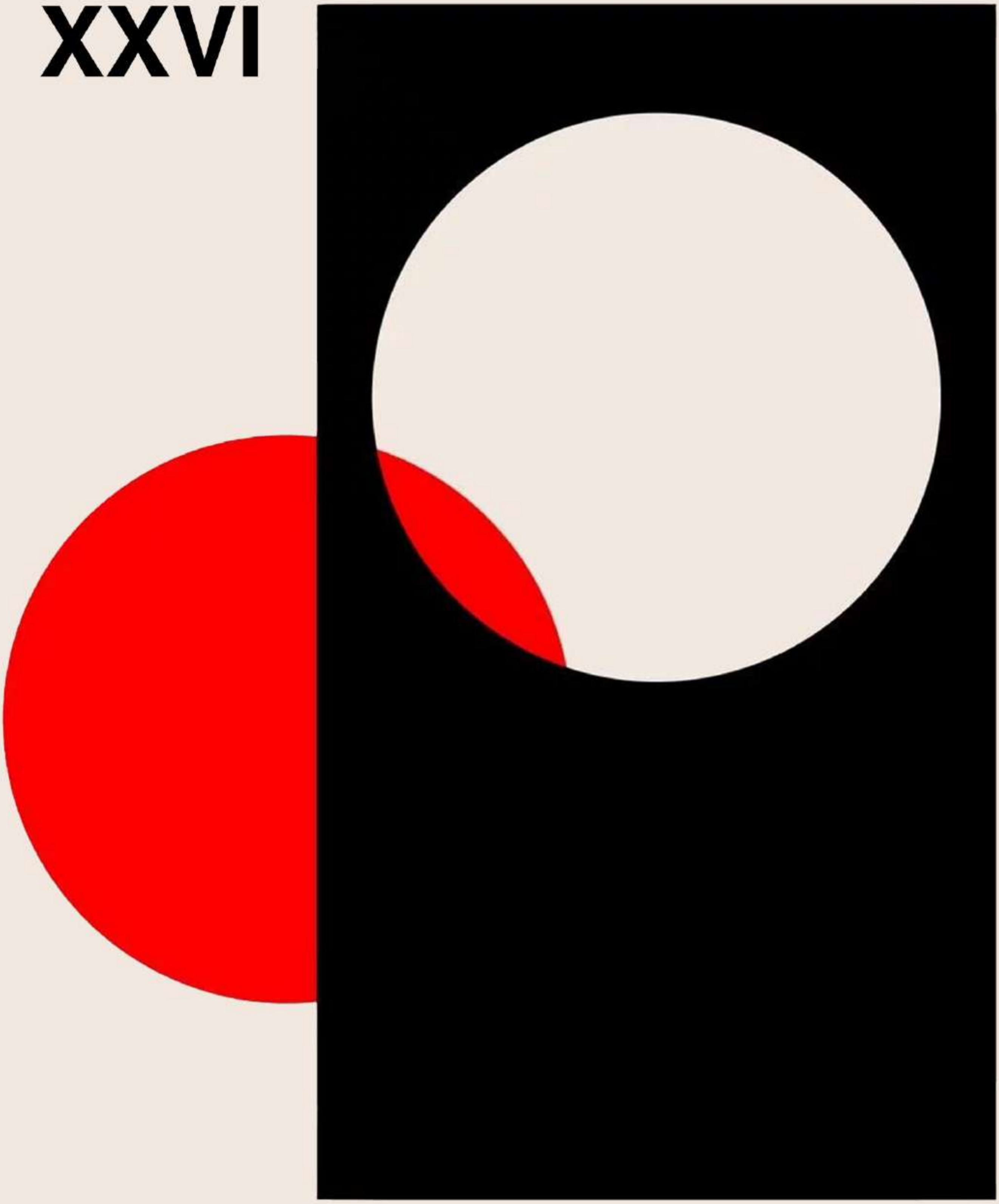
APORIα

XXVI



Aporia

Undergraduate Journal of the St Andrews Philosophy Society

VOLUME XXVI

Aporia is funded by the University of St Andrews Philosophy Society, which receives funds from the University of St Andrews Department of Philosophy, the University of St Andrews Students' Association, and independent benefactors.

