

4-5-2023

THE NON-APPREHENSION OF SELFHOOD AND METAPHYSICAL FREEDOM

Christopher Rinderle
Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://repository.lsu.edu/gradschool_theses



Part of the [Metaphysics Commons](#), and the [Philosophy of Mind Commons](#)

Recommended Citation

Rinderle, Christopher, "THE NON-APPREHENSION OF SELFHOOD AND METAPHYSICAL FREEDOM" (2023). *LSU Master's Theses*. 5773.
https://repository.lsu.edu/gradschool_theses/5773

This Thesis is brought to you for free and open access by the Graduate School at LSU Scholarly Repository. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Scholarly Repository. For more information, please contact gradetd@lsu.edu.

THE NON-APPREHENSION OF SELFHOOD AND METAPHYSICAL FREEDOM

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Philosophy

in

The Department of Philosophy and Religious Studies

by
Christopher Rinderle
B.A., Louisiana Tech University, 2020
May 2023

Table of Contents

| | |
|---|-----|
| Abstract | iii |
| I. The Buddhist Anātman | 1 |
| 1.1. The Ātman | 1 |
| 1.2. Eternalism and Personalism | 5 |
| 1.3. Annihilationism | 8 |
| 1.4. Non-Apprehension Theory | 9 |
| 1.5. Meditation | 12 |
| II. An Investigation of Contemporary Theories of Selfhood | 16 |
| 2.1. I*-Thoughts | 17 |
| 2.2. Phenomenal Self-Model | 18 |
| 2.3. Transparency and Opacity | 19 |
| 2.4. Critiques of the PSM | 21 |
| 2.5. Strawson's Synchronic Self | 23 |
| 2.6. Phenomenological Approach | 25 |
| 2.7. Strawson and Descartes | 26 |
| 2.8. Critiques of the Synchronic Self | 29 |
| 2.9. A Terse Reflection on Thought | 32 |
| III. Free Will and the Dark Mystery of Appearing | 34 |
| 3.1. Determinism and the Consequence Argument | 34 |
| 3.2. A Problem of Definition | 37 |
| 3.3. Indeterminacy and Randomness | 39 |
| 3.4. Reason Responsiveness Argument | 44 |
| 3.5. The Dark Mystery of Appearing | 48 |
| 3.6. Summary | 52 |
| IV. Conclusion | 55 |
| 4.1. The Basic Argument | 55 |
| 4.2. Responsibility | 57 |
| 4.3. Reward and Punishment | 58 |
| 4.4 A Positive Future | 60 |
| Bibliography | 61 |
| Vita | 63 |

Abstract

In this work I take on one of the most axiomatic assumptions humans possess. The sense of self is embedded so deeply inside each one of us that to question it seems utterly ridiculous. Fortunately, I do not outright reject the notion of selfhood. I investigate what it is and is not through multiple angles – ancient and contemporary – which leads to investigations of free will and responsibility.

In the first chapter, I discuss the Buddhist *anātman* or non-self. I argue that the historical Buddha did not endorse the idea that the self must be overcome through enlightenment – this would affirm a self – but that the Buddha endorsed the theory of non-apprehension. That is, to discuss a self's existence or non-existence is to utterly miss the point of *anātman*. Selfhood, rather, is an incoherent concept from the beginning – like pointing North as one stands on the North pole. The non-apprehension theory colors the rest of the work in contemporary views of selfhood, free will, and ultimate responsibility. I close the chapter with my personal experience with Buddhist meditation which leads me to the claim which grounds the whole thesis – *that we are not the authors of our thoughts*.

The second chapter presents an overview of two contemporary theories of the self and non-self by Thomas Metzinger and Galen Strawson. Since these views are mostly empirical and analytic, mystical language found in the non-apprehension theory of Buddhism does not fit well. Through contemporary language, I investigate their complicated arguments for and against a self. Metzinger posits a phenomenal self-model in replace of the traditional self whereas Strawson affirms a synchronic self. I critique both of their theories and reject them. In rejecting them, I discuss the ultimate grounding of my argument against egoism – that thoughts are not our own. Like the preceding chapter's section on meditation, I close the second chapter with a brief reflection of *thinking* to further develop my thesis.

From the first two chapters, which focus on my withholding of judgment on the non-concept of ātman or self, it is only natural to ponder what this says about metaphysical freedom. The theme of the third chapter was the inspiration of the whole work – that free will is an illusion. Although it is nearly impossible to write outside the purview of the binary of existence, I have slightly backed off from that theme. I argue, rather, that the illusion of free will is itself an illusion. Free will, like the self or non-self, is not a coherent concept worthy of discussion.¹ I, nevertheless, present contemporary theories involved in the free will debate such as (in)determinism, (in)compatibilism, metaphysical libertarianism, and randomness. As an anti-libertarian incompatibilist, I argue against competing theories and justifications for free will with what I call the dark mystery of appearing. This is merely a formal name for the idea that we are not the authors of thoughts or that they appear from nowhere. I find that if there is no coherent self worthy of discussion and thoughts are not our own, any notion of free will whether it is grounded in reason, effort, or desire is groundless.

I close this work with a consideration of a world where egos and metaphysical freedom are considered incoherent concepts. What does this mean for responsibility, in particular? If we are mere happenings occupying a universe of constant flux, does it make any real sense to hold “myself” responsible for “my” actions? More importantly, is it coherent to hold others responsible for choices and actions? I conclude that responsibility, too, is an incoherent concept on an ultimate level. Conventionally, however, our choices still matter due to the fact that they affect the people and world around us. Ultimately, giving up on responsibility does take away our sense of accomplishment and pride, but a *non-ethical ethic* crops up where a more forgiving and empathetic world emerges for the less fortunate.

¹ Does that make my whole thesis worthless? Perhaps. Or, maybe through its worthlessness, it is worth something.

I. The Buddhist Anātman

1.1. The Ātman

Prior to any investigations into the non-self, or the anātman in Buddhism, it is vital to discuss what the self is supposed to be. In most philosophical circles, the essence of the self is what makes me *me*. Is the self my body, my consciousness, my desires, or dispositions? Many may argue for the existence of a self through these avenues. The Buddhist tradition considers most, if not all, conceptions of a positive self and ultimately rejects them. There is an array of differing ways to justify rejecting the self (annihilationism and non-apprehensionism), but I will interpret the Buddha as not necessarily rejecting the self but questioning the very essence of selfhood as a coherent concept. I will claim that there is a middle way between eternalism (the positing of a self or ego)² and annihilationism (overcoming an ego) which I call non-apprehensionism. From this, I present why this is the most appropriate interpretation of the anātman and how there is no author of thoughts but only an illusory “taking credit” for those mental and physical experiences.

For all intents and purposes, the Buddhist interpretation of the self is the idea of an abiding, immortal substance in a person or outside, whether it is called *Ātman*, I, soul, self, or ego.³ Edward Conze breaks down the self into five potential functions which are invariably denied by him and most Buddhist scholars. They are as follows:

- I. The ‘self’ is that which *appropriates* and *owns*.
- II. The ‘self’ is that permanent factor within the concrete personality which unites and maintains its successive activities.
- III. The ‘self’ is that which *acts* and *initiates*.
- IV. The ‘self’ is the subject which *knows* and *sees*.

² I use self and ego interchangeably.

³ Walpola Rahula. “What the Buddha Taught.” Second ed., New York, NY, Grove Press, 1959, 1974. 55.

V. The 'self' is that which *distinguishes* one person from another. It separates me from others and others from me.⁴

It is through Jay Garfield, however, that the most useful interpretation of the self can be found since it unifies the five points from Conze. He explains how the self is that which is primordial and goes beyond mind and body. He gives a great analogy⁵ in *Losing Our Selves*, when he asks that you imagine yourself as another human being – *inside* another body, if you will. I always thought Peyton Manning was a great football player and have wondered what it would be like to throw the football like him. I can only imagine myself *as* Peyton Manning. That is, I do not want to be *the* physical Peyton Manning. I want to, rather, imagine myself as Chris Rinderle occupying Manning's body. From this, it shows that my body, Chris Rinderle's body, is not my self since I can imagine my self "in" someone else's body and remain "me."

In the same vein, imagine yourself in possession of another's mind. Imagine having Einstein's mind for one day, being able to understand the complicated equations of general relativity. That is, you would still be yourself but capable of the mental exercises that Einstein was capable of. You still would be *you* but with *his* mind. Thus, what it means to be your self consists of something even more primordial than your mind.⁶

From this thought experiment, it necessarily follows that if there is a self, it is something about us which follows us through time and helps constitute our identity beyond our mind and body. The self that I am now is the *same* self that I was ten minutes ago or ten years ago. My present self is also the *same* self that I will be ten minutes or ten years into the future from now. This is the common notion of a diachronic self which is discussed in detail in the next chapter of this work. But what

⁴ Edward Conze. "Buddhist Thought in India." Second printing. US, The University of Michigan Press, 1967,1970. 103-104.

⁵ Jay Garfield. "Losing Our Selves." Princeton, New Jersey, Princeton University Press, 2022. 8.

⁶ Jay Garfield. "Losing Our Selves." 9.

exactly is meant by the word ‘*same*’? This term is quite ambiguous especially when referring to the self. Sameness can be explained qualitatively and numerically. Mark Siderits explains this well when he says,

When we say ‘x and y are the same’, there are two things we might mean. We could mean that x and y are qualitatively identical, or we could mean that x and y are numerically identical. To say that x and y are qualitatively identical is to say that they share the same qualities, that they resemble one another or are alike. To say that x and y are numerically identical is to say that they are one and the same thing, that ‘x’ and ‘y’ are really just two names for one entity.⁷

When I say that my present self is the same as the self that existed ten minutes ago, I am appealing to numerical identity because they are to be taken as the same entity through time. They are not similar entities but the exact same unchanging entity. For the self now versus ten minutes ago are not distinct – this would counter the whole notion of a primordial and unchanging self. The difficulty with the idea that the self must be qualitatively unique is that it once again confuses the notion of the self with the notion of what one is like, one’s properties and characteristics. And properties may be shared between two things, whereas numerical identity may not.⁸ The self, rather, is the same now as it was then. In saying this, I am merely expressing a distinguishing temporal aspect of the same primordial self.

If this primordial self is to exist, it must be located in the individual or in the ‘I.’ For Buddhists, an individual consists of the five *skandhas* (heaps or bundles). If the ātman does exist, it should be apprehended in at least one of the five skandhas. Buddhists usually reject any notion of selfhood because these five bundles do not produce a satisfactory version of the primordial self. The skandhas are:

- Form (Rūpa): anything corporeal or physical.
- Feeling (Vedanā): sensations of pleasure, pain, or indifference.

⁷ Mark Siderits. “Buddhism as Philosophy: An Introduction.” Indianapolis, IN, Hackett Publishing Company, Inc., 2007. 33.

⁸ Mark Siderits. Buddhism as Philosophy. 34.

- Perception (Samjñā): those mental events whereby one grasps the sensible characteristics of a perceptible object.
- Volition (Samskāra): the mental forces responsible for bodily and mental activity. For example, hunger or attentiveness.
- Consciousness (Vijnāna): the awareness of physical and mental states.

The human subject is seen as a process, and all aspects of experience – the objects of experience, the sense organs, feeling, consciousness, and contents of the mind – are seen as indissolubly linked.⁹ The Buddha grants the skandhas, but through investigating each one, he fails to notice a self. He argues that the self which is posited to be primordial and unchanging fails on two accounts: 1) impermanence (Anicca) and 2) the fact that the skandhas are not under my control. As the Buddha illustrates,

For if now, O monks, the *rūpa* were the self, then this *rūpa* would not tend towards destruction, and it would be possible to say of *rūpa*, “Let my *rūpa* be this way; let not my *rūpa* be that way!” ... And that which is transitory, painful, and liable to change – is it possible to say of it: “This is mine; this am I; this is my self?”

‘Certainly not, Reverend Sir’¹⁰ replied a Monk.

For if the five skandhas are impermanent, which they most certainly are, then the self cannot be said to exist within them. For example, most modern readers would grant the skandha of consciousness as the closest to what one perceives as the self. For what am I more than my recognition of my internal workings and happenings? For the Buddhist, however, consciousness is thought of as the farthest away from what is ultimately true since it is in constant flux. One’s thoughts are forever changing from moment to moment and even this is not said of rūpa.¹¹ Consciousness is never constant and permanent.¹² Because consciousness is everchanging, it is no longer simple and eternal – it cannot be the placeholder for a self, according to the Buddha.

⁹ Brian Morris. “Anthropology of the Self.” *American Ethnologist* 25, no. 4 (1998): 746–47.
<http://www.jstor.org/stable/645864>. 59.

¹⁰ Mark Siderits. *Buddhism as Philosophy*. 38.

¹¹ Rūpa still changes overtime, but seemingly not in every moment.

¹² I will discuss this topic in greater detail when I bring up meditation and the lack of controlling one’s thoughts.

The everchanging nature of all five skandhas segues well into the more important criticism of selfhood. The fact that they are not under “my” control fits well with one of the main claims of this work involving free will. No one would argue that my body is the result of “my” making it that way. In other words, the bodily makeup of someone is not the result of their doing – they are not responsible for their body.¹³ As individuals, rather, we are thrown in the world with a certain genetic makeup, completely devoid of our approval or disapproval. We experience the world through sensation – which we also cannot control. It may be argued that I may control which sensation I want at a certain time, but the sensation, itself, is something outside of my control. The same argument can be made for perception and volition. Consciousness, again, is the problematic skandha, at least for the non-Buddhist. Most would suggest that our thoughts are authored by the locos of causation –the self – inside each one of us. But this is an assumption that I find issue with and it accompanies this work from beginning to end.

1.2. Eternalism and Personalism

If the five skandhas of “me” in no way represent what we mean by the self, then what does “me” even mean? The most popular answer to this question comes in a form of the eternalist theory known as Personalism. Nearly all schools of Buddhist thought agree on the illusory nature of the ātman. An eternalist view is the belief that as a living being, one has an essence which goes on forever.¹⁴ The personalists claim the “I” is a ‘convenient designator’ or a word that refers to something that is just a useful fiction. The *person* is that useful fiction.¹⁵ We are not selves but selfless persons. A personalist is, thereby, not a true eternalist but what I would call a minimalist. The minimalist denies

¹³ Many would blame the obese for their bodily makeup, but this is more of a discussion regarding free will and their poor dietary decisions made by those people. In this sense, I mean to say that no human being has right to claim responsibility for their genetic makeup.

¹⁴ Richard Gombrich. “Theravāda Buddhism: A Social History from Ancient Benares to Modern Colombo.” Second ed., New York, NY, Routledge & Kegan Paul Ltd, 1988, 2006. 62.

¹⁵ Mark Siderits. Buddhism as Philosophy. 49.

the primordial essence but posits a conventional entity or a substitute for a self. But in order to grasp what this means exactly, it will be necessary to explain the Buddhist distinction between conventional and ultimate reality.

This distinction is best illustrated through an analogy by the ancient Buddhist sage named Nāgasena. Jay Garfield spells out the analogy well,

Nāgasena asks the King to consider the chariot on which he rode. The King grants that he did ride a chariot, and so that the chariot he rode exists. But what, Nāgasena asks, is that chariot, really? He points out that the chariot is neither identical to its wheels, nor to its axels, nor to its poles, and so on. It cannot he argues, be identical to any of its parts, for that would be to leave some others out; to select one part as the *real* chariot would be arbitrary, as well as clearly false. So, he invites the King to conclude, it must be possible for there to be real chariots despite the fact that there is nothing to which *this* chariot refers.¹⁶

But, it may be argued, could not the chariot be identical to *all* the parts arranged in the correct manner? Nāgasena suggests not, since any one of those parts may be changed out or rearranged. As we saw with the skandhas, when something is not permanent, eternal, and is subject to change, it cannot be said to ultimately exist. This is a common mereological reductionist view. The chariot is not actually a real thing. The parts are real, but the whole that is made up of those parts is not. The whole can be reduced to the parts, it is not anything over and above the parts.¹⁷

If the chariot is not real, then what is it? Though the chariot cannot possess ultimate reality, it still is not 100% illusory. For the Buddhist, when we use such expressions in our daily life as “I”, “you”, “being”, “individual”, or “chariot”, we do not lie because there is not a self or being as such, but we speak a truth conforming to the conventional world. But the ultimate truth is that there is no “I” or “chariot” in reality.¹⁸ We can, however, still refer to “chariot” and understand what that means.

¹⁶ Jay Garfield. “Losing Our Selves.” 14.

¹⁷ Mark Siderits. Buddhism as Philosophy. 54.

¹⁸ Walpola Rahula. “What the Buddha Taught.” 55.

We use chariots for an array of useful purposes and just because the ultimate truth of the chariot may be illusory, there is still something there by which we refer. As Mark Siderits illustrates,

It is just a ‘conceptual fiction’: something not ultimately real that is nonetheless accepted as real by common sense because of our use of a convenient designator. Our common-sense ontology is full of things that we think are real, but are also wholes made of parts. The early Buddhist view is that strictly speaking none of these things is really real.¹⁹

The chariot is a thing made of parts but is no ultimately true thing – only a useful thing. The chariot might be a fiction, but it is not an utter fiction.²⁰ Buddhist personalists are not positing the view that the chariot does not exist – it most certainly does, to some degree. They are, rather, questioning the mode of the chariot’s existence. Through this discussion of conventional chariots and fictional yet useful chariots, we can grasp the analogy when applied to the ātman and the *person*.

While the self is said not to exist, personalism claims that this denial has nothing to do with the denial of personhood. For reasons given by non-Buddhists philosophers for positing the existence of a self – unity of synthesized representation, memory, relation to past and future, the centering of experience – the personalists posit a person. Some may argue that the personalists are merely using the term “person” for what most would call the self, but I think the is an uncharitable interpretation.

As noted previously, Nāgasena’s chariot analogy did not eliminate the chariot. That is, he did not proclaim its non-existence. He, rather, proclaimed that the chariot did not exist in ultimate reality – it was useful as a fiction. In the same way, the self is said to be illusory, and the recognition of that fact is ultimate reality. But the person, “I”, still exists. As Brian Morris states, “He (the Buddha) did not deny the reality of the person in the ordinary sense of the term – for it has a practical usage – nor did he deny the reality of the material world. But the essence of existence was for Buddha its

¹⁹ Mark Siderits. *Buddhism as Philosophy*. 55.

²⁰ Mark Siderits. *Buddhism as Philosophy*. 56.

impermanent and processual nature.”²¹ For the self can only be transcendent, primordial, and immaterial. And thanks to reflection on the five skandhas, it is said to be illusory. The person, on the other hand, is social, biological, and psychological. We share a name, many properties, a causal history, and a social role; and that, while not involving a self, is enough.²² The person becomes, through its lifetime, a useful conventional fiction. We can, for example, make a truth-claim about the character of Hamlet. For he is a fictional character even if he does not ultimately exist. In the same way, we can still make truth claims about the *person* which, too, does not ultimately exist. This means that my self is illusory, and my being “Chris Rinderle” is no less illusory but is, in fact, useful for convention. Under personalism, therefore, I *exist* but not on the ultimate level. This may break reality up into multiple realms which we may or may not have epistemic access to. In ultimate reality, “I” am nothing. “I” is not a coherent concept. On a conventional level, perhaps the level of reality humans naturally occupy, the concept “I” exists. Below I will endorse a different view, that the concept of “I” is not worthy of discussing as ultimately or conventionally true – both allow for the concept to be mapped onto reality, ultimate or otherwise.

1.3. Annihilationism

For context, I believe it is beneficial to briefly mention the unpopular view, among Buddhists, known as annihilationism. It is the belief that one has an essence or self which can be annihilated, typically at death or Enlightenment.²³ This view differs with personalism in that personalism suggests no self is annihilated but a conventional entity is posited. This presents notions of specific versus general negation within a certain context. As Conze points out,

One must distinguish between specific negation, stating that the self cannot be identified with a clearly defined range of items, such as the skandhas, and general

²¹ Brian Morris. “Anthropology of the Self.” 59.

²² Jay Garfield. “Losing Our Selves.” 20.

²³ Richard Gombrich. *Theravāda Buddhism*. 62.

negation, which says that ‘the self does not exist anywhere.’ The latter is a universal theoretical proposition, which is of no use in any context except that of philosophical disputation, answers no worthwhile questions, removes no misunderstanding, and does nothing to further salvation.²⁴

The personalist delves in the realm of specific negation – negating the primordial, ultimate self while granting a conventional entity. The annihilationist, on the other hand, posits a general negation by claiming that the self cannot exist at all if we come to the correct understanding of reality and see things as they truly are. The Buddhist term for this is known as Nibbāna. It means “going out” like a flame.²⁵ It is usually misunderstood that Nibbāna is the going out of the self-essence. But to assume this, it invariably follows that there must be some “thing” to go out. This is the main issue for annihilationist view and why it is an unpopular view amongst Buddhist scholars. For there to be a non-self through Enlightenment or Nirvana via Nibbāna, the self must have existed prior to. The name “annihilationist” fits well since they are not claiming the self never existed, but rather, that the self exists and must be overcome or annihilated. And through this annihilation comes enlightenment and ultimate reality. As I will show in the proceeding section, the annihilationist theory is an erroneous interpretation of the Buddha’s anātman. While personalism is a respectable theory in Buddhism, annihilationism fails to understand the true nature of the anātman and fails to meet the most basic criteria of it – that the self never existed.²⁶

1.4. Non-Apprehension Theory

Upon illustrating the eternalist and annihilationist theories, I would like to present the case why they both are incorrect views, and why the non-apprehension theory is the correct interpretation of what the Buddha meant by anātman. The eternalist view contradicts the notion of anātman from

²⁴ Edward Conze. “Buddhist Thought in India.” 130.

²⁵ Richard Gombrich. *Theravāda Buddhism*. 64.

²⁶ It could be stated more accurately. That the self is not even a coherent concept from the beginning.

the beginning. Though a minimalist interpretation, in personalism, can be charitable in conventional contexts, ultimate truth and reality fail to entertain any notion of a self or person.²⁷ The annihilationist view may sound convincing within the contexts of Enlightenment and the “going out” of Nibbāna, but this assumes a state in which the individual must find their authentic ‘self’, which is a mistaken view. Nirvana is a synonym of truth, it is a state of psychological awareness that is lived and experienced. It does not involve the annihilation of the self, because there is no self to annihilate.²⁸

Context matters when inquiring into selfhood when one posits it (as personhood), rejects it, or neither. As Conze puts it,

It may be useful to teach that there is a self, under others that there is none, under others again that there is neither a self nor a not-self. But all these statements are circumscribed by their context, and outside it they lose their significance. In the context of salvational practices an absolute ‘is’ or ‘is not’ is useless and misleading.²⁹

If we are to concern ourselves with ultimate truth and reality, which I think we should, we need to understand what the anātman means in that context. Since personalism, as a type of eternalism, and annihilationism both fail to produce a coherent explanation of the ultimate truth of selfhood, it follows that the non-apprehension theory comes closest to fitting in that context. It is called the ‘doctrine of the middle,’ and it stresses that the extremes of eternalism and annihilationism are both false doctrines.³⁰

The self as a concept does not make sense from the beginning. The Buddha did not say that a positive notion of the self is the wrong answer, but that it was the wrong question in the first place.³¹ To propose that the self does not exist is to presuppose the notion of selfhood as a coherent concept

²⁷ This was shown in the five skandhas.

²⁸ Brian Morris. “Anthropology of the Self.” 57.

²⁹ Edward Conze. “Buddhist Thought in India.” 130.

³⁰ Brian Morris. “Anthropology of the Self.” 64.

³¹ Richard Gombrich. *Theravāda Buddhism*. 65.

to begin with. This is not what I will do. I will show, rather, how the ultimate true self is the non-self. Alan Watts puts it beautifully when he says, “Anātman might be expressed in the form, ‘The true Self is non-Self,’ since any attempt to conceive the Self, believe in the Self, or seek for the Self immediately thrusts it away.”³² As soon as one realizes that any apprehension of the non-self takes a grasping of that which one is trying to reject, one can finally start to see ultimate truth.³³ Through a non-grasping, we grasp the non-grasping, which is always, in turn, a form of grasping. We must, therefore, learn to cease of thinking of things in terms of grasping and non-grasping or self and non-self.

There is a story of the Buddha which, I think, presents this sentiment well, and backs up the non-apprehension theory. A certain Parivrājaka (Wanderer) named Vacchagotta asked the Buddha whether there was an ātman or not. The story is as follows:

Vacchagotta comes to the Buddha and asks:
‘Venerable Gotama, is there an *Ātman*?’

The Buddha was silent.

‘Then Venerable Gotama, is there no *Ātman*?’

Again the Buddha was silent.

Vacchagotta gets up and goes away.³⁴

Through his silence, the Buddha gave the only true answer. All words are finally convicted of nullity and the highest wisdom lies in silence.³⁵ Not only is the self an illusion, but the illusion of a self, itself, is illusory. This non-response by the Buddha is also a practical one. For if he gave an answer, it would only have presented more problems and questions. But with his silence and wisdom, the wanderer was given the best possible opportunity to the middle path of true realization – that the self is no

³² Alan Watts. “The Way of Zen.” Second ed., New York, Vintage Books, 1957, 1985. 47.

³³ Notice that while I am using the language of differing realities, it is only because I am forced to. To talk of these ideas at all necessitates a limit. Therefore, by “ultimate reality” or “ultimate truth,” I do not necessarily mean to divide reality up – I use it to suggest that there is no truth, ultimate or otherwise, in grasping. Ultimate truth is a letting – it is not an ontological stance.

³⁴ Walpola Rahula. “What the Buddha Taught.” 62-63.

³⁵ Heinrich Dumoulin. “The Person in Buddhism: Religious and Artistic Aspects.” *Japanese Journal of Religious Studies* 11, no. 2/3 (1984): 143–67. <http://www.jstor.org/stable/30233321>. 163.

longer an appropriate entity or non-entity to ponder because it is nothing. Both the eternalist and annihilationist theories, in their pondering of the self's essence or lack thereof, proclaim the erroneous view of "I AM." We ought to see "us" as we truly are: constant mental and physical processes in flux, completely independent of our wills or desires, leading to the point of realization that nothing about us is everlasting and permanent. 'We are' nothing. "We are", under this interpretation, is incoherent. Upon this realization, the true "self" becomes nothing (anātman).

1.5. Meditation

I would like to discuss the reasons for the sentiment I hold regarding the anātman in Buddhism through my personal experiences and realizations in meditation and tying them to our lacking authorship of thought and action. I would also like to preface this with the fact that I am in no way an expert in meditation practices. I have, however, decided that it benefits my mental and physical wellbeing. It just so happens that meditation also brought me life altering thoughts about being and freedom.

Meditation, in the most general sense, is an effective practice to improve mental health, equilibrium, and tranquility.³⁶ It is not about escaping from society or individualism in order to produce some mysterious trance. Meditation is concerned with a mental cleansing of the poisons of life which ultimately lead to a life of suffering. The word itself is a very poor substitute for the original term *bhāvanā*, which means 'culture' or 'development', or mental culture and development.³⁷

Bhāvanā can be separated into two forms. The first being known as *samatha* or *samādhi*, or mental concentration. The meditator is to home in on one single aspect of consciousness for the purpose of getting lost in a kind of mystic state. One could be said to 'lose oneself' in this mystic form through meditation. But as we have already stated, this is the wrong approach, in ultimate Buddhist

³⁶ Walpola Rahula. "What the Buddha Taught." 67.

³⁷ Walpola Rahula. "What the Buddha Taught." 68.

contexts. This mental concentration, however, is not wholly useless because it does produce some sense of happiness and a peaceful outlook on life. In my very little experience with mental concentration of this form, it reduces stress levels and allows me to gain a greater appreciation for living in each moment. And for this, I am grateful for it.

The second and more important form of meditation for my purpose is known as *vipassana*, or insight. This is essentially Buddhist meditation and Buddhist mental culture. It is an analytic method based on mindfulness, awareness, vigilance, and observation.³⁸ When you first start the practice of mindfulness (*smṛ-ti*), it helps to concentrate on one aspect of your experience³⁹ such as the breath or the pressure created by gravity between you and the Earth. You breathe in and out all day and night, but you are never mindful of it, you never for a second concentrate your mind on it.⁴⁰ Through mindfulness, you should begin to take notice of these banal happenings of your conscious experience. As you practice, you will invariably notice how easily your thoughts wander – losing your concentration and awareness.⁴¹ We begin to tell ourselves stories, through endless thought, about what we ate yesterday or, perhaps, the most embarrassing event that happened to you in high school, for example. This is not a problem, so long as you notice the thought(s) as a thought of consciousness and “come back” to the original concentration on the breath. The point is not to focus solely on the breath – the breath is merely a ‘home base’ to come back to upon noticing the wondering. The ultimate goal of mindfulness is to be aware of every conscious phenomenal and mental experience you have in the very moment it *comes* to you.

Through gaining enough experience with mindfulness in the form of meditation, you will begin to become aware and mindful of whatever you do, physically or verbally, during the daily routine

³⁸ Walpola Rahula. “What the Buddha Taught.” 68-69.

³⁹ In a similar way the practice of samatha, above, describes meditation.

⁴⁰ Walpola Rahula. “What the Buddha Taught.” 70.

⁴¹ A similar phenomenon comes when one is driving down the road and coming to the realization that you have been lost in some mental conversation with your inner interlocutor while paying nearly no attention to driving.

of work in your life.⁴² We can train ourselves to become so aware of each moment regardless of the present state of our existence, either meditating or not. The practice of mindfulness goes beyond the meditative state of mental concentration while sitting straight with your eyes closed. In our everyday lives, we can be mindful. Anytime you feel anger, for example, and identify your ‘self’ with that anger, usually negative results ensue. Your stress level increases, or you may lash out at someone for being the ostensible reason for that anger. If you can become mindful, on the other hand, and notice that anger is merely a retelling of a disapproving story, unauthored by you, in your mind – which are thoughts that merely appear to you in consciousness – the anger soon evaporates. Mindfulness is a total clarity and presence of mind, actively passive, wherein events come and go like reflections in a mirror.⁴³ And this is the ultimate point of mindfulness meditation – that our thoughts are not our own.

One may ask, however, how this has anything to do with the notion of the *anātman*? If there is no concept of the self to be discussed, what is ‘creating’ these thoughts? Nothing is creating thoughts; they merely appear to us in consciousness. In order to think a thought, with some internal locus of causation initiating the thinking of the thought, one would have to think *to think* the thought, leading to a regress. Upon partaking in mindful mediation, one may notice that we are *experiencers*, whatever “we” are, of thoughts and not the *authors* of them. By being aware of one’s own physique, feeling, states of mind and thoughts, the Buddhist will cease to identify with them as their ‘self’, to introject a sense of ego into what are transient phenomena, constantly coming into being and passing away.⁴⁴ We are, therefore, not an active locus of causation with freely caused thoughts. We are nothingness that possesses some illusory ability to take credit for the thoughts that merely happen to us. We do not first experience a self, and then experience experiences and wonder whether they are

⁴² Walpola Rahula. “What the Buddha Taught.” 71.

⁴³ Alan Watts. “The Way of Zen.” 53.

⁴⁴ Richard Gombrich. *Theravāda Buddhism*. 65.

associated. We first *have experiences* and use them to construct a representation of a self.⁴⁵ Through meditation and mindfulness, we must see that if our thoughts do not originate with us, we can no longer be said to possess some primordial “I” or ego. This leads to a lack of causal power, since there is no locus of causation, making freedom of the will a problematic conception as well. This is the ultimate goal of this work⁴⁶ – to show that 1) we have no notion of ‘self’ worth discussing, 2) there is a lack of original self-causal thought processes, and from this, 3) any coherent notion of freewill is nowhere to be found.

⁴⁵ Jay Garfield. “Engaging Buddhism: Why It Matters to Philosophy.” New York, NY, Oxford University Press, 2015. 118.

⁴⁶ The thesis as a whole.

II. An Investigation of Contemporary Theories of Selfhood

In this section, I will discuss many of the most notable and well defended contemporary interpretations of the self and non-self. Through the analysis of Thomas Metzinger's "Being No One: The Self-Model Theory of Subjectivity" and Galen Strawson's "Selves," we will better grasp contemporary notions of selfhood and phenomenal self-models which are not mentioned in Buddhist contexts. I begin with Metzinger's theory of the PSM or the phenomenal self-model. While he denies the existence of a traditional self, he posits an empirical and evolutionary explanation for the first-person perspective and experience. While I tend to agree with Metzinger on his conclusion that selves are mere fictions or things not worth discussing, I cannot help but disagree with him due to the problem of subjectivity which I present with a twin earth example. Strawson's view, which affirms selfhood, albeit synchronically, is also justified through the phenomenological experience in each moment. Selves exist synchronically, which is to say that they are born and die in each moment of experience giving rise to another self in the next moment. I also draw a parallel with Strawson's material synchronic self and Descartes' "thinking thing." While this thesis does not ultimately conclude that selves exist (or not-exist), Strawson's argument is the most convincing. His argument is faulty, however, which causes me to abandon it. I conclude with a brief meditation on thinking serving as a summation of my critiques of Metzinger and Strawson.

I will outline contemporary analytic theories of selfhood which include arguments that conclude in differing ways. While these theories are both empirical and mostly materialist theories, they each come to glaringly different conclusions. Metzinger posits a self-denying theory, or so it seems, while Strawson argues for a self-affirming theory. I found that this manner of investigating the contemporary self was most profitable, but not quite profitable enough to adopt either one.

2.1. I*-Thoughts

As a naturalist, Thomas Metzinger attempts to explain the human, first-person experience through the third-person. There is an important distinction between possessing a first-person perspective and conceiving oneself as the *experiencer* of this first-person perspective. That is, there is a clear difference between mere sentience and self-consciousness. While the latter is a subset of the former, the reverse is not the case. Humans possess the latter. We can experience ourselves as “I” in each moment, as many animals also can, but there exists a deeper level whereby we possess the concept of “I” doing the thinking as the thinker, or experiencing an experience as the experiencer. As Metzinger says “It is not only necessary to have thoughts that can be expressed using ‘I.’ What is necessary is the possession of a concept of oneself as the *thinker* of these thoughts, as the *owner* of a subjective point of view.”⁴⁷ For example, when a dog desires a treat, it knows the treat is something to be desired, otherwise it would not want the treat. The dog does not necessarily know, through reflection or in real time, that it is the entity desiring of the treat. Humans, as mentioned above, do possess this ability. One could even say that we are forced to think this way. For how could I not think of myself as the thinker of these thoughts or the experiencer of my experiences?⁴⁸ I know I am writing this sentence, but I also cannot help but realize that I feel as though I am the entity which possesses the knowing that I am writing this sentence. Self-consciousness or a first-person perspective is a kind of sentience on a meta-level. Without the bare minimum of the first-person perspective, it is difficult to imagine an animal surviving in evolutionary terms. But self-consciousness allows one to engage in activities beyond the mere biological. I could not reflect on higher ideas or write this sentence, for example, if I did not possess some notion of myself as the thinker of the thoughts that appear to me. These thoughts are known as I*- thoughts – we can see ourselves as ourselves. While this explanation

⁴⁷ Thomas Metzinger. *Being No One: The Self-Model Theory of Subjectivity*. Cambridge, MA: Bradford Books. 396.

⁴⁸ While I do end up denying this on an ultimate level, I cannot help but recognize that this is how the human being is hardwired, even if it creates an illusion.

of I*-thoughts seems rather intuitive since we all experience them naturally, Metzinger denies this.⁴⁹ For him, I am not a self thinking of myself. I am, rather, a phenomenal self-model, an *it*⁵⁰, not an “I*.”

2.2. Phenomenal Self-Model

We are not selves in the traditional sense, but we are information-processing systems that activate a world model and a self-model. The self-model that is activated is the phenomenal self-model (PSM). When Metzinger rejects selves in the “traditional” sense, he is referring to the Cartesian immaterial/dualist cogito – a substance that possess the property of thinking. This traditional self could also consist of a primordial self which goes beyond the body and mind. “I” am something which is not my body, but I am also not purely my mind since I can imagine myself existing with the mind of another. The primordial self is an entity which transcends mind and matter.⁵¹ But if there is no self, in this sense, what are we? Metzinger attempts to answer this question through his theory of the PSM.

As he says, “The content of the PSM is the content of the conscious self: your current bodily sensations, your present emotional situation, plus all the contents of your phenomenally experienced cognitive processing. One could say you are the content of your PSM.”⁵² The self-model is created by the representational system that is always currently activated within itself.⁵³ Through the content given to the system, a world and self-model emerge. Metzinger claims the PSM is grounded in the phenomenology of our experience. Experience supervenes on brain states and are characterized by how it feels or “what it’s like” to have them.⁵⁴ My conscious self, which was previously used to describe I*-thoughts (my thinking of myself as the thinker), is also my phenomenal experience which occurs due to the existence of the self-model which conceptualizes the content of my experience. This model,

⁴⁹ Metzinger does not deny the experience described here. No one can deny the experience, but he thinks this experience is illusory.

⁵⁰ This *it* will turn out not to be a thing, but a process.

⁵¹ Jay Garfield. “Losing Our Selves.” Princeton, New Jersey, Princeton University Press, 2022.

⁵² Thomas Metzinger. *Being No One: The Self-Model Theory of Subjectivity*. 299.

⁵³ Thomas Metzinger. *Being No One: The Self-Model Theory of Subjectivity*. 302.

⁵⁴ Lynne Rudder Baker. “DOES NATURALISM REST ON A MISTAKE?” *American Philosophical Quarterly* 48, no. 2 (2011): 161–73. <http://www.jstor.org/stable/23025085>. 167.

however, is not epistemically justified according to Metzinger. In the same way that I own a self-model during dreams or hallucinations, where none of the representations of these models are said to exist, the representation of the PSM, too, cannot be known to be real.⁵⁵ A brain in a vat could generate the conscious experience of thinking [*I am certain that I* exist*] as easily as I can.⁵⁶ The self-model merely allows for the activation of a representation of a world and a self in that world.

2.3. Transparency and Opacity

There are intrinsic aspects of this self-model which are described as *transparent* and *opaque*. For I*-thoughts to ever occur, an opaque representation must be integrated into a preexisting transparent self-model. As Metzinger spells out,

The phenomenology of transparency and opacity is this: transparent experience is the experience of not only knowing but also knowing that you know while you know; opaque experience is the experience of knowing while also (nonconceptually, attentionally) knowing that you may be wrong.⁵⁷

The standard way of defining transparency would be to say that only the content properties of the phenomenal *representata* are introspectively available to the system, and not the vehicle properties.⁵⁸ In other words, through transparency, the content is available to the system – which is not epistemically justified – while the mechanisms involved in interpreting the content are completely invisible to the system itself. This experience takes place in a *medium*, making it inaccessible to conscious experience.⁵⁹ For example, take the experience of looking through a perfectly clean window where there is a bird. The content includes the bird and everything around the bird. In the moment of experiencing the bird, we do not think of the clean window through which the bird is being viewed. The window, in this analogy, is the transparent representational aspect of the PSM. We are unable to see it, but without

⁵⁵ Lynne Rudder Baker. "DOES NATURALISM REST ON A MISTAKE?" 168.

⁵⁶ Thomas Metzinger. *Being No One: The Self-Model Theory of Subjectivity*. 404.

⁵⁷ Thomas Metzinger. *Being No One: The Self-Model Theory of Subjectivity*. 171.

⁵⁸ Thomas Metzinger. "Phenomenal transparency and cognitive self-reference." *Phenomenology and the Cognitive Sciences* 2 (2003): 353-393. 358.

⁵⁹ Thomas Metzinger. "Phenomenal transparency and cognitive self-reference." 359.

it there would be no means for the content to make any sense. In the same way, without the window, the bird would not be visible to us, the vehicle.⁶⁰ The transparent representation is that which must exist, but one that cannot be apprehended by the very system which possesses it.

There is another accessible aspect of our representation which Metzinger calls opacity. It is through this opaque representation that I*-thoughts emerge. Opaqueness means that we can recognize or take credit for the representation showing itself. Metzinger mentions,

Opacity appears precisely when darkness is made explicit – at the moment we consciously represent *that* something actually is a representation, not by propositional knowledge or a conscious thought, but first by our attention being caught by the fact that what is currently known is known through a medium.⁶¹

The very fact that we are able to consciously acknowledge the possession of the PSM, which constructs the models, is the ultimate example of opaque phenomenology. We can *know* that we know these representations and the ways in which they are known. It is as if the window, to use the earlier analogy, was to become dirty or foggy – I would be forced to acknowledge the vessel through which I view the bird, the window.⁶²

This self-model theory with two major components of transparency and opacity allows Metzinger to posit that we are processing models that can activate an experience which “we” take to be our own. The opaqueness of the self-model maps onto the preexisting transparent aspect, giving rise to the illusion that I am the thinker of the thoughts that appear to me. As he says, “Cognitive self-reference is a process of phenomenally modeling certain aspects of the content of a preexisting transparent self-model, which in turn can be interpreted as the capacity of conceiving oneself as oneself.”⁶³ The first-person perspective is an illusion making “us” mere phenomenological

⁶⁰ Phenomenal transparency is a special kind of darkness. This darkness is said to be beneficial to the system because it creates a world and self-model that is taken to be the realistic versions of ultimate reality, even when this is not epistemically justified. For a biological being focused only on survival, however, justification is of little concern.

⁶¹ Thomas Metzinger. *Being No One: The Self-Model Theory of Subjectivity*. 170.

⁶² Consciously experienced thoughts, lucid dreams, and certain types of hallucinations where the subject is aware of the hallucination as such – these are examples of opaque phenomenal experience.

⁶³ Thomas Metzinger. *Being No One: The Self-Model Theory of Subjectivity*. 405.

information-processing systems in order to better grapple with the difficulties of survival in a material world where evolution randomly determines who is to survive. My thinking that I am the entity writing this sentence is an I*-thought referring to a representation in my opaque self-model representing itself as the thinker of this thought which, in turn, must be integrated and mapped onto my preexisting transparent self-model.⁶⁴

2.4. Critiques of the PSM

With a theory of selfhood as complex as this, there comes many potential problems. Primarily, if we are only content processing models, how do we have the experience of subjectivity or mine-ness? In other words, why would a model think of itself as something that is the opposite of its actual nature? Metzinger attempts to answer this question by suggesting that since the model possesses the capacity to self-reference (mapping opacity onto transparency), the process is forced to view itself as an “I.” The model generates experiences – one of which is the notion that I am the one generating the experience. But what makes this PSM “me” and your PSM “you”? Metzinger explains this through the fact that my PSM is distinct from yours because we are not experiencing the same phenomenological qualia. This can be pushed even further with a twin earth example. Imagine you, the reader, have a twin who lives on twin earth where you each possess the exact same PSM. You both have the same genetics, brain states, character traits, memories, desires, and qualia. The only difference lies in that you both do not occupy the same space – that is, there exist two separate, yet identical, PSMs. In each moment, you and your twin have the exact same phenomenal experiences. There is a problem though: one of these PSMs is *yours* and one is *not*. Thus, while the phenomenal content is undoubtedly the same for each PSM, you are completely unaware of one of them. This shows that one physically embodied PSM is no more justified in being explained as “you” as anyone else and if one were to attempt to do so, only an answer from arbitrariness would result. That is, there is no

⁶⁴ Lynne Rudder Baker. “DOES NATURALISM REST ON A MISTAKE?” 167.

explanation for you possessing your PSM while your twin possesses their PSM. It could have easily been the opposite.

Furthermore, there is a clear paradox resulting from Metzinger's theory. Since the self-model produces an illusion, the experience that I undergo when I feel like it is "I" who thinks is just a *confused* position of mine, according to Metzinger. But *who* is being confused? If it is not truly me, then how can there be anything which is experiencing the confusion? Since Metzinger is a materialist, or at least attempts to explain selfhood in the realm of materialism, the brain ultimately produces the self-model meaning that "I" cannot be confused about my non-existence. My brain matter, rather, confuses itself and I resonate with this confusion and create an illusory model that aids me in my survival. The process which occurs in my brain confuses itself, but there is no subject to be confused. For Metzinger, the PSM resulting from the innerworkings of the brain – whether in a skull, vat, or some other medium – is confused, not a self.

I am close to agreeing with this idea that we are merely processes happening in each moment, if I were to grant a materialist explanation found in the PSM. I cannot help but criticize Metzinger's labeling the PSM as some repudiation of selfhood. He calls it a process but he refuses to acknowledge its essence as a self. He admits the PSM can be substituted for the traditional self, but then claims that selves do not exist, only processing models do.⁶⁵ What he means to say is that there is a self, but only a special kind of self – one that is found in his theory of the PSM. The problem with a physical explanation for a PSM is the locality of it. It cannot exist beyond the brain – meaning there exists a specific makeup of atoms in a certain part of the brain, or a certain makeup of the whole brain, which curates this version of the self (PSM). In short, a brain state makes up this self in each moment. Any purely material explanation for the existence of a self fails since there is the problem of the infinite

⁶⁵ Thomas Metzinger. *Being No One: The Self-Model Theory of Subjectivity*. 626.

substitution of matter. If the brain or a part of the brain locates the self, we cannot add nor remove a single atom to this recipe. But mental flux occurs in each moment.

Thought and consciousness are in constant flux and are completely out of the control of any sense of inner causation of thought. This is laid out well in an analogy by Douglas Hofstadter when he says,

Most of the time, any given symbol in our brain is dormant, like a book sitting inertly in the remote stacks of a huge library. Every so often, some event will trigger the retrieval of this book from the stacks, and it will be opened and its pages will come alive for some reader. In an analogous way, inside a human brain, perceived external events are continually triggering the highly selective retrieval of symbols in dormancy, and causing them to come alive in all sorts of unanticipated, unprecedented configurations. This dance of symbols in the brain is what consciousness is. (It is also what thinking is.)⁶⁶

We are ever-changing beings living in a constantly evolving world leaving no room for an entity – a self or PSM – capable of being incessant, constant and, more importantly, capable of determining the goings-on in the mind.⁶⁷ This is the gravamen of my argument which should be familiar by now.

2.5. Strawson's Synchronic Self

Galen Strawson argues in favor of a self. Although most self-affirmers are unsatisfied with his argument, since he argues for the unpopular synchronic self, I find it to be slightly convincing for a self-affirming theory.⁶⁸ The strength of his argument comes through the domain of the undeniable phenomenological experiences we all find ourselves in from moment to moment.

He begins by defining self-experience as the experience informed by a certain complex experience-determining mental element: the idea or feeling of the self. It can exist whether or not

⁶⁶ Douglas R. Hofstadter. 2008. *I Am a Strange Loop*. London, England: Basic Books. 276.

⁶⁷ One could argue, as many do, that the self will never be found materially which allows them to posit an immaterial self. My critique of the coming and passing of conscious thought as out of “my” control applies to a material or immaterial sources of selfhood.

⁶⁸ While it may be a convincing theory, it still fails to completely convince me that selfhood is a coherent concept.

selves do, just as pink-elephant-experience may exist whether or not pink elephants do.⁶⁹ Note that because Strawson is a materialist, when he mentions self-experience, he does not mean the experience of the physical human being as a whole. For Strawson, materialism is the view that ‘everything is physical,’ making him willing to endorse the view that experience may just be neurons firing.⁷⁰ Due to our ignorance in the fields of mind science, he makes the common distinction between the human being as a whole and the mental being – the inner mental presence that nearly never ceases. It is the latter that gives us the feeling that we are a subject, a self.⁷¹ In common parlance, “I” can be used to refer to both the physical human being or the inner mental presence. While some identify with their bodies, others frequently use language suggesting they identify with something beyond the physical human being as a whole. We move naturally between conceiving of ourselves primarily as a human being and as some sort of inner subject. Sometimes we mean to refer to the one, sometimes to the other, sometimes our semantic intention hovers between both, sometimes it embraces both.⁷² The most common use, which groups us all together as believers in the inner subject, is the phrase “my body.” If I am my body, then the phrase makes little sense – it becomes redundant. If my self is an entity distinct from my body as a whole, then this phrase makes perfect sense. I possess a body as opposed to my body being “me.”

In short, Strawson attempts to explain the inner presence of our minds as something no different from our eyes, our heart, or any other part of us when considering the human being as a whole. He calls our inner subject of experience *thin* subjects. As he says,

Are we thin subjects? In one respect, of course, we are thick subjects, human beings considered as a whole. In this respect we are, in being subjects, things that can yawn and scratch. In another respect, though, we are, in being subjects of experience, no

⁶⁹ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. London, England: Oxford University Press. 2.

⁷⁰ E. J. Lowe. “Selves: An Essay in Revisionary Metaphysics.” *Analysis* 71, no. 3 (2011): 587–92. <http://www.jstor.org/stable/41237366>. 588.

⁷¹ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 32.

⁷² Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 337.

more whole human beings than hands or hearts: we are – literally – inner things, thin subjects no more things that can yawn or scratch than eyebrows or thoughts.⁷³

The human being as a whole is purely material – the human body – but the inner mind, phenomenologically, is distinct and goes beyond the human being as a whole. This distinction is grounded in our phenomenal experience from each moment in time. While a distinction like this may seem to strongly suggest that Strawson is some kind of dualist, he is persistent in maintaining materialism, as I will show below. But first, I will describe his phenomenological approach to self-experience.

2.6. Phenomenological Approach

For Strawson to posit an inner subject or self, he thinks it must be rooted in the undeniable. This undeniable aspect of ourselves comes in the phenomenology of what it is like to experience each moment in time. For him, phenomenology is the general study of the character of experience, the experiential ‘what-it’s likeness.’⁷⁴ This experience, while also a sensory kind, is more importantly cognitive. We constantly experience physical sensations of the body, but Strawson focuses on the mental goings-on – which is true self-experience. This phenomenal, cognitive self-experience becomes the distinctive entity separate from the human being as a whole.

Our experience is necessarily a single thing, in contrast with Metzinger’s PSM. Strawson says,

The self is not figured in ordinary human self-experience as a mere state or property, or as a mere event or series of events, or, given the ordinary understanding of the word ‘process’, as some sort of mere process. To that extent, I propose, there is nothing else for it to be figured as, other than as a thing or entity of some sort. It isn’t apprehended as being a thing in the way a stone or a chair is, but it is none the less apprehended as a thing of some kind. For the self is, centrally, experienced as something that can undergo experiences, have thoughts, decide things, perform actions, and, most simply, be in some experiential state.⁷⁵

⁷³ Galen Strawson. “What Is the Relation between an Experience, the Subject of the Experience, and the Content of the Experience?” *Philosophical Issues* 13 (2003): 279–315. <http://www.jstor.org/stable/3050535>. 285.

⁷⁴ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 38.

⁷⁵ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 65.

We are not processes, but single things or entities which occur in each moment of our cognitive experience.

It is easy to confuse the language of “thing” or “entity” as some part of the human being as a whole. But this would be a mistaken view of what Strawson is saying. Strawson views the human being as a physical, material entity and the subject of experience (self) as a mental entity when he says,

A human being considered as a whole is ordinarily figured as something that has mental properties or mental being, but it isn't ordinarily figured as 'something mental,' and is certainly not ordinarily figured as a 'mental thing.' A self, by contrast, is ordinarily figured not only as something that has mental properties, or mental being, but also as 'something mental,' in some sense, and indeed as a 'mental thing.'⁷⁶

Does this imply that Strawson is, in fact, contradicting himself when he also claims to adhere to materialism? He thinks not. He says, “If you ask me whether I think the self is merely mental and immaterial, metaphysically speaking, I'll reply, 'Not at all. I take the self to be wholly a brain phenomenon.'”⁷⁷ He suggests that we are in the infancy of understanding the physical brain and ought not conclude, based upon an epistemic gap, that the mind is purely immaterial. While the phenomenological experience of a self may feel as though it is immaterial, this may be an utter illusion, Strawson grants. But the illusion of this experience does not justify concluding that said illusion is irrelevant. If it is illusory, it is a useful one.

2.7. Strawson and Descartes

Reflecting upon this synchronic, phenomenological, and cognitive self posited by Strawson, I could not help but think of the similarities it has with Descartes immaterial cogito. The language used by Strawson, such as the ‘single and mental entity,’ reads like the language Descartes used in the second meditation in his *Meditations on First Philosophy*.

⁷⁶ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 68.

⁷⁷ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 69.

In part seven of *Selfes*, Strawson translates many of the concepts outlined above into more convenient terms, which caused me to refer to Descartes. He starts by explaining the two conceptions of the subject, and then creates a new third conception which he ultimately endorses. This third conception is merely a materialist version of Descartes' argument that the self is a thinking thing. The first two conceptions are:

[A] the *thick* whole-creature conception; and

[B] the *traditional inner* conception to which a subject of experience is an inner presence of some sort.⁷⁸

Both of these conceptions allow for a continued existence of the subject without any experience – say complete unconsciousness. For this reason, Strawson introduces his *thin* conception of the subject.

[C] The *thin* conception according to which a subject of experience does not and cannot exist at any given time unless it is having experience at that time.

This *thin* subject is the synchronic subject outlined above where in each moment of experience, a new self is born and dies in each subsequent moment of experience.

Since Descartes was a dualist, [B] would still, technically, apply to his cogito. But Strawson's 'experience' in each moment of selfhood – a thing having experiences in each moment – reads very similar to Descartes' 'thinking thing.' In his second meditation, he says,

What about thinking? Here I make my discovery: thought exists; it alone cannot be separate from me. I am; I exist – this is certain. I am therefore precisely nothing but a thinking thing; that is, a mind, or intellect, or understanding, or reason. Yet I am a true thing and am truly existing; but what kind of thing? I have already said: a thinking thing. ... A thing that thinks. What is that? A thing that doubts, understands, affirms, denies, wills, refuses, and that also imagines and senses.⁷⁹

⁷⁸ Galen Strawson. "What Is the Relation between an Experience, the Subject of the Experience, and the Content of the Experience?" 283.

⁷⁹ René Descartes. 1993. *Meditations on First Philosophy*. Translated by Donald A. Cress. 3rd ed. Cambridge, MA: Hackett publishing. 19-20.

In other words, the fact that Descartes finds himself in a state of constant thought, regardless of the nature of that thought, affirms the existence of his self since there must be a thing doing the thinking. For Descartes, “thought” is used in the same manner that Strawson uses “experience.” And since, for Strawson, the self is a thing that necessarily experiences cognitively, he can only mean exactly what Descartes means. We are thinking or cognitively experiencing things that can only happen to an entity for which the experience “is for.” As Strawson says, “Experience necessarily involves experiencing (experience is just experiencing), and the existence of a thin subject is guaranteed by the fact that there is experiencing.”⁸⁰ For there to be an experience or a thought for Descartes and Strawson, an experiencer and thinker are necessary.

The only⁸¹ discernable difference between the two is their attitudes towards materialism. Since Descartes holds to substance dualism, the self must be that substance which goes beyond the physical body – it is mental, not physical. For Strawson, the thin subject is *literally* an inner thing. That is, it is located in the brain of the subject. He alters the third conception, [C], of the subject to:

[C]* the thin conception according to which a subject of experience is an inner thing of some sort that does not and cannot exist at any given time unless it is having experience at that time.⁸²

As he says directly below this conception, “As a materialist, I take the subject to be inner in a robust spatial sense because the physical goings-on that wholly constitute its existence consist entirely of parts of a brain in a certain state of activation.”⁸³ He believes the self to be located in the brain but is not divorced from the phenomenological experience that occurs to us in each moment in time. The self

⁸⁰ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 325.

⁸¹ Not to mention the difference between their conclusions regarding synchronicity and diachronicity.

⁸² Galen Strawson. “What Is the Relation between an Experience, the Subject of the Experience, and the Content of the Experience?” 284.

⁸³ Galen Strawson. “What Is the Relation between an Experience, the Subject of the Experience, and the Content of the Experience?” 284.

exists for Descartes and Strawson, albeit the former as persisting and the latter as synchronic, for similar reasons yet different assumptions.

2.8. Critiques of the Synchronic Self

The self argued for by Strawson is a clear example of synchronic identity, directly opposing diachronic or persistent⁸⁴ identity that is more closely associated with Descartes' cogito. While Descartes and Strawson argue for their respected "selves" in a similar manner (the cognitive experience in each moment), Strawson breaks from Descartes by endorsing a special, and unpopular, version of identity – synchronic identity. For context and clarification, a single diachronic self is figured as something unified when considering it as something that lasts longer than the lived present of experience.⁸⁵ Strawson rejects this identity and endorses the synchronic, which figures the self as something unified when considering it as something existing in the lived present of experience.⁸⁶ Since the phenomenal experience cannot be denied, this notion of synchronic selfhood seems to make sense.

It is interesting to note that Strawson does not entirely give up on the diachronic self though. As he says, "The self is figured as diachronically single (persisting) specifically in so far as it is something that is mentally propertied."⁸⁷ To pick out a single property of an entity, however, to ground a persisting self through time is cause for concern. Although we are mentally propertied, I fail to see how anyone could argue that it is the same "me" that continues to possess the property itself. For it certainly does not follow that the property of *flux* is grounds for suggesting an entity is the same entity through time. Similarly, when mental property is granted, the entity possessing the property can change through time. The question is this: is it necessary that an entity remain the same because it possesses

⁸⁴ I will use persisting and diachronic interchangeably. They will both represent an idea of a self that continues through the linear time of a subjective existence.

⁸⁵ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 75.

⁸⁶ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 75.

⁸⁷ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 77.

at least one discernable property? The property certainly remains the same, but the entity can and, perhaps, must change through time.

Strawson abandons the diachronic self, however, for the seemingly more promising synchronic self. He states,

Is persisting a necessary component of any genuine self-experience? Must one have the persistence belief, belief that one (I*) is something persisting, in order to have self-experience? Must one have persistence experience, experience of oneself* as something persisting, in order to have self-experience? I think not. It seems that self-experience can be vivid and complete at any given time even if it has to do only with the lived present of experience and never involves a sense or conception of the self as something that has any diachronic singleness or continuity.⁸⁸

The phenomenal experience in each moment cannot be denied even by some of the strongest of skeptics, but where can the subject be located in this “moment” of experience? There is a problem of regression with any argument of momentary experience, especially when it is grounded phenomenologically. Any slice(s) of linear time can be forever divided into or added up to constitute a “moment.” If my self is born through one cognitive experience but then dies in the next, when exactly does this “next” moment begin? If it is subjective – or up to the subject to determine – then this undeniable subjectivity is a possible argument even for a diachronic self. Some individuals experience their lives as a continuing moment, such as a movie where they are the main character. Others see their lives from moment to moment – not identifying with their past or future self.⁸⁹ Because these two phenomenal experiences of living are equally valid, due to their undeniability, I fail to see where any notion of selfhood, either synchronic or otherwise, can be reconciled. Their equal claim to validity may be an argument for both, but since synchronic and persisting selves are mutually exclusive, it seems to be an implausibility. In short, due to the strong impossibility of locating a self in the correctly divided slice of time, the self which is not persistent seems unconvincing. Even if we

⁸⁸ Galen Strawson. 2011. Selves: An Essay in Revisionary Metaphysics. 199-200.

⁸⁹ These two notions are called endurantist and impermanentist views of living a life, respectively. For more on this, see Strawson (2009) 221.

could all agree on what this “moment” is, our differing psychologies (endurantist and impermanentist individuals) would never reconcile the different phenomenal experiences of selfhood, whether they are momentary selves or persisting selves. The point is not that a synchronic self is non-existent, but that it is an implausible explanation for feeling like a thinking thing in each moment. Moments can feel infinitely short or immensely long. For example, when one ponders carefully in each moment, it is quite easy to divide time into miniscule slices. There are other times, where one gets lost in some action, such as driving a car, where a moment can make up a quite large sum of time. For the reason of the imprecision of momentary slicing, and our differing mental make-ups, a synchronic self fails.

A stronger critique has to do with Strawson’s “inner” thing of the thin subject. He says, “As a materialist, I take the thin subject to be inner in a robust spatial sense because the physical goings on that wholly constitute its existence consist entirely of parts of a brain in a certain state of activation.”⁹⁰ He ultimately reduces the self to a physical brain state or some part of the brain where selfhood is located. He does this to embrace materialism and to avoid a mind-body problem, but it creates new issues that he does not address. Primarily, if this part of the brain is constantly impermanent, then this self would manifest as a materialist cogito. If my self, in my brain, is unchanging, there would need to be an explanation for why my self has changes in emotion, attitude, and worldview. But the synchronic self as a part of the physical brain does not account for these changes. If the self is in constant flux, aligning with his theory of phenomenal experience in each moment, how can this self be physically and stagnantly located in the brain? Assuming our brains are in flux, in a purely material universe, there is no constancy allowing Strawson to claim the self’s locality in the brain. He can still conclude that the self is *somewhere* in the brain but could never jump to concluding exactly where it is. He is, therefore, stuck with a dilemma. On the one hand, he can grant the physical location of the self in the

⁹⁰ Galen Strawson. “What Is the Relation between an Experience, the Subject of the Experience, and the Content of the Experience?” 284.

brain, while, on the other hand, rejecting the notion that phenomenal cognition is ever changing – both of which he tries to affirm.

Most importantly for my purposes, I find the notion of a necessary experiencer to be a false one, or at least illusory. For even I cannot deny the phenomenal experience I am currently living out this very moment. But does that necessarily amount to “me” having the experience? Could not an experience – phenomenal or otherwise – occur to an entity, which I happen to identify with, without it being “me”? I would answer positively.

It is most obviously the case that I do not actively engage in the process of sense data being given to me. Experiences occur through sensations of light, sound, or vibration. There is no necessity of an ego or witness for this process to occur exactly as it does every moment of my life. But since Strawson makes the clarifying point that by experience he means a cognitive type of experience, perhaps there is room for a self in each moment of cognition – since thought and reasoning are at the core of this issue. For without *thought*, a thin conception of selfhood is devoid of meaning.⁹¹

2.9. A Terse Reflection on Thought

I am, whatever I am, experiencing thought. This experience cannot be denied. Is thought *really* occurring? Perhaps not, but the feeling of the experience of thought is as real as real. But what is thinking and is it something “I” can do? I certainly talk as though I am the one doing the thinking.⁹² This Cartesian language permeates the cross-cultural zeitgeist. For Strawson, the point of whether or not it is “me” doing the thinking is not as important as the idea that there be an experiencer for thoughts to appear in the first place. In other words, it is not a question of will but a question of a necessary witness. When considering the will, which I discuss in much detail in the succeeding chapter, it seems quite trivial to acknowledge the lack of causation from any kind of central locus of thought.

⁹¹ Galen Strawson. “What Is the Relation between an Experience, the Subject of the Experience, and the Content of the Experience?” 283.

⁹² “I just had an amazing idea!” or “I cannot think straight today.”

For I cannot will what I think in the next moment since willing takes thought. Thought is prior to causation.

The more important question of a necessary witness is much more difficult to grapple with. Strawson claims for there to be an experience at all, there must exist a subject for which the experience “is for.”⁹³ This is where I disagree with Strawson’s argument. I would not venture as far as to claim that there is no subject or even an experience as a subject. I would, rather, suggest that this experience in this moment is not done by “me” or for “me” at all – I reject both. It cannot be done by “me” since cognitive thought is not done by “me,” it merely happens to me. In the same way, an experience merely occurs, upon which “I” take credit. An experience is happening to itself by itself without any necessary concept of “I,” which is nearly impossible to prove due to some illusory sense of an ego standing in the way. Experience of cognition and reasoning certainly feel as though it is “for me,” but the concept “me” is unnecessary and most likely illusory. It would be more accurate to say, “the experience happened to it, as a process of happening” as opposed to “the experience happened to, or by, me.” Upon reflecting on this meditation, I find no room for justifying any notion of a synchronic self even with my admission of the undeniable phenomenal experience in each moment.

⁹³ Galen Strawson. 2011. *Selves: An Essay in Revisionary Metaphysics*. 419.

III. Free Will and the Dark Mystery of Appearing

This investigation delves into the differing schools of thought concerning determinism and indeterminism: anti-libertarian incompatibilism, compatibilism, and metaphysical libertarianism. I will cover issues ranging from Peter van Inwagen's deterministic *Consequence Argument*, the compatibilist's *Harmony Condition*, quantum physics, self-determinism, randomness, reason responsiveness, and, most importantly, the origin of thoughts. In this work, I will outline the views of each school of thought, and I will object to free will as an anti-libertarian incompatibilist would. In response to Van Inwagen, E. J. Lowe and Philip Goff attempt to defend free will with what I will call the *Reason Responsiveness Argument*. I will present a case as to why such an explanation for free will does not suffice due to, what I am calling, the dark mystery appearing: the appearing of thoughts. I do not argue any point derived from the issues of responsibility in regard to the lack of free will.⁹⁴ This work remains in the purview of metaphysics and some basic logic. In my closing remarks, I lay out my views on why the self is at the center of all considerations involving free will leading to the conclusion of this work where I find that a more loving world would result from the recognition of the incoherency of selfhood and metaphysical freedom.

3.1. Determinism and the Consequence Argument

It is most appropriate to commence with anti-libertarian incompatibilism and a defining of first principles. Anti-libertarians believe that if determinism is true, there is no room left for the possibility or actuality of free will. Determinism is the thesis that at every moment, the way things then are determines a unique future that only *one* of the alternative futures, which may exist relative to a given moment, is a physically possible continuation of the state of things at that moment.⁹⁵ Free will is defined, by incompatibilists, in a two-fold manner.⁹⁶

⁹⁴ I save this for the concluding remarks of the work.

⁹⁵ Peter Van Inwagen. *Metaphysics*. 3rd ed. Boulder, Colo: Westview, 2009. 254.

⁹⁶ As we will see, differing schools of thought define free will differently in order to aid and justify their theory.

It is necessary for the agent to have the ability to have done otherwise. As Robert Kane puts it, “the existence of alternative possibilities is a necessary condition for acting freely, or acting ‘of one’s own free will.’”⁹⁷ Imagine the timeline of the universe was rewound one billion years in the past – with the same conditions as existed one billion years ago – and it were to play out again. Would I have the ability to *not* write this paper and go for a walk instead? If determinism is true, I would not. Thus, a detriment would be placed upon the possibility and actuality of me freely willing to write my paper or go for a walk. The discovery of causal explanations for our actions, preferences, and decisions shows that we could not have done other than we have done.⁹⁸

Second, it is necessary that there exists *ultimate self-causation* – that is, my being the very first and ultimate uncaused cause of my actions.⁹⁹ Self-causation presents many questions underlying what it means to be a self or what the self even is. But the idea is that an agent – some notion of a self – can act without being fully determined to act by causal factors outside the agent’s control. This accords with exactly how we feel in every moment of everyday existence.¹⁰⁰ So, for the time being, assume self-causation is metaphysically possible. The anti-libertarian incompatibilist recognizes the rather obvious contradiction between the two requirements for free will and physical determinism. Assuming our brains are subject to the same deterministic laws of nature, for the anti-libertarian incompatibilist, free will is an impossibility¹⁰¹. The best explanation of this theory is found in Dr. Peter van Inwagen’s *Consequence Argument*.

⁹⁷ Robert Kane. "Introduction: The contours of contemporary free will debates." In *The Oxford handbook of free will*. 2002. 8.

⁹⁸ A. C. Macintyre. "Determinism." In *Free Will and Determinism*, edited by Bernard Berofsky, 240-256. Harper and Roe Publishers, 1966. 241.

⁹⁹ Ken M. Levy. *Free Will, Responsibility, and Crime: An Introduction*. First published. New York London: Routledge, Taylor & Francis Group, 2020. 9.

¹⁰⁰ Ken M. Levy. *Free Will, Responsibility, and Crime: An Introduction*. 74.

¹⁰¹ What about the immaterial existence of the self or soul where the laws of nature do not apply? I discuss this below.

To better ascertain this argument, van Inwagen defines what he calls *untouchable facts*. These facts are those which no human being can do anything about and that no human being in all of history could ever have done anything about.¹⁰² If a person has no control over x, and x necessarily results in y, it follows that the person has no control over y. For if I have no control over the last saber tooth tiger dying before my birth, it follows that I also have no control over my not being able to witness a living saber tooth tiger, making it an *untouchable fact*. In the same sense, the *Consequence Argument* suggests that if my present action is the result of past conditions of the universe plus (+) the laws of nature (which are also out of my control), it follows that my action in the present moment is not of my own doing. It is, rather, that a mere happening occurs in this present moment whereby I have no claim to legitimate initiation.¹⁰³ The past condition of the universe is an *untouchable fact*, to use van Inwagen's term, just like the laws of nature. So, my present action is deemed to be what it is, and thus my present action would have to be an *untouchable fact* also. Thus, I cannot do otherwise – I cannot be said to possess free will.¹⁰⁴ If this is true, my *self*, if it is a real entity, could not have caused itself to act as I do in the present moment. The self-causing was anything but self-caused – for the causation was determined by the prior conditions of the universe and the laws of nature. Therefore, an anti-libertarian incompatibilist and defender of the *Consequence Argument* would suggest there can be no ultimate self-causation under this view. For van Inwagen, this argument undermines both necessities for the existence of free will. However, some do believe in the compatibility of determinism and free will – known as compatibilists.¹⁰⁵

¹⁰² Peter Van Inwagen. *Metaphysics*. 260.

¹⁰³ This is a quite important point when discussing the self as an incoherent concept. For if the self is some fiction, then it becomes less about what "I" did and more about what happened (to "me").

¹⁰⁴ To save free will or allow for a differing presentation in this context, one would have to reject determinism (which many do) or change the laws of nature which seems to be an impossibility.

¹⁰⁵ In fact, compatibilists believe determinism is necessary to justify free will. An indetermined universe, for them, would not allow for free will. I say more on this below.

3.2. A Problem of Definition

Since I agree with van Inwagen and therefore identify as an anti-libertarian incompatibilist, the previous section did not possess critiques of each point made. I will, rather, present the appropriate critiques of anti-libertarian incompatibilism for each school of thought as I proceed. In addition, I will answer the critiques and positions held with an anti-libertarian incompatibilist response.

Primarily, the compatibilist disagrees with the incompatibilist (anti-libertarian or otherwise) on the definition of free will. While the latter views free will as necessarily possessing the “could have done otherwise” condition and ultimate self-causation, the former defines free will as nothing more than the ability to act in accordance with what the subject wants or desires.¹⁰⁶ If I desire to eat a piece of cake more than I desire not to and I act in accordance with my desires and eat a piece of cake, then I have freely willed and acted. Many enlightenment philosophers such as David Hume and Thomas Hobbes agreed with this sentiment. A. J. Ayer represents the compatibilists’ view best when he says, “for it is not causality that freedom is to be contrasted with, but constraint.”¹⁰⁷ A causally determined action does not necessarily mean one is constrained to do it. As long as the agent is not physically prevented from, or coerced into, acting or willing in accordance with their desires, they are said to possess free will. And since our desires may (and probably do under a deterministic universe) result from the past conditions and laws of nature, there is no contradiction between determinism and free will according to the compatibilist.

The anti-libertarian incompatibilist responds by first rejecting this definition of free will.¹⁰⁸ It may be true that no one in this moment has compelled me to get up and walk across the room, but if my doing so can be causally explained in terms of the past conditions of the universe and the laws of

¹⁰⁶Ken M. Levy. *Free Will, Responsibility, and Crime: An Introduction*. 17.

¹⁰⁷ A. J. Ayer. “Freedom and Necessity.” In *Free Will*, edited by Gary Watson, 15-35. Oxford University Press, 1982. 19.

¹⁰⁸ Free will as acting in accordance with my desires or lacking any constraint.

nature, then how am I any more free than if some other person had compelled me?¹⁰⁹ As Ken Levy states,

Free will seems to require both the ability to do otherwise and the ability to want otherwise, it also seems to require the ability to want to want otherwise, which would have to be translated as: I would have wanted to want otherwise if I had wanted to want otherwise. So traditional compatibilists' attempts to reconcile the ability to do otherwise with determinism seem to lead to an infinite, and therefore nonsensical, regress of conditionals.¹¹⁰

For there is no manner in which an agent can be said to determine their own desires – they merely appear to us in consciousness.¹¹¹ It is for this reason, lacking the ability to self-determine one's desires even if they are fulfilled, that compatibilism fails. They fail to ask: "Did 'I' cause the desire for a piece of cake?"

Compatibilists make an argument contrasting with metaphysical libertarians, as I will show below. This argument has to do with the fact that free will necessarily requires determinism. Self-determinism or ultimate self-causation, in an indeterministic universe, is void of sense and should be viewed as absurd. Just as spontaneous motion by the body is undetermined, not self-determined, so too with spontaneous eruption in the brain. Like a twitch, convulsion, or spasm, an uncaused choice or action just happened *to* me rather than *by* me.¹¹² Any freedom incompatible with determinism would require indeterminism; and what is undetermined would happen by chance or luck and could not be free and responsible action.¹¹³ I will say more on this in the subsequent section.

Thus far, two theories on free will affirm determinism but differ on its compatibility with free will – anti-libertarianism and compatibilism. Disagreement on the definition of free will is fundamental to defending or rejecting free will in a deterministic universe. This is best understood through analogy.

¹⁰⁹A. C. Macintyre. "Determinism." In *Free Will and Determinism*. 21.

¹¹⁰ Ken M. Levy. *Free Will, Responsibility, and Crime: An Introduction*. 24

¹¹¹ One may argue that we can refuse or accept these acts based on said desires, however this is not a compatibilist argument but a libertarian argument.

¹¹² Ken M. Levy. *Free Will, Responsibility, and Crime: An Introduction*. 11.

¹¹³ Paul Russell and Oisín Deery, eds. *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*. New York: Oxford University Press, 2013. 194.

Plato describes what seems to be a fictional island, Atlantis. Say we have two people, person X and person Y. Person X suggests that Atlantis was never a real island and was a creation in Plato's mind. On the other hand, person Y suggests Plato was writing about the island of Sicily and used Atlantis as its referent. While person X rejects the actuality of Atlantis, person Y changes the definition of what Plato meant by Atlantis to a real island: Sicily. Person X would represent the anti-libertarian incompatibilist view, while person Y would represent the compatibilist view regarding Atlantis' existence. For the compatibilist to affirm free will or Atlantis, they change the definition to justify their claim. Free will becomes not that ability to do otherwise but acting without being physically prevented or coerced. If this was the agreed upon definition, free will would be an unfalsifiable position – one that does not get at the underlying issue. This definition of free will is insufficient because it fails to consider whether or not one's acting without coercion was actualized by the "self" or things outside of it.¹¹⁴ In a way, the two schools are arguing past one another due to their lack of agreement on the meaning of key terms.

3.3. Indeterminacy and Randomness

While the anti-libertarian and compatibilist agree that we live in a deterministic universe, there is a third school of thought, metaphysical libertarianism, which rejects determinism. Since metaphysical libertarians affirm indeterminism and affirm free will (hence the appeal to liberty), they too are considered incompatibilists. Metaphysical libertarians believe the agent, through a variety of justifications, owns the moment in which a choice is made due to the indeterminacy of the universe. In the moment of effort or reasoned deliberation, as we will see below and in the next section, an agent is credited full responsibility for the choice made. For many philosophers in the last few centuries, however, the universe functioned as a physical law-possession, deterministic machine of

¹¹⁴ As stated above, this is the reason why compatibilism fails. There is a deeper question begging to be answered that goes beyond the fulfillment of desires.

cause and effect. It made little sense to affirm free will in the absence of deterministic laws. These philosophers, rather, scrupulously attempted to marry determinism with free will. With the advent of quantum physics, indeterminism has become a much more justified position. Modern quantum physics, according to its unusual interpretations, introduced indeterminism into the physical world, giving us a more sophisticated version of the Epicurean chance “swerve of the atoms.”¹¹⁵

Unfortunately for libertarians, according to Robert Kane,

Some twentieth-century scientists and philosophers have suggested that free will might be rescued by supposing that undetermined quantum events in the brain could be amplified to have large-scale effects on choice or action. Unfortunately, this modern version of the ancient Epicurean “swerve” of the atoms seems to be subject to the same criticisms as its ancient counterpart. It seems that undetermined events in the brain or body, whether amplified or not, would occur spontaneously and would be more of a nuisance – or perhaps a curse, like epilepsy – than an enhancement of freedom.¹¹⁶

Ultimately indeterminism reduces events to randomness or mere chance. At the very least, randomness can be reduced to those immaterial laws which the human conceptual makeup cannot ascertain or comprehend.

Imagine two agents who experienced the exact same past up to the present moment whereby they can choose to lie or tell the truth. One chooses to tell the truth and the other lies. If they did in fact live identical past lives and their choosing was not determined¹¹⁷ but was the result of chance or randomness due to the indeterminacy of the universe, could we really hold the liar accountable and the truther praiseworthy?¹¹⁸ The rather intuitive idea of settling on the randomness of indetermined events, as anti-libertarian incompatibilists and compatibilists both do, is best expressed in the Mind Argument.

¹¹⁵ Robert Kane. "Introduction: The contours of contemporary free will debates." 4.

¹¹⁶ Paul Russell and Oisín Deery, eds. *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*. 195.

¹¹⁷ Either by the laws of nature, self-causation, or some other form of causation we do not yet understand.

¹¹⁸ Paul Russell and Oisín Deery, eds. *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*. 197.

The Mind Argument may be stated like this: if one's actions do not follow *deterministically* from one's previous states, then it is the result of an indeterministic process, and one is unable to determine the outcome of an indeterministic process.¹¹⁹ While the anti-libertarian incompatibilist may focus solely on the mere randomness of one's actions, under indeterminism, a compatibilist would focus on this Mind Argument. For the self to determine its own decisions based upon motives, it is supposed that these motives necessarily play a deterministic role in one's deciding to act.¹²⁰ Without determinism (of a kind), those actions are derived from motives outside of the agent and are separated from the "choosing" of the self. That is to say, without a determined causal link – from the self, deliberation of the options, reasoning, choosing, and then acting – it makes the choosing of an action and the action itself utterly random since there was no reason, from the self, why it landed on choice A over choice B. Nevertheless, metaphysical libertarians address these concerns in many profound ways.

There is a lucid counter argument to this: the potential immateriality of the universe. While most would grant the existence of physical objects in the objective realm, perhaps this is not all there is and non-physical entities *are* within the universe's possession. If there are non-material goings-on outside the domain of human experience or understanding, this seemingly opens the possibility to an ultimately willed self without prior physical necessities coercing the self onto a specific path. Immaterialism would allow for an unexplained method (outside the realm of physics) for the conscious agent to consider past and potential future events to help determine a completely self-determined decision and action. The material and objective realm would no longer be the only necessary nexus the self interacts with. There would exist properties of the universe that do not align with the physical determined properties, thus allowing for an unexplained explanation for self-causation. That is, an

¹¹⁹ Peter van Inwagen. *Thinking about Free Will*. Cambridge, United Kingdom; New York, NY, USA: Cambridge University Press, 2017.162.

¹²⁰ This is very similar to the argument made by E. J Lowe Philip Goff, whereby the agent adheres to reasons in order to freely respond, which I will refute below.

explanation we are ignorant of – creating an epistemic gap. Those who promulgate such an argument, however, fail to understand the nature of determinism. For even an immaterial universe would be influenced by immaterial laws. For it would have to in order to avoid the randomness argument. In a sense, the physical universe runs on immaterial laws already, for the law is merely an explanation, a concept of the mind, not a physical entity. It makes no difference whether the universe is purely physical, purely immaterial, or a combination of the two. The essence of all randomness is the lack of ability to predict. In other words, that which we cannot fully engulf with natural physics is said to be unknown or completely random. That which is unknown could potentially be labeled as that which is random. There would exist “laws” for which these beings would naturally tap their essence into, for they would have no other choice. Otherwise, they would cease to be.

Take a similar argument as applied to the physical body and its potentially immaterial soul. Intuitively, we apprehend the idea that our bodies are determined by past events, genetics, eating habits, exercising habits, etc. and yet we never think this way when considering our immaterial soul. We believe the soul could possess qualities beyond their being determined by material existence. In other words, the soul is the creator of itself and exists as a placeholder for that which cannot be explained by the physical body. Whereas the body is left to the precariousness of the objective. This way of thinking is erroneous due to the creation of an unwarranted line of distinction. While I would agree that the body is physically affected by that beyond my influence, the immaterial soul should be thought of in the same way. For if the soul is assumed to exist, there is no one who could take credit for their possessing their soul as opposed to another’s soul. My soul, or my character, is no more of my doing than my body is. This is like suggesting that the chair chose to possess the soul of said chair when, in reality, that chair could do nothing to change its essence as that chair. For if it did, it would cease to be that chair. Therefore, the individual who claims responsibility for their material body or their immaterial essence and soul are under a complete misapprehension. In the same vein, the

materiality or immateriality of the universe makes no difference to me when discussing the will of the agent. Neither explanation of the cosmos allows for the agent to exercise any kind of self-causation. Simply put, regardless of the material or immaterial essence of the universe, each are either completely determined or random.

Metaphysical libertarians attempt another way around the randomness objection by claiming freedom is saved through the *effort* of the agent. If an agent has two options, for simplicity's sake, and they struggle in the process of coming to a decision, it is said that this saves their freedom from randomness. As opposed to a random ceasing of oscillation between two choices, when an agent puts in the effort of landing on the correct choice, said choice was the ultimate result not from chance but from the agent themselves. In short, even if a level of effort is random, it can be fully self-realized.

The first problem with justifying free will through effort is that it makes metaphysical freedom seem like a concept on a spectrum. That is, there exists varying degrees of freedom depending on how much or little effort was put in by the agent. In terms of politics, varying degrees of freedom makes sense to discuss, but metaphysically speaking, freedom is an all or nothing concept. Perhaps I am misinterpreting the theory of effort. It could be the case that when effort is present at all, at any degree, there exists ultimate metaphysical freedom and when there is no effort of will, like in an inanimate object, there is no ultimate metaphysical freedom. For argument's sake, I will grant the more charitable reading.

The issue with freedom originating in the effort of the agent is that the origination of the idea to give more or less effort is a thought of the agent. In saying it is a thought of the agent, it seems as though the argument still holds, but what I mean by "a thought of the agent" is that the agent experiences a thought to perform more or less effort – that is, the thought is *not* of the agent's own creation. So, for an agent to give more or less effort in trying to make a choice, there must be a reason

for them engaging in or not engaging in the effort. A reason¹²¹ is a thought – so to engage in effort or not originates in thought or consciousness. The question then becomes: is the agent responsible for this thought? Certainly not – for we are not the conscious authors of thoughts, they merely appear to us in consciousness. I will grant that effort comes from the thing we call an agent, but the thought of performing the effort came from nowhere – so there is no justification for crediting an agent with the effort performed.

3.4. Reason Responsiveness Argument

E. J. Lowe¹²² has developed an explanation for free will which, I think, poses the greatest response to the Randomness Argument. Going forward, I will label Lowe’s argument the Reason Responsiveness Argument or RRA, for short. Philip Goff, a proponent of Lowe’s argument, acknowledges the Consequence Argument when he states that if materialism is true, all such explanations of human affairs are false since people are physical objects, and hence what they do is determined not by considerations of reason or rational considerations, but by mechanical causes.¹²³ In the assumption of a purely physical universe, which Goff does not endorse, the notion of free will becomes absurd. For if all there is are atoms in motion, Laplace’s demon could predict every future event of the universe – our brain states and the actions they produce included – by calculating the mass, velocity, and direction of each to determine, with 100% accuracy, the future states of the universe. Since there are three approaches in regard to free will – physical determinism, the randomness of indeterminacy, and ultimate self-causation – and RRA does not fall squarely into one of these, it becomes an approach all its own. Lowe and Goff both reject physical determinism and in

¹²¹ More on reason responsiveness in the next section.

¹²² Philip Goff also adds to and references the work of E. J. Lowe which I will cite and respond to below.

¹²³ Philip Goff. *Galileo’s Error: Foundations for a New Science of Consciousness*. 1st ed. New York, NY: Vintage Books, 2019. 196.

order to answer the objection of the randomness argument, they posit RRA as a special version of ultimate self-causation.

Advocates of RRA respond to the arguments made by those who believe there is a major issue with an indeterminate universe leading to the randomness argument. To rehash, this argument claims that for there not to exist physical or non-physical laws of nature, events are no more the result of each event coming about through a cosmic roll of the dice. This would apply to those events that many would already identify as random, such as a gust of wind or the picking of a single card from a deck. However, it would even apply to those decisions and actions which are seemingly the events of mental causes of an agent. If our actions are not the result of some long line of causal necessities, they must be the result of some random chance event(s) which grants an agent no more claim to freedom than the deterministic laws would. One's decision to enjoy vanilla as opposed to chocolate ice cream would be determined not by the agent – even though it may feel this way in the moment and upon reflection – but by random mental roll of the dice. Lowe and Goff think there exists a route around this randomness conundrum – reason responsiveness.

RRA claims free choices can be distinguished from random events by the fact that they involve responsiveness to rational considerations.¹²⁴ For example, a person in a classroom could decide *to* raise their hand or *not to*. The reasons *for* doing so could include, but are not limited to, the fact that they desire to have a question answered, and perhaps a desire to gain participation points in class. On the other hand, they are quite taciturn and knows the aforementioned reasons can be fulfilled by visiting the professor during office hours, reasons *for not* raising their hand. After taking both the reasons for and against raising their hand into account, they ultimately decide not to raise their hand. The reasons *for not* raising their hand outweighed those reasons *for* raising their hand and thus, they do not raise

¹²⁴ Philip Goff. *Galileo's Error: Foundations for a New Science of Consciousness*. 1st ed. New York, NY: Vintage Books, 2019. 198.

their hand. But even in the course of action in which they decided to raise their hand, this action, too, would have been in response to reasons which they deliberated upon. Choice, by its very nature, is never exercised ‘blindly’, but is always informed by or responsive to reasons for action.¹²⁵ Being aware of these reasons and deliberating on them is a process which is not determined, and it goes beyond the realm of randomness. Reasons, of which agents are fully aware, cannot merely be chalked up as random since the human condition reacts to reasons that play a fundamental role in determining our choices and subsequent actions. As Goff says, “Any explanation is dependent on the assumption that human affairs are determined by the responsiveness of people to rational considerations.”¹²⁶ However, I am afraid Goff has shot himself in the foot by acknowledging the determining aspect of this process. For reasons are pondered, a decision is made deterministically, and an action results from said decision and reasons for it. It is held that the pure event of the deliberation process is uncaused and not subject to chance. Yet this only makes sense when it is assumed that the process is a result of an agent, the uncaused causer, who is in control of their thoughts prior to those thoughts appearing. It is said that the agent can choose not only how to move their body but also how to direct their thoughts, and the latter kind of choice is essentially involved in all processes of deliberation.¹²⁷ It is with this argument that I think the RRA has failed, but I will address this in my argument on the darkness of appearing below.

The picture painted by RRA thus far can be best summarized by E. J. Lowe when he says, In certain situations, a rational agent is confronted with two or more alternative courses of action, which are genuine alternatives in the sense that, at the time at which the agent must choose between them, no sufficient cause already exists of one of the prospective outcomes. The agent then deliberates, by weighing the reasons for and against the various different alternatives. Finally, the agent makes a decision and chooses to act in one way rather than another. His mental act of choice may then, and

¹²⁵ E. J. Lowe. *Personal Agency: The Metaphysics of Mind and Action*. First published. Oxford, New York: Oxford University Press, 2008. 194.

¹²⁶ Philip Goff. *Galileo’s Error: Foundations for a New Science of Consciousness*. 196.

¹²⁷ E. J. Lowe. *Personal Agency: The Metaphysics of Mind and Action*. 193.

normally will, contribute causally to a chain of subsequent events issuing in the performance of the chosen action.¹²⁸

Prior to an agent's choice, there were no necessary conditions which caused them to choose the way they did. In other words, there was no need for any antecedent cause for action. The next question invariably presents itself: what *explanation* accounts for one action over another? There are four potential explanations for action:

[A] mental (brain) states (pure physicalism),

[B] freedom from effort and RRA,

[C] desire and belief accounts,

[D] the dark mysterious appearance of thoughts of consciousness.

[A] has been dealt with – for if pure physicalism is true, Van Inwagen's Consequence Argument would be sound.¹²⁹ [B] has two potential objections which are found in [C] and [D]. [C], desire and belief accounts, is not to be credited to me and will be discussed below. The ideas in [D] are my own thoughts and arguments against freedom from effort and RRA which is the explanation I will endorse. I will commence with [C], desire and belief accounts as an explanation for action.

Take the example of a construction worker on site who hears a loud noise above them which causes them to look up and witness a 4x4 bearing down upon them. If all stays the same, they would be the victim of gravity and the 4x4. The worker believes that a 4x4 is imminently about to fall on them, they believe this will result in their injury, and they desire not to be injured. Considering these beliefs and desires, according to the RRA, the worker can only be said to be free if they use these as reasons to *choose* to side-step and avoid the 4x4. Possessing a desire not to be injured, their action of jumping out of the way of the 4x4 will only qualify as a free and rational action if they *choose* to jump out the way in the light of this desire. If the desire merely *causes* them to jump out the way and the

¹²⁸ E. J. Lowe. *Personal Agency: The Metaphysics of Mind and Action*. 187-188.

¹²⁹ It very well may be true, but why stop investigating?

power of choice is not exercised by the agent on this occasion, it is not an instance of rational action.¹³⁰ RRA views beliefs and desires as a certain manifestation of reasons. In what sense, however, can the worker be the author of their beliefs and desires? If they are truly responsible for them, it follows that their reasons and actions that follow are also under their jurisdiction. I find that the worker could not *but* believe a 4x4 was bearing down on them and could not help *but* desire not to be injured. Imagine what it would take for the lack of this belief and desire. The belief would never originate if the 4x4 never started falling – but with this originating outside the control of the worker, what could have made them not hold the belief that a 4x4 was bearing down on them as one *was* certainly doing just that? There seems to be nothing. In the same way, the desire not to be injured could not *not have* manifest itself through the worker. For any person either does or does not desire to be injured – so no ultimate decision made there.

All things considered, we cannot grant the worker ultimate self-causation when the reasons for their decision and action were the results of beliefs and desires completely out of the control of the worker. Sure, it may be argued, but not all reason responsiveness actions are willed because of their connection with one's beliefs and desires. There exists situations where purely rational thought allows one to act in such a way that is consistent with RRA. While I grant this, I always come back to the question: where do these reasons come from? And the answer is what I am calling 'the dark mystery of appearing.' It is through this mystery that I will present the gravamen of my case; that thoughts are no more of my doing than my beliefs and desires are.

3.5. The Dark Mystery of Appearing

To prove any theory or concept erroneous, one necessarily stoops down to the level of said concept itself. Free will fits this quite well, as did the ātman and self. In arguing for the illusory nature

¹³⁰ E. J. Lowe. *Personal Agency: The Metaphysics of Mind and Action*. 186.

of free will, I necessarily grant credence to the existence of the concept of free will. For this reason, I believe ‘free will is an illusion’ should be taken as illusory as well. As Lao Tzu says in the *Tao*,

The ways that can be walked are not the eternal Way;
The names that can be named are not the eternal name.
The nameless is the origin of the myriad creatures;
The named is the mother of the myriad creatures.¹³¹

There is no illusion of free will just as there is no illusion of self and the more one attempts to disprove these concepts, the more they permeate through. I have been convinced that there is no free will; I cannot view the world or my self as anything which possesses this illusion of free will. I say this because I reject any notion of an egoistic explanation of “me” and the workings of that entity, whatever it may be. And it follows from this that free will is not only an illusion, but is an entirely incoherent concept, or non-concept. The self and free will are obviously two sides of the same coin, but in this final section on metaphysical freedom, I will address the illusion of the illusory nature of free will as the result of the dark mystery of appearing thoughts.

I have been presenting free will as the ability to do otherwise in addition to some nucleus of self as the ultimate authorship of thought and action. In this latter respect of free will, the dark mystery applies. Through understanding the mystery as an explanation for action, RRA cannot be entertained as a support for free will or any notion of metaphysical freedom of will.

It is thought that we are the conscious authors of our thoughts, reasons, choices, and actions – each of which have a causal connection to its predecessor. But what comes before thoughts? It is here that the self-causing agent is posited. The illusion of agent causation is the illusion that we are the free, uncaused agents of our actions, acting on the world, but causally unconstrained by it.¹³² Randolph Clark presents the self-causal account well when he says,

¹³¹ Lao Tzu. *Tao Te Ching*. Trans. Victor H. Mair. New York, New York: Bantam Books Publishing, September 1990. 59.

¹³² Jay Garfield. *Losing Ourselves: Learning to Live Without a Self*. 33.

An agent's control over her behavior resides fundamentally in her *causing* what she does. Her control does not reside fundamentally in her performing some special sort of action. Since causing is bringing something about, producing it, or making it happen, causing seems to be the right sort of thing on which to base an agent's control over her behavior.¹³³

Why does one choose A over B? Because the reasons, brought about by thoughts, for A determined the uncaused self to choose A over B in the process of deliberation which is said to be the agent's *causing*. For RRA, this is what it means to be free, but this merely pushes the enquiry further whereby one should realize that this explanation does anything but explain. Where do these reasons and thoughts come from? Since thoughts cause reasons to be pondered, the real question is simple: *where do thoughts come from if not from the authoring self?*

Are you responsible for the thoughts that appear in consciousness? It most certainly feels like you are, but perhaps we can reject this false sense of certainty by presenting rather obvious ways in which you are not responsible. For example, you would not argue that you are responsible for your genetic makeup. That is, your natural hair color and your height are not aspects you would claim responsibility for – for they are completely out of your control. Similarly, the country you were born in, and the socioeconomic status of your family are, also, completely out of your control. Yet these things help define who you are in a very impactful manner, but you cannot truly take credit for them happening to you. What about the processes occurring in your body as you read this? Are you controlling your heartbeat? Can you take credit for your fingernails growing? Certainly not. This is all to say that many, if not all, aspects and occurrences which happen to “you” are happening without an ego – without the ultimate authorship of action that we all think we possess. I am merely arguing that the conscious happenings of our minds are no different than the heart producing red blood cells or

¹³³ Paul Russell and Oisín Deery, eds. *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*. 2013.

your fingernails growing. In a similar way, thoughts appear to us without any explanation whatsoever, including self-authorship. This is the dark mystery of appearing.

I want you, the reader, to think of an animal; any animal that has ever existed. It could be extinct or living, but just think of an animal. Got it? Now, reflect on the process that just occurred. Were you free to pick any animal or merely the *one* that came to you? You certainly were not free to think of any of the animals which you are unaware of existing, past or present. So, of the animals that you are aware – those in your mental filing cabinet – you decided upon one. Did you choose to think of that animal, or did it come to you without explanation? For if you chose to think it, you would have had to think to choose to think it. And if this were the case, you would have had to think to choose to think to choose to think it, ad infinitum. It makes no sense to think of a thought before you think it. There is never an original source by which a thought appears in consciousness. It is always appearing from a dark chasm. We notice the thought after it appears to us, and we immediately take credit for it – we take responsibility for the thought. It becomes *mine*. But a thought appearing is no more my doing than my heart beating. They both have an equal claim to the notion of “me” doing it – that is, none at all. We are not the authors, but we are the conscious observers of our lives. Thoughts are mere happenings which give rise to reasons which give rise to choices, which give rise to actions. But as they are causally connected, and since we have no claim upon the primary causer – the thought – it follows that we have no claim upon the rest of the chain.

It may be argued by the metaphysical libertarian, however, that even if we do not possess agency over the thoughts that appear, the agent still possesses the ability to follow or reject said thoughts. That is, there is no agency in the thoughts themselves, but agential freedom comes in determining how to use those thoughts, which is totally consistent with our everyday experience. For it certainly feels as though I am the author of my thoughts, but assume I am not. I still feel like I am the one that debates my interior self on these thoughts. I am the one that adopts an idea, finds a piece

of artwork beautiful, or sees a person as immoral. These thoughts and their antitheses came to *me*, and *I* accepted one while rejecting the other(s). Thoughts are organized by me while my ultimate metaphysical freedom comes through the way I interact with them.

I would counter this in a very similar manner in which I countered the idea that freedom comes through the effort of the agent. This theory fails to recognize that the interaction with thoughts is, itself, a thought. Accepting or rejecting a thought necessarily exists as a thought, otherwise it would not have appeared to the agent. In other words, an agent's rejecting to eat a piece of cake, a thought, is also its own thought appearing in consciousness. When I have a thought, it is immediately followed by another succeeding thought – either to accept it, reject it, or observe it. This boils down to the fact that, if I am not the conscious author of my thoughts, I am also not the conscious author of the thoughts involving preceding thoughts. From this, even though it feels like I am the one accepting or rejecting thoughts as they appear to me, the “accepting” or “rejecting” of thoughts are not done by “me” but occur to “me” without my causing them – as all thoughts do.

3.6. Summary

I identify as an anti-libertarian incompatibilist for the objections raised throughout this work against compatibilism and metaphysical libertarianism. From the outset, compatibilists define free will erroneously. For if doing what I desire is freely willing – a much more fundamental question presents itself: in what way did I choose what I desire? Coercion and constraint are the only preventors of a free will according to the compatibilist. On a deeper level, my acting in accordance with my wants or desires gives me no freedom whatsoever. Say I want to save my latest paycheck, so I deposit the money in the bank for safekeeping. My wanting to save the money – and the subsequent action reflecting that sentiment – is fully compliant on a free will for the compatibilist. But could I choose to want to waste the money? This is not the same question as asking whether I have the ability to choose to waste the money – rather, could I *choose to desire* to waste it? Of course not. I desire to save the money and there

is nothing, in that moment, that could change that. This desire to save the money is presumably caused by some deterministic or indeterministic law which orients my brain, physically or randomly, in the manner of desiring to save money. Therefore, my willing to save money, while acted upon, is still determined by my desires which are determined by previous conditions of the universe and the laws of nature – thus it is an untouchable fact or out of my control.

The metaphysical libertarian's self-determination doctrine of free will contradicts itself in its affirming indeterminism. All events in an indeterministic universe must necessarily reduce to random chance¹³⁴ which gets the libertarian nowhere closer to freedom. My choosing anything becomes a random stopping of the oscillation between two or more options. It becomes less about the self choosing a certain action of their own doing and more to do with mere random happenings.

Reason responsiveness addresses the randomness argument by purporting that the human being, the agent of causation, responds to the weighing of reasons for a particular action. Through the process of deliberation, the agent considers the best course of action by having their choices and actions determined by those reasons in favor of a certain action. As we saw, the belief and desire explanation of action undermines this, but only if one views beliefs and desire as separate from reasons contemplated. RRA theorists identify those beliefs and desires as reasons and thus the theory holds its weight. Nevertheless, the dark mystery of appearing as an explanation for action does not allow for reasons, through the form of thoughts, to be considered of the person's doing. Thoughts merely appear, through a happening, which cannot be of the person's volition.

Neither determinism nor indeterminism grant the individual a free will. The human interpreter has set us up for a fall. It has created the illusion of self and with it, the sense we humans have agency

¹³⁴ Unless a dualist framework of the mind is posited however, I would reject this. Not to mention the randomness of the immaterial "soul" or "essence" of the person. How could one choose their soul or character, since this, too, would be subject to chance of indeterminism, one is still no closer to freedom of the will with an immaterial soul than with a monist approach to the mind.

and “freely” make decisions about our actions.¹³⁵ I do not believe the self can cause events without being the effect of prior causes – random or determined. For reasons mentioned in other sections of this thesis and my intuition on the nature of the self as a non-concept, the conclusions on free will have resulted. To recap, the “I” is an illusory illusion or at the very best, an evolutionary benefit for survival, but not something independent or primordial to a mental or bodily essence. For the self is at the center of any discussion connected to free will. Without the self, these debates and conversations are futile. It is my belief that the self is something which cannot be denied in affectivity, similar to free will. There is no denying my feeling that I possess this author or thinker of thoughts within me. Upon a slight reflection, however, about how thoughts – and thereby actions – originate into consciousness, one realizes that they are not the author of these thoughts. Thoughts appear in consciousness without invitation, yet we take credit for their happening. For my “self” to have authored a thought would consist in my thinking my thought before I think it. And to think a thought before I think it assumes I thought my thought in order to think my thought before I think it, ad infinitum. This is the true essence of the dark mystery of appearing. Ultimately, the first thought, wherever it originates, necessarily originates in utter darkness, of which I am in no control. Therefore, I am not the author of my thoughts but a thing which experiences thoughts happen. I, thus, possess no traditional notion of freedom in this process.

¹³⁵ Michael S. Gazzaniga. *Who's in Charge? Free Will and the Science of the Brain*. 1st ed. Gifford Lectures 2009. New York, NY: HarperCollins, 2011. 105.

IV. Conclusion

In my closing remarks, I will give a rather terse consideration of the ramifications that result from what has already been stated.¹³⁶ What can we expect in a world where the concepts “self” and “free will” are considered incoherent? The choices we make still affect the material conditions of society and nature, but are we to be held ultimately responsible for said choices? I will consider an argument against the notion of ultimate responsibility – making responsibility, along with the self and free will, an incoherent concept. Galen Strawson calls this argument the “Basic Argument.” I will defend it while I also, and more importantly, investigate the potential consequences of adopting such a view. Without concepts like “self” and “free will,” I ultimately find that, we can never be held truly responsible. While this does take something away from us, we gain so much more in this life through compassion and empathy for others. I conclude this work by suggesting that we gain an ethic through a non-ethic (or non-responsibility). In acknowledging the notion that the traditional sense of freedom, as a self-causal linkage, is incoherent – we gain a much greater sense of freedom and responsibility by extending empathy and love to those living amongst us.

4.1. The Basic Argument

The Basic Argument is a simple two-premise syllogism:

- (1) Nothing can be *causa sui* – nothing can be the cause of itself.
- (2) In order to be truly morally responsible for one’s actions one would have to be *causa sui*, at least in certain crucial mental respects.
- (3) Therefore, nothing can be truly morally responsible.¹³⁷

A helpful rewording of the first premise goes as follows: (1) I cannot bring out how I am physically, and more importantly, mentally. That is to say, to be truly responsible for a certain choice or action, I

¹³⁶ I reserve the right to add to this work. If I were to do so, I would need to consider much of the literature on responsibility and the potential side effects of what I stated above. What follows in this conclusion is not grounded in research but in my intuition based upon what has previously been claimed.

¹³⁷ Galen Strawson. “The Impossibility of Moral Responsibility.” *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 75, no. 1/2 (1994): 5–24. <http://www.jstor.org/stable/4320507>. 5.

would also have to be responsible for how I am mentally since choices or actions are the result of mental being. My physical or mental being, if it is to be considered “mine” or that which I am responsible for, must be brought about by me, and what is brought about by me is the result of how I am. But am I the sole reason, or cause, for how I am, physically or mentally?

Obviously, no one would argue that I brought about how I am physically – or that I made my own genetics or the skills I possess. Those were brought about and given to me by nature, so it would be ridiculous to take responsibility for them. At the same time, I cannot be upset by the fact that I do not possess certain skills, such as being a good basketball player. I may be able to develop a certain level of skill through practice, but there exists a ceiling on my ability to play basketball which, I am sure, is not very impressive. This lacking a certain skillset is also completely out of my control, and it would be ridiculous to be ridiculed by others for not possessing it. Physically speaking, most would agree that no one can lay claim to how they are and if they do, they are self-deluded.

What about the person who wakes up at 5AM every day to exercise? Surely, they are the way they are, physically, because they brought it about. On a conventional level, yes, but then we must ask: did they bring about their mental being possessing the idea of waking up at 5AM to exercise every day which was the cause of them waking up at 5AM to exercise? Of course not. There is no explanation for why this person desires what they do, whether it be a workout or a piece of chocolate cake. So, it still follows that no one is responsible for how they are physically.

I am also unable to discern how I can lay claim to my mental being since I am not the way I am due to some act of making myself that way. But it could be argued that I have a nature and this nature is the cause of how I am, mentally. For me to be responsible for how I am mentally, however, I would have had to intentionally bring about my nature. In order to bring about my own nature, I would have had to intentionally bring it about with some prior nature, a nature which I intentionally

brought about with an even prior nature, ad infinitum.¹³⁸ So again, due to an infinite regress in the necessity of some prior nature(s) causing my mental being, and since I am not the cause of how I am mentally, I am no more justified in claiming to be the cause of my mental being than my physical being.

4.2. Responsibility

What does the Basic Argument mean for responsibility? Strawson acknowledges that we cannot get around the natural sense of responsibility most human possess.¹³⁹ As he says,

When it comes to questions of responsibility, we tend to feel that we are somehow responsible for the way we are. Even more importantly, perhaps, we tend to feel that our explicit self-conscious awareness of ourselves as agents who are able to deliberate about what to do, in situations of choice, suffices to constitute us as morally responsible free agents in the strongest sense, whatever the conclusion of the Basic Argument.¹⁴⁰

Moral responsibility may be a useful fiction and, perhaps, ought to be preserved as such. It certainly feels like I am the reason for my current situation in life. To use my education as an example, I made good grades in high school and college which played a significant role in allowing me to enter this graduate program. Most would congratulate me on these achievements because they feel as though I am, at least partially, responsible for them. Likewise, it certainly feels like I am the thinker of my thoughts, or that I am the way I am, mentally, because I brought it about. But we should not take ultimate responsibility seriously if it is grounded only on what we *feel* to be the case – for we are extremely fallible. Therefore, since I am not the intentional cause of my internal goings-on, I am not responsible for them. Since I am not responsible for the cause of my internal goings-on, I am also not responsible for their effects – my choices and actions. So, the achievement of being accepted into a

¹³⁸ Galen Strawson. “The Impossibility of Moral Responsibility.” 14.

¹³⁹ I should mention that I am not a psychopath or arguing in favor of psychopathy. I do possess this feeling of responsibility which usually produces decent behavior and self-reflection, but I believe the feeling itself is ultimately groundless.

¹⁴⁰ Galen Strawson. “The Impossibility of Moral Responsibility.” 16.

graduate program feels like something I ought to take credit for and take pride in, but I know that this is a silly feeling to give true credence to. The most successful person is no more responsible for their success than the most impoverished. While this notion of non-responsibility does take away a sense of accomplishment from us, while reducing pride, it forces us to be more forgiving, loving, and empathetic to those less fortunate. Giving up responsibility, accomplishment, and pride in exchange for love and empathy is a tradeoff I am happy to make.

4.3. Reward and Punishment

Responsibility is ultimately an incoherent concept, but I believe our choices still matter. The choices we make every day – even if we do not technically *make* them – affect the material conditions which affect all of us. For this reason, we ought not adopt a nihilistic approach to ethics and moral conduct. We should, however, change the narrative from crediting or blaming people to acknowledging the reality of the situation granted under the Basic Argument. Society can decide to reward those *lucky* enough to possess a mental state which society approves of while rehabilitating or even punishing those who are *unlucky* to possess a mental state which society disapproves. I will argue that rewarding is naturally pernicious and should be repudiated for the negative effects it has on individuals, so long as “reward” is not defined as encouragement. I will make a distinction below between two kinds of rewards, one worth keeping and the other worthy of dismissal. Rehabilitation, on the other hand, cannot and should not be thrown out merely because responsibility is incoherent because some actions do negatively affect others which any society ought to limit.

When considering rewards, they can be divided into two kinds: encouragement and praiseworthiness. The former should be preserved in instances of valuable lesson learning, such as teaching children. Encouraging children with the assumed sense of absolute responsibility, even today, is not practiced. The opposite is the case. We tend not to truly praise or blame children because they do not know otherwise; due to their ignorance, they are not to be truly rewarded nor punished. These

tools are used, rather, to encourage children in the right direction in order to produce certain desired outcomes in behavior.

Praiseworthiness as a kind of reward occurs when students are put on the Dean's list or when an athlete breaks an important record, for example. We grant full responsibility to the agent as if their action was completely of their own doing. But if the Basic Argument is correct, this kind of reward makes no sense and only produces negative traits in people, such as pride, which only increase their sense of ego. Through praiseworthiness as a reward, a person can become quite narcissistic and unpleasant to encounter due to their close identification with these "accomplishments" as *theirs*. Praiseworthiness should never be defended since it results negatively for the individual and because this strong sense of responsibility, resulting in the praise, is grounded in a farce.

On the other side, I cannot endorse a view that totally disregards rehabilitation. Only an argument from practicality can justify rehabilitation for those unlucky¹⁴¹ individuals who happen to commit even the most heinous of crimes. For it is exactly as just to punish or reward people for their actions as it is to punish or reward them for the natural color of their hair or the shape of their faces.¹⁴² In the same way that most would grant that Donald Trump is lucky to be in the position he is in today – due to his fortunate situation – people who commit the worst of crimes are merely unlucky to possess their material situation or mental state, or both, that resulted in their crime. Their unfortunate actions do, however, have negative repercussions that affect those around them which we, as a society, should try to limit. For the sole purposes of empathy for these individuals and in order to curate a decent society, rehabilitation for these unlucky souls is necessary. We should recognize that their negative action was the result of unfortunate mental goings-on and not the result of their inner being

¹⁴¹ Yes, I mean unlucky in all cases. It even applies to the rapist, mass murderer, or fascist. It does not follow from this that we should not openly criticize those who commit these crimes. I am a huge proponent of criticizing these kinds of people and their terrible ideas/actions, but only due to the fact that criticism may produce effects that lead to positive change. Ultimately, these horrible people are not truly horrible, they are victims. I admit it becomes quite difficult to find a balance here.

¹⁴² Galen Strawson. "The Impossibility of Moral Responsibility." 16.

as a causal agent. We would never see a murderer as an evil doer but as a very unfortunate person who was the victim of poor mental states.

4.4. A Positive Future

A world with less ego and an acknowledgement that we are not the ultimate cause for our situations and actions results in a more loving and empathetic world. Most of society's ills result from hyper egoism and narcissism. When a parking spot is taken by another, for instance, we often get upset with the other person, but why? Because our egos identify with the parking spot and the only person to blame, the person to be held responsible, is the other. It was "mine" and now it is no longer. Upon reflecting that the parking spot has nothing to be identified with, the claim to the parking spot dissolves and there is no longer a need to be upset.

Through the acknowledgement of the incoherency of a self, free will, and responsibility, themes I have argued for throughout this work, a more loving and positive world will emerge. An ethic is developed from a non-ethic. By rejecting traditional responsibility, we become more ethical and empathetic towards others through our realization that they could not have done otherwise. While this non-ethical ethic takes away our sense of accomplishment, I still find it to be a worthy goal. Also, through the recognition that free will is a non-concept worthy of discussion, one gains a sense of freedom whereby I accept the course of events, inside or outside of one's being, as mere happenings. We have nothing to gain but empathy, love, and ultimate freedom in this arduous self-reflection, but since my urging the reader to adopt this argument for their own would be futile, due to my conviction to the argument itself, I can only hope that, upon considering it, the happenings in your mind render you incapable of rejecting it.

Bibliography

- Ayer, A. J. "Freedom and Necessity." In *Free Will*, edited by Gary Watson, 15-35. Oxford University Press, 1982.
- Baker, Lynne Rudder. "DOES NATURALISM REST ON A MISTAKE?" *American Philosophical Quarterly* 48, no. 2 (2011): 161–73. <http://www.jstor.org/stable/23025085>.
- Conze, Edward. "Buddhist Thought in India." Second printing. US, The University of Michigan Press, 1967,1970.
- Descartes, René. 1993. *Meditations on First Philosophy*. Translated by Donald A. Cress. 3rd ed. Cambridge, MA: Hackett publishing.
- Dumoulin, Heinrich. "The Person in Buddhism: Religious and Artistic Aspects." *Japanese Journal of Religious Studies* 11, no. 2/3 (1984): 143–67. <http://www.jstor.org/stable/30233321>.
- Garfield, Jay. "Engaging Buddhism: Why It Matters to Philosophy." New York, NY, Oxford University Press, 2015.
- Garfield, Jay. *Losing Ourselves: Learning to Live Without a Self*. 1st ed. Princeton, New Jersey: Princeton University Press, 2022.
- Gazzaniga, Michael S. *Who's in Charge? Free Will and the Science of the Brain*. 1st ed. Gifford Lectures 2009. New York, NY: HarperCollins, 2011.
- Goff, Philip. *Galileo's Error: Foundations for a New Science of Consciousness*. 1st ed. New York, NY: Vintage Books, 2019.
- Gombrich, Richard. "Theravāda Buddhism: A Social History from Ancient Benares to Modern Colombo." Second ed., New York, NY, Routledge & Kegan Paul Ltd, 1988, 2006.
- Hofstadter, Douglas R. 2008. *I Am a Strange Loop*. London, England: Basic Books.
- Kane, Robert. "Introduction: The contours of contemporary free will debates." In *The Oxford handbook of free will*. 2002.
- Levy, Ken M. *Free Will, Responsibility, and Crime: An Introduction*. First published. New York London: Routledge, Taylor & Francis Group, 2020.
- Lowe, E. J. *Personal Agency: The Metaphysics of Mind and Action*. First published. Oxford, New York: Oxford University Press, 2008.
- Lowe, E. J. "Selves: An Essay in Revisionary Metaphysics." *Analysis* 71, no. 3 (2011): 587–92. <http://www.jstor.org/stable/41237366>.

- Macintyre, A. C. "Determinism." In *Free Will and Determinism*, edited by Bernard Berofsky, 240-256. Harper and Roe Publishers, 1966.
- Metzinger, Thomas. *Being No One: The Self-Model Theory of Subjectivity*. Cambridge, MA: Bradford Books.
- Metzinger, Thomas. "Phenomenal transparency and cognitive self-reference." *Phenomenology and the Cognitive Sciences* 2 (2003): 353-393.
- Morris, Brian. "Anthropology of the Self." *American Ethnologist* 25, no. 4 (1998): 746-47. <http://www.jstor.org/stable/645864>.
- Rahula, Walpola. "What the Buddha Taught." Second ed., New York, NY, Grove Press, 1959, 1974.
- Russell, Paul, and Oisín Deery, eds. *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*. New York: Oxford University Press, 2013.
- Schopenhauer, Arthur, and Konstantin Kolenda. *Essay on the Freedom of the Will*. Dover Philosophical Classics. Mineola, N.Y: Dover Publications, 2005.
- Siderits, Mark. "Buddhism as Philosophy: An Introduction." Indianapolis, IN, Hackett Publishing Company, Inc., 2007.
- Strawson, Galen. *Selves: An Essay in Revisionary Metaphysics*. London, England: Oxford University Press. 2011.
- Strawson, Galen. "The Impossibility of Moral Responsibility." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 75, no. 1/2 (1994): 5-24. <http://www.jstor.org/stable/4320507>.
- Strawson, Galen. "What Is the Relation between an Experience, the Subject of the Experience, and the Content of the Experience?" *Philosophical Issues* 13 (2003): 279-315. <http://www.jstor.org/stable/3050535>.
- Tzu, Lao. *Tao Te Ching*. Trans. Victor H. Mair. New York, New York: Bantam Books Publishing, September 1990.
- Van Inwagen, Peter. *Metaphysics*. 3rd ed. Boulder, Colo: Westview, 2009.
- Van Inwagen, Peter. *Thinking about Free Will*. Cambridge, United Kingdom ; New York, NY, USA: Cambridge University Press, 2017.
- Watts, Alan. "The Way of Zen." Second ed., New York, Vintage Books, 1957, 1985.

Vita

Chris Rinderle was born in Metairie, Louisiana in 1997. He graduated from Mandeville High School in 2016. In 2020, he received his B.A. in political science from Louisiana Tech University. While studying at LA Tech, he developed a passion for philosophical enquiry. In 2021, he enrolled in the philosophy graduate program at Louisiana State University. In graduate school, he wrote his thesis on selfhood and free will. He is a candidate in 2023 for a master's degree in philosophy. Chris has been supported by his family and friends throughout his journey and is extremely grateful for their contributions and love along the way.