. Pascal's Wager

Pascal's wager is a very basic means of weighing the costs and benefits of believing or disbelieving a claim. Pascal's original application to the question of the existence of a Christian God prompted substantial criticisms due to his over-simplification of the problem and questionable recommendations for implementation.¹⁴ His theoretical framework for analysis, however, independent of his application to God, is a sound one. The free will question, insofar as it is the only way to grant morality, is in fact a perfect candidate for analysis. The question is based on the assumption, outlined in earlier sections, that a metaphysically possible version of free will is the only means of achieving true morality. Thus, all alternatives are morally arbitrary and insignificant.

Your options are as follows: If you choose to act as though we have free will and we do, you earn moral approbation for your good deeds. If you choose to act as though we

¹⁴ Pascal has a further claim that, once you have determined the option that is in your best interest, you should condition yourself to believe in it wholeheartedly, regardless of your initial position on the subject. I do not advocate for this. As previously stated, to employ Pascal's framework is conditional upon the idea that you accept both options as possible and part of a currently unanswerable question. Furthermore, choosing to act as though determinism is false does not preclude you from accepting it as a metaphysical possibility. You can believe that both are possible, while letting only one influence your actions.

have free will and we do not, it does not matter. If you act as though we do not and we do not, it also does not matter. If you act as though we do not and we do, however, then you are morally culpable for all your wrong deeds committed in the non-reflective state (referenced earlier).

The only upside yields from the assumption of free will and the only downside yields from the opposite. The other options are morally insignificant. There is no incentive to prefer the determinism thesis. Even any intrinsic value of truth does not factor in if we grant that Kant has precluded us from knowing conclusively one way or another. Thus, we should prefer free will.

2. The Will to Believe

In his paper “The Will to Believe”, William James proposes that there are some questions in which we cannot determine conclusively, but which demand an answer. In these situations, he suggests that we are justified in believing as we please.¹⁵

James defines the relevant questions as those that are living, forced, and momentous. The free will debate is clearly all of these. The case is very much alive in the minds of most philosophers and even most non-philosophers, capturing our curiosity in cutting edge neuropsychology, questions of criminal justice, and even literature and movies that play with our intuitions about the subject. It is forced, because there is no neutral standpoint between the two and the way we act in any given moment assumes one or the other. Even if you ignore the question, as Strawson does, placing faith in your reactive

¹⁵ For reasons I will not go into here, I disagree with aspects of his thesis; I have attempted to control for these aspects, however, so similar critics need not automatically disregard this line of argument.

attitudes implicitly assumes free will. Finally, it is certainly momentous, for what you decide determines whether or not your life and any of your decisions are meaningful.

Having held other things largely constant, including plausibility, scientific explanation, and my own experience, I prefer the framework that gives me a wider range of options: the ability to act in accordance with either my preferences or morality, when they do not align. I prefer the one that gives me something to strive for and accountability for whether I make it or fail. By this account, I cannot appeal to those who would rather have a way out or easy access to comfort. But I can appeal to those who are inclined to believe that they possess a free will, but are not sure they can in good conscience act as though they do.¹⁶

3. Inference to the Best Explanation

The application of inference to the best explanation in this scenario is a difficult one because the scope is so vast. By holding many things constant between the two frameworks, however, we can do a credible job. The key areas of comparison will be (1) consciousness/agency and (2) basic metaphysical questions. These aspects of our experience are some of the most difficult to account for, the most hotly debated, and the most telling parts of an evaluation.

To tackle (1) completely would be to engage fully with the entire philosophy of mind debate. Can physical/brain states completely account for mental states? To put it briefly, the conversation is inconclusive. For some dualists, it will not be much of a stretch to

¹⁶ Were I to go on to derive a substantive view of morality from this version of free will, it is possible that I could appeal to those not already inclined to prefer free will. This remains an open area for investigation.

posit a free will existing outside of the bounds of our knowledge because there is already much explaining of our mental life left to be done that is left unattended by physicalism. Of them, for those who think that consciousness cannot be accounted for by the laws of nature, determinism has already been ruled out. Non-dualists, however, will argue that consciousness could very plausibly arise out of a purely physical world governed exclusively by laws of nature. For these philosophers, the illusion of free will is no obstacle; while Bok's Pocket Oracle may not have yielded morality, it did show that the experience of free will is common to both deterministic and non-deterministic realities. While determinism posits this experience as a systematic error in our perception, they would argue that it is a full description and its simplicity is enough to make it preferable.

In this case, we must turn to basic metaphysical questions of the sort raised in Kant's Antinomies, particularly 1-3 (on space/time, the foundations of matter, and causality). These are the questions that prompted us to look for free will outside of the bounds of our possible knowledge in the first place. They, unlike topics of consciousness and other modern science are not simply unexplained; they are contradictions. They are contradictions that arise, not out of lack of knowledge, but out of logical incoherency. While determinism seems a promising framework from which to extend our knowledge, it seems unsound if we are to resolve such dilemmas. We must therefore allow ourselves a framework that can account for these issues described by Kant to one that limits us to our reason, which got us into these tangles in the first place. From this angle, determinism, rather than indeterminacy, seems unlikely, if not impossible. Like Newton's laws, while simple, it does not seem to have the explanatory capacity to account for the intricacies of the world in which we live. Explanatory capacity

supersedes simplicity. When we give up determinism, we land in the realm of explanations that lie outside of our possible knowledge. Among these of potentially equal explanatory power, the simplest way to account for our experience of free will is to posit that we do in fact have it.

If any of the three grounds for preference outlined above appeals to you, then you have reason to act as though you have free will.

Clarifications and Delineations

I will clarify before moving on that these reasons to prefer *only* give us grounds to assume these basic principles of free will and not any substantive moral claims. The question at hand is whether *something* matters, as opposed to *nothing* mattering. This cannot be extended weigh substantive moral theories against each other. Perhaps this theory of free will logically entails just one moral framework, and while this would be an interesting conclusion, that is not my claim. I argue simply that we should prefer a metaphysically possible, incompatibilist version of a free will, which exists whenever we make intentional decisions. This in itself moves us a significant distance in the debate. It gets us morality in some form and, while it does not tell us how to use that morality, it does tell us when it applies. This becomes very important in practical applications.

Also worth noting is that it becomes very difficult to refute me. It is much easier to argue against compatibilism, which lies within the bounds of our knowledge, than Kantian incompatibilism, which does not. Even if it looks unlikely, as long as it cannot

be proven wrong, then it seems you must prefer it to the alternatives, or integrate other theories where possible.

There do exist some possible escapes from my conclusions, however. I identify them, not to do the work of critics for them, but to clearly delimit my claims. I can see three distinct avenues, though not how one would approach them exactly. The first is to come up with a viable form of compatibilism, which we should theoretically prefer over incompatibilism, presuming Ockham's Razor. In my eyes, this is a daunting task, for it must reconcile free will with a world determined by knowable laws, which seems to me to be inconceivable. The second, and the one that most obviously distinguishes me from Strawson, is to discredit Kant and prove conclusively that free will cannot exist outside of reason, leaving us with determinism. Unlike Strawson's claims, here, the question of morality is still tied to the metaphysical possibility of free will. Perhaps the most accessible ways to do so would be to prove that nothing could exist outside of our possible knowledge or that we can reach perfect predictive power. In this case, we could no longer assume that we have free will and would be forced to treat our reactive attitudes and intuitions of agency as illusions.

The final way is to show through neuropsychology that our decisions can be predicted. This is an empirical undertaking, as opposed to the philosophical second option. While it may leave open things outside of the bounds of our knowledge, we would have no reason to think that they applied to our decision-making.

Thankfully, I am doubtful that any of these things can ever be done, so until that point, we can go on assuming that we have what we need to assume responsibility for our rational and intentional decisions.

Conclusion

We have thus set compatibilism aside, made room for plausible alternatives, and grounded a preference of one that supports a robust account of moral responsibility. Any one of these areas is worth protracted consideration that simply is not feasible in this medium, particularly given the unresolved views of existing literature. It is a depiction, however, of the arc of the argument and grounds to consider the continued conversation in a new light.

In philosophy, we often want an answer, rather than a set of possible options. In this case though, a set of possible options is an answer. It is equally as important to rule out faulty solutions as it is to find plausible ones. Compatibilism is no more than gauze over the wounds of disheartened determinists. Morality requires a drive to do better, to make laudable decisions that determinism cannot provide. We owe it to ourselves and to the people around us to hold people to a higher standard.

When you think hard about how free will actually works, it is easy to extrapolate existing neuroscience to theoretically encompass all of our decision-making capacity. I encourage you to resist this temptation. I believe this stems, not from a grounded capacity of neuroscience to explain all of these cases, but from our tendency to rest on the theories and laws that we can know. In principle, this is not a bad inclination and has driven many incredible scientific advances. In this case, however, doing so sacrifices perhaps the most valuable part of human existence: accountability.¹⁷ As long as we are able, it seems we should prefer this possibility above all others.

¹⁷ For many moral relativists, sacrificing accountability sacrifices value altogether.

Bibliography

- Allais, Lucy. *Manifest Reality: Kant's Idealism and His Realism*. Oxford Univ Pr, 2015. Print.
- Allison, Henry E. "Transcendental Realism, Empirical Realism, and Transcendental Idealism." *Essays on Kant* (2012): 67-83. Web.
- Bok, Hilary. *Freedom and Responsibility*. Princeton, NJ: Princeton UP, 1998. Print.
- Chisholm, Roderick. "Human Freedom and the Self". On *Metaphysics*. NED - New edition. University of Minnesota Press, 1989. 5–16. Web. 30 November 2015.
- Cleve, James Van. *Problems from Kant*. New York: Oxford UP, 1999. Print.
- "Could Have Done Otherwise." *The Informational Philosopher*, Web. 30 Nov. 2015. Dodds, E. R. *The Greeks and the Irrational*. Berkeley: U of California, 1951. Print.
- Ginet, Carl. "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing." *Noûs* 30 (1996): 403. Web.
- Grant, C. K., and P. F. Strawson. "The Bounds of Sense: An Essay on Kant's Critique of Pure Reason." *The Philosophical Quarterly* 18.70 (1968): 84. Web.
- Hobbes, Thomas. *Of Liberty and Necessity*. www.theinformationalphilosopher.com. Web. 30 November 2015.
- Hume, David, and Charles William Hendel. *An Inquiry Concerning Human Understanding*. New York: Macmillan, 1955. Print.
- Hume, David. "A Treatise of Human Nature." *David Hume: A Treatise of Human Nature (Second Edition)* (1739). Web.
- Inwagen, Peter Van. *An Essay on Free Will*. Oxford: Clarendon, 1983. Print.
- James, William. "The Will To Believe." *And Other Essays in Popular Philosophy The Will to Believe* (2009): 1-31. Web.
- Kant, Immanuel. *Critique of Pure Reason*. Cambridge: Cambridge UP, 1998. Print.
- McKenna, Michael, and D. Justin Coates. "Compatibilism." *Stanford Encyclopedia of Philosophy*. Stanford University, 26 Apr. 2004. Web. 30 Nov. 2015.
- Pascal, Blaise, and W. F. Trotter. *Pensées*. London: J.M. Dent, 1931. Print.
- Pereboom, Derk. "Determinism Al Dente." *Noûs* 29.1 (1995): 21. Web.
- Rohlf, Michael. "Immanuel Kant." *Stanford University*. Stanford University, 20 May 2010. Web. 30 Nov. 2015.
- Rosefeldt, Tobias: *Dinge an sich und sekundäre Qualitäten*, in: Stolzenberg, Jürgen (Hrsg.) *Kant in der Gegenwart* (316 S., Ln., 2007, de Gruyter, Berlin). English Translation.
- Strawson, P. F. *Freedom and Resentment, and Other Essays*. London: Methuen, 1974. Print.