Letter from the Editor

Dear Reader,

Before you is the twenty-sixth edition of Aporia, the University of St Andrews Philosophy Society's undergraduate journal.

Aporia exists to celebrate and promote outstanding undergraduate philosophy, providing a platform for students to share their ideas and engage meaningfully with academic research. This edition continues that tradition, bringing together work that reflects both the breadth and depth of philosophical inquiry undertaken by our contributors.

Producing a journal such as this is always a collaborative effort, one that truly takes a village (or, perhaps more fittingly, a town on the east coast of Scotland). This edition would not have been possible without the dedication, generosity, and hard work of everyone involved.

I owe particular thanks to our Deputy Editor-in-Chief, Hoochang Yi, whose thoughtful advice, careful editorial eye, and unwavering commitment have been invaluable throughout the process. I am equally grateful to every author and editor who devoted their time and effort to this journal. Your enthusiasm, diligence, and willingness to engage with one another's work are what make Aporia possible.

I hope you enjoy reading this edition as much as we enjoyed bringing it together, and that you find within its pages ideas that engage, inspire, and invite further reflection.

Yours faithfully,

Joseph Bradstreet
Editor-in-Chief, *Aporia*

Acknowledgements

Editors

Isaac Copley

Dingrui Deng

Christina Landys Herre

Olivia Kendall

Mihika Kulkarni

Zibi Laird

Zack Ledesma

Rowan Liddell

Yi Ma

Agata Mala

Claire Mizrahi

Sofia Mona

Daniel Playle

Wiktor Przyborowski

Rory Todd

Patricio Ávila

Wendy Zuo

Joseph Bradstreet
Editor-in-Chief

Hoochang Yi
Deputy Editor-in-Chief

Hanaa Khan
Technical Officer

Natalie Urusov
Cover Artist

Contents

Polishing the Causal Stream: A Pluralistic Reading of the *Platform Sutra* and a Parfitian critique of the traditional interpretation

Ng Kang Qi, *University of Cambridge*

The Platform Sutra of the Sixth Patriarch is a pivotal text in Chinese Buddhist philosophy, traditionally interpreted as establishing the orthodoxy of the Southern "Sudden" school of Chan Buddhism while relegating the Northern "Gradual" school to a provisional teaching. To understand this schism, one must understand the mechanics of Buddhist soteriology: enlightenment fundamentally requires liberation from the "psychological stream"—the continuous, causally linked flow of mental states, habits, and karmic dispositions that perpetuates the illusion of a permanent self. This paper argues that the traditional interpretation fails to properly address this causal reality by making a fatal logical leap from necessity to sufficiency. By claiming that a sudden, propositional insight into the emptiness of the self is entirely sufficient for enlightenment, the orthodoxy invites antinomianism and dogmatically dismisses the Northern method of cultivation. Drawing upon Derek Parfit's reductionist metaphysics and Gilbert Ryle's epistemology, I argue that the Northern school's gradual approach does not mistakenly reify a Cartesian ego but is rather strictly necessary to convert the propositional truth of emptiness (knowing-that) into the dispositional purity of enlightened living (knowing-how). Ultimately, this paper rehabilitates the *Platform Sutra* as espousing a methodological pluralism: true liberation requires both the sudden, conceptual destruction of the ego and the continuous, gradual cultivation of one's dispositions.

1. Introduction

In Chinese Buddhist philosophy, the *Platform Sutra of the Sixth Patriarch* is regarded as the definitive manifesto of the "Southern School" of Chan Buddhism. The "Southern School" represents an alleged doctrinal divergence from the "Northern School", with both branches being rooted in a historical Mahayana tradition descending from the shared "East Mountain" teachings of the Fifth Patriarch, Hongren.[1] The doctrines coincide in their understanding of enlightenment as a liberation from *Samsara*—the cycle of life and death—through a detachment from worldly desire. However, both schools diverge in their interpretations of what this detachment consists of, and how it can be achieved.

According to the traditional interpretation as established by Southern polemicists, the most salient soteriological differences between these schools can be represented in both their views of the nature of enlightenment and the method through which enlightenment is attained[2]:

According to the traditional interpretation, the doctrinal divergence illustrated above is the main rhetorical angle of the *Platform Sutra* with the text dramatising this schism through a pivotal narrative event: a succession contest wherein the learned head monk Shenxiu and the illiterate layman Huineng are framed as rhetorical opposites. While Shenxiu's submitted verse advocates for gradual character cultivation, Huineng's

[1] John R. McRae, *The Northern School and the Formation of Early Ch'an Buddhism* (Honolulu: University of Hawaii Press, 1986), 30–34. (Note: McRae thoroughly details Hongren's "East Mountain Teaching" as the shared progenitor of both schools).

[2] The traditional interpretation can be traced primarily to the rhetoric of Heze Shenhui (684–758), a disciple of Huineng. In his *Treatise Establishing the True and False according to the Southern School of Bodhidharma*, Shenhui explicitly attacked Shenxiu's lineage for advocating "gradual" cultivation. He condemned their practices—such as "freezing the mind to enter concentration" and "fixing the mind to view purity"—as dualistic errors that failed to apprehend the sudden, unconditioned nature of enlightenment. See John R. McRae, *Seeing through Zen* (Berkeley: University of California Press, 2003), 33–35.

Table 1.1: The Traditional Interpretation's Characterisation of the Soteriological Divide.

	Nature of Enlightenment	**Method of Attainment**
Northern School	A dispositional and recurrent state of purified moral and cognitive habits; the continuous maintenance of an unattached mind, free from worldly striving, which ultimately secures liberation from samsara and the cessation of suffering.	Gradual character cultivation, sustained meditation, and continuous moral vigilance to prevent moral defilement.
Southern School	An occurrent, transformative realisation of a propositional truth: the ontological emptiness of the self and the inherent purity of the mind, which directly results in liberation from samsara and the cessation of suffering.	Direct, unmediated intuitive awakening; a sudden paradigm shift that bypasses conceptual study.

verse characterises enlightenment as a sudden flash of insight into a reductionist ontology of the self. Within the narrative of the *Platform Sutra*, the outcome is decisive: the Fifth Patriarch acknowledges Huineng's verse as depicting the ultimate truth and crowns him the Sixth Patriarch.

Building upon this narrative outcome, the traditional interpretation asserts three conclusions. First, regarding the *nature* of enlightenment, it concludes that an occurrent realisation of Huineng's metaphysics is entirely sufficient for liberation. Second, it dismisses Shenxiu's views entirely, claiming his framework of gradual cultivation is an elementary metaphysical error that reifies a non-reductionist self. Third, regarding the *method* of enlightenment, it asserts that the only means through which liberation can be attained is by sudden insight.

The objective of this paper is to challenge these interpretative conclusions by arguing that the traditional interpretation overplays the extent to which the narrative victory of Huineng should represent a corresponding dismissal of Shenxiu. While the *Platform Sutra* places great emphasis on the importance of "Sudden Enlightenment" and a reductionist ontology of the self, I argue that Shenxiu's views on the nature of enlightenment as a disposition attained through gradual cultivation are compatible with such an ontology. Ultimately, I claim that Huineng's soteriology must necessarily contain some aspects of Shenxiu's teachings to be tenable. To demonstrate this, we must re-evaluate the implications of the traditional interpretation and show that the three main soteriological conclusions it draws are unjustified.

The paper proceeds as follows. In Section 2, I argue against the traditional interpretation's claim that propositional knowledge is sufficient for enlightenment by showing that this view leads to antinomianism, and thus cannot be an accurate reflection of the *Platform Sutra*. Consequently, the Northern view of dispositional purity is necessary to a tenable Chan soteriology. In Section 3, drawing on Derek Parfit, I demonstrate that the traditional interpretation misinterprets Shenxiu's ontology, and that the pedagogy of cultivating dispositions need not reify a non-reductionist self. In Section 4, drawing from Gilbert Ryle's thesis regarding the independence of propositional knowledge and dispositional skill, I argue against the traditional interpretation's view of the method of attaining enlightenment; because sudden insight is insufficient to build dispositional purity, the Shenxiu's teachings of gradual cultivation are strictly necessary to a tenable pedagogy of working towards enlightenment. Finally, in Section 5, I elaborate on why Huineng's propositional insight remains a necessary prerequisite for this gradual cultivation. This clarifies the narrative logic of the *Platform Sutra*, justifying the Fifth Patriarch's prioritisation of Huineng while exposing the root of the traditional interpretation's error: conflating necessity with sufficiency. Ultimately, I propose that the *Platform Sutra* should be viewed as espousing a pluralistic theory of enlightenment: one that defines the nature of enlightenment as both propositional insight and dispositional purity, and the method of attainment as both sudden awakening and gradual cultivation. This synthesis, corroborated by Guifeng Zongmi, seamlessly integrates the doctrines that the traditional interpretation falsely forces into a mutually exclusive "Northern" and "Southern" dichotomy.

2. The key verses of the Platform Sutra

The philosophical core of the *Platform Sutra* and the source of the tension addressed in this paper is found in the first section of the Sutra, where two verses are presented at the behest of the Fifth Patriarch—the great master of the Baolin monastery—as a means of finding a successor. According to the Sutra, the verse contest would allow the Fifth Patriarch to determine which disciple had apprehended the deeper truth of enlightenment and could therefore lead the monastery towards an appropriate soteriology.[3]

The first verse, presented by the learned head monk Shenxiu, is a representation of the "Northern" school of Chan:

> The body is a bodhi tree (*shēn shì pú tí shù*, 身是菩提树),
>
> The mind is like a bright mirror stand (*xīn rú míng jìng tái*, 心如明镜台),
>
> At all times we should diligently polish it (*shí shí qín fú shì*, 时时勤拂拭),
>
> and prevent dust from collecting (*wù shǐ rě chén āi*, 勿使惹尘埃).[4]

Within the narrative, Shenxiu's verse is treated as an incomplete apprehension of the true nature of enlightenment, and he is described by the Patriarch as "unable to see the self-nature," suggesting his framework is provisional rather than ultimate. After Shenxiu, the illiterate layman Huineng's verse is revealed:

> Bodhi originally has no tree (*pú tí běn wú shù*, 菩提本无树),
>
> The bright mirror also has no stand (*míng jìng yì fēi tái*, 明镜亦非台),
>
> Fundamentally, there is not a single thing (*běn lái wú yī wù*, 本来无一物),
>
> Where could dust arise? (*hé chù rě chén āi*, 何处惹尘埃)[5]

In contrast to Shenxiu, Huineng is celebrated by the Fifth Patriarch in private as having truly perceived the fundamental, unconditioned nature of the mind, thereby earning the secret transmission of the robe and Dharma, and the title of the Sixth Patriarch.

2.1. The traditional interpretation

According to the traditional interpretation initially published by Southern polemicist Shenhui and later accepted by standard Chan historiography, the reason for Huineng's rhetorical victory lies in how these verses map onto a fundamental ontological dispute. In Western analytic terms, this debate directly parallels Derek Parfit's distinction between non-reductionist and reductionist metaphysics of the self.

According to this view, Shenxiu's verse espouses a non-reductionist metaphysics of the self. Because the verse makes a distinction between the mind ("Mirror Stand") and its transient attachments or impurities ("Dust"), it is claimed that Shenxiu implies that the mind is a container that exists independently of its contents. In modern metaphysical terms, Shenxiu is accused of positing a "Further Fact" of identity: a substantive, unchanging self that exists over and above the stream of experience.

However, this substantive self is incompatible with the fundamental Buddhist doctrine of *Anicca* (impermanence)—because *Anicca* posits that all phenomena are in a state of constant, interdependent flux, the existence of a permanent, unchanging, and independent essence or soul must be rejected.[6] The traditional orthodoxy thus uses this rhetoric as a means of discrediting Shenxiu's view. According to this interpretation, because

[3] Philip B. Yampolsky, trans., *The Platform Sutra of the Sixth Patriarch* (New York: Columbia University Press, 1967), 128–129.
[4] Yampolsky, *The Platform Sutra*, 130
[5] Yampolsky, *The Platform Sutra*, 132.
[6] Mark Siderits, *Buddhism as Philosophy: An Introduction* (Indianapolis: Hackett Publishing, 2007), 32–34.

Shenxiu commits a fundamental metaphysical error, his practical teachings must be dismissed as a category mistake. If there is no permanent, underlying entity to accumulate defilement, the act of continuous purification is logically incoherent, since one cannot polish a non-existent 'mirror'. Therefore, the gradual method is tolerated within the Southern narrative only as a remedial crutch designed solely to prevent practitioners with 'dull faculties' from falling into grave moral error.

The traditional narrative then frames Huineng's verse as the welcome reductionist reply: by asserting "The mirror has no stand," Huineng denies the existence of some 'further fact' about the self, moving away from the problematic non-reductionism that Shenxiu apparently posits. Crucially, the traditional orthodoxy goes even further: it concludes that because Huineng is right about the ontology of the self, the nature of enlightenment is constituted merely by a propositional understanding of this ontology.

This conclusion is problematic for two primary reasons. First, it uses Huineng's rhetorical victory to justify a dangerous view of enlightenment that is likely to lead to antinomianism. It assumes that because Huineng holds the correct reductionist view of the self, an occurrent understanding of this metaphysical truth is inherently sufficient for enlightenment. Second, this narrative is extremely uncharitable to Shenxiu, prematurely reducing his gradualist pedagogy to an elementary metaphysical mistake.

2.2. **Semantic ambiguity and Huineng's ontology**

To demonstrate why treating this propositional knowledge as sufficient is fatal to Chan soteriology and precisely how it invites the threat of antinomianism, we must first clarify the exact reality Huineng is describing. This requires a close analysis of his main metaphysical claim: "fundamentally, there is not a single thing" (*běn lái wú yī wù*, 本来无一物). The main challenge in understanding this assertion lies in the key term "a single thing", which is semantically overloaded. Depending on how we interpret this phrase, the verse generates three entirely distinct and mutually exclusive ontological claims:

Reading 1: The Nihilistic Reading ("Single thing"= Anything)

If "a single thing" is read as 'anything', then the statement denies the existence of any objects (variables) whatsoever within the domain.

$\neg\exists x(x = x)$

Reading 2: The Anti-Monistic Reading ("Single thing"= One thing)

If the word "single" is read instead as a numerical quantifier designating one, then the statement denies that there is exactly one object within the domain (meaning there is either zero or more than one object).

$\neg(\exists x(x = x) \land \forall x \forall y(x = y))$

Reading 3: The Anti-Pluralistic Reading ("Single thing"= Distinct thing)

If "a single thing" is read as a substantive, independent entity possessing inherent existence, then the statement denies the existence of boundaries, or of distinctness between objects. It asserts that there are no distinct things in the universe.

$\neg\exists x \exists y(x \neq y)$

Which of these readings is the most plausible? We can immediately reject Reading 1 (Nihilism), as it is directly incompatible with the rest of Huineng's soteriology. An empty universe precludes the possibility of the experiencing mind and sudden awakening, which are the foundations of Huineng's brand of Chan Buddhism. Furthermore, elsewhere in the *Platform Sutra*, Huineng explicitly states:

> "Emptiness includes the sun, moon, stars, and planets, the great earth, mountains and rivers, all trees and grasses, bad men and good men, bad things and good things, heaven and hell; they are all in the midst of emptiness."[7]

[7]Yampolsky, *The Platform Sutra*, 146.

This passage serves a crucial philosophical purpose: it definitively populates the ontology. Because the nature of emptiness encompasses the phenomenal world, the domain of discourse cannot be empty, and the nihilistic reading must be rejected. Furthermore, the rhetorical structure of this passage eliminates Reading 2 (Anti-Monism). Huineng deliberately lists the vast pluralities and dualities of the phenomenal world, including the good and bad, heaven and hell, only to immediately collapse them into the singular concept of "emptiness." From this we can surmise that he is not advocating for pluralism but trying to dismantle it.

Therefore, we are left with Reading 3. When Huineng asserts "there is not a single thing," he is using the formulation $\neg\exists x\exists y(x \neq y)$ to make a mereological claim that denies the existence of boundaries and distinctness. In standard logic, the statement "no two things are distinct" is logically equivalent to the universal affirmation of identity: $\forall x\forall y(x = y)$ (everything is identical to everything else). When we combine this anti-pluralistic premise ($\forall x\forall y(x = y)$) with our rejection of the nihilistic reading, the conclusion is that Huineng is asserting that there exists *exactly one* thing $\exists x(x = x \wedge \forall y(y = x))$.

This conclusion that reality is an undifferentiated whole is grounded in the broader textual tradition of Chan Buddhism. In the *Platform Sutra* and wider Mahayana literature, this singular underlying reality is explicitly identified as the inherently pure Buddha-nature (*fó xìng*, 佛性). The denial of pluralistic boundaries is the ontological equivalent of the Chan doctrine of non-plurality, which posits that there is no substantive metaphysical division between ordinary sentient beings and the enlightened Buddha.

2.3. **Monism and Antinomianism**

This brings us to the core tension. While Huineng's metaphysical monism aligns with the broader Mahayana doctrine of Buddha-nature, the traditional interpretation weaponises this ontology to make an unjustified conclusion regarding the nature of enlightenment. The orthodoxy assumes that because reality is an inherently pure, unpartitioned whole, the mere propositional realisation of this metaphysical truth is sufficient to *constitute* enlightenment.

However, this cannot be the case. Treating propositional knowledge as the sole requirement for enlightenment invariably leads to antinomianism: the dangerous theological doctrine that moral laws, ethical habituation, and behavioural precepts are no longer binding. If enlightenment is reduced to a sudden flash of conceptual insight into the underlying purity of reality, it implies there is nothing left to do, and that the everyday Buddhist practice of continuous moral cultivation is rendered useless. Under this view, one could grasp the monistic ontology while continuing to engage in harmful behaviour, justifying their actions under the guise that "all things are already the Buddha-nature."

This antinomian outcome is unacceptable for two reasons. First, it is practically disastrous for the Chan tradition, which is premised on the idea that Buddhist practitioners must aim to alleviate real-world suffering. Second, it would be unable to account for other sections of the *Platform Sutra*.

For example, in Chapter 6, Huineng explicitly confers the "Formless Repentance" (*wú xiàng chàn huǐ*, 无相忏悔) upon his disciples, commanding them to actively repent for past and future sins arising from ignorance and deceit. Crucially, he does not merely ask them to realise the abstract emptiness of these sins; he explicitly demands that they "awaken to [their] own errors and **reform** [their] past evil."[8]

This pedagogical mandate refutes the traditional interpretation's claim that a sudden flash of propositional understanding is sufficient to constitute enlightenment. The act of "[reforming] past evil" inherently requires a functional, temporal separation between past karmic defilement and present moral action. It demands a continuous ongoing change in psychological habituation and behaviour, beyond a change in metaphysical understanding. If a practitioner only possesses the propositional knowledge that everything is one with the Buddha-nature, yet remains acting in a manner that brings more suffering to the world, they cannot be enlightened according to the doctrines of Chan. Therefore, the nature of enlightenment must necessarily include a dispositional purity alongside propositional insight.

If this dispositional purity is practically necessary, as Huineng's own repentance rituals demonstrate, then the dismissal of Shenxiu's verse is premature. Shenxiu's instruction to "diligently polish" the mirror represents the exact conceptual framework required to execute the behavioural changes that the traditional interpretation abandons. To save Chan practice from the complacency of the absolute, a more charitable

[8] Yampolsky, *The Platform Sutra*, 143–144.

interpretation of Shenxiu's gradualism is in order. We must reinterpret his verse not as an incorrect metaphysical claim, but as a practical and behavioural framework for maintaining dispositional purity.

This rehabilitation, however, presents a significant philosophical hurdle: how can this be done without risking a problematic, non-reductionist ontology of the self that Shenxiu seems to imply? To resolve this tension, we require a philosophical framework capable of distinguishing between the metaphysical subject (which Huineng rightly targets) and the psychological continuum (which the traditional reading inadvertently discards).

3. The Parfitian Solution

Derek Parfit's reductionist metaphysics in *Reasons and Persons* provides precisely the analytical vocabulary needed to resolve this tension. Parfit allows us to strictly separate the question of identity (the metaphysical subject) from the question of survival (the psychological continuum), revealing that the "Sudden" and "Gradual" approaches are not mutually exclusive, but rather address two entirely distinct ontological categories.[9]

To map Parfit's framework onto the Chan debate, we must first understand his distinction between two rival views of persons:

Non-reductionism (The Cartesian View): this view holds that personal identity is ensured by a "Further Fact", or a separately existing entity, such as a Cartesian Ego or a Soul, that remains distinct from the brain and body.

Reductionism (The Bundle View): Parfit, aligning with Hume and orthodox Buddhism, holds that a person is not a separate underlying entity, but a composite bundle of interrelated mental and physical events. For the Reductionist, personal identity is constituted solely by Relation R: psychological connectedness and continuity (memory, character, and causal links).[10]

Applying these categories provides the necessary lens to interpret Shenxiu without uncharitably reducing his view to non-reductionism. Rather than thinking of Shenxiu's verse as reifying a "Further Fact" about the self, we can interpret his verse as a necessary pedagogy regarding Relation R.

3.1. Identity vs. Survival: The Independence of Relation R

To understand why this pedagogy regarding Relation R remains practically necessary, we must utilise Parfit's distinction between identity and survival. Parfit argues that what matters in personal existence is not the persistence of a "Further Fact" (identity), but the preservation of relation R (survival).[11] The traditional interpretation correctly identifies Huineng's verse as a metaphysical attack on the "Further Fact." Huineng's assertion that there is "not a single thing" simultaneously asserts a Buddhist monism and attacks the idea of the Cartesian Ego, aligning with the Buddhist doctrine of *anatman* (no-self). However, the dismissal of Shenxiu leads directly to antinomianism because it overlooks a crucial reality: the negation of the "Further Fact" does not negate Relation R. Even in a strictly reductionist universe where there is no permanent self to 'own' experiences, the causal stream of psychological continuity continues to flow. Mental state M1 (e.g., a moment of anger) remains causally connected to mental state M2 (e.g., a disposition toward violence) via memory and psychophysical causality. This causal chain exists independently of the ontology of the "Further Fact." Ignoring the necessity of regulating Relation R is what leads the traditional interpretation towards antinomianism.

Therefore, when Shenxiu commands practitioners to "diligently polish" the mirror, we should not read this as the reification of a non-reductionist view of the self. Rather, it is the necessary pedagogical mechanism for managing Relation R, where Shenxiu is addressing the causal reality of the psychological continuum. Just because the "polisher" and the "mirror" lack an independent, substantial existence does not mean the causal links between immoral thoughts and harmful actions cease to operate. Gradual cultivation

[9] Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), 215.

[10] Parfit, *Reasons and Persons*, 216–217. See also David Hume, *A Treatise of Human Nature*, 1.4.6.

[11] Parfit, *Reasons and Persons*, 255. Parfit argues that "Identity is not what matters in survival."

is the essential practice of purifying the causal stream, and it is the only way to make sense of Huineng's command for "formless repentance" and avoid antinomianism.

One might object that the syntactical structure of Shenxiu's verse contradicts this interpretation and commits him to a non-reductionist ontology. By utilising nouns like "mirror stand" and "dust," in a subject-copula-predicate form ("The body **is** a bodhi tree"), he appears to be hypostatising the mind, treating it as a substantive container. However, a closer reading of the original Chinese text reveals this to be a functional analogy rather than a strict identity statement. While his first line uses the copula to assert "The body *is* [*shì*, 是] a bodhi tree," his second line deliberately shifts to a simile: "The mind is *like* [*rú*, 如] a bright mirror stand." From this, we can grant that Shenxiu is not making a non-reductionist claim about the self but is instead utilising an analogy to highlight the mind's capacity to reflect reality clearly or be obscured by passing phenomena.

Furthermore, the syntactical phrasing is likely a byproduct of pedagogical necessity. Language is inherently dualistic; the very syntax of a behavioural imperative requires a grammatical subject and object. To describe the stewardship of the mental stream without using the convenient shorthand of a "mirror" and "dust" would require a cumbersome, technical circumlocution that would be pedagogically inert and perhaps even more opaque to the average practitioner.

Thus, Shenxiu adopts a pragmatic "intentional stance." Just as a modern biologist might say "the gene wants to replicate" without literally ascribing conscious desire to a molecule, Shenxiu commands his monks to "polish the mirror" without literally ascribing a permanent Cartesian soul to the practitioner. He is utilising conventional language to point toward necessary practical action. Once we strip away this grammatical hypostatisation, we can accurately reconstruct the underlying logical form of Shenxiu's verse. When he commands his practitioners to "diligently polish" and "prevent dust from collecting," he is fundamentally asserting that there is a causal sequence of mental states that requires continuous regulation.

In Parfitian terms, we can map his metaphor directly onto the mechanics of Relation R. The "dust" is not a physical property attached to an underlying soul, but rather the causally salient factors within the psychological continuum, such as fleeting thoughts of greed, anger, or ignorance. If the person at time T1 allows this "dust" to settle by reinforcing a habit of ignorance, the person at time T2 will inherit those dispositions via Relation R. "Polishing," therefore, represents the precept of regulating a particular causal stream so that future psychological states develop the correct moral dispositions. The traditional interpretation, by fixating exclusively on Huineng's negation of the "Mirror," throws out the baby (Relation R) with the bathwater (the "Further Fact"). It leaves the practitioner with a metaphysical theory, but without a practical method to address the continuity of their conditioned habits. By utilising Parfit, we successfully vindicate Shenxiu's verse as necessary framework for understanding the dispositional nature of enlightenment.

4. The epistemology of enlightenment with lessons from Ryle

Having established that the nature of enlightenment inherently requires dispositional purity (the purification of Relation R), we must now turn to the method of attainment. A defender of the traditional interpretation might concede that the nature of enlightenment necessarily requires a dispositional aspect, but resist the idea that the method of attaining enlightenment requires gradual cultivation by contending that a sufficiently profound, sudden propositional insight would automatically generate these dispositions.

However, drawing upon the epistemology of Ryle reveals why this argument must fail. Ryle would call such a view the "intellectualist legend": the false assumption that intelligent, skilful action is merely a derivative of theoretical knowledge, or that to "know" a truth is to automatically be able to act upon it. In *The Concept of Mind*, Ryle argues against the view that performance is merely the execution of a recognised rule. He distinguishes between two cognitive categories:

Knowing-That (Propositional Knowledge): the possession of information or theoretical truth (e.g., knowing *that* a bicycle functions via gyroscopic force).

Knowing-How (Dispositional Skill): the capacity to perform a task or navigate a domain (e.g., knowing *how* to ride a bicycle).[12]

[12] Gilbert Ryle, *The Concept of Mind* (London: Hutchinson, 1949), 27–32.

Crucially, Ryle argues that these two categories are fundamentally distinct and irreducible: one can possess propositional knowledge in a subject matter but lack the corresponding dispositional skills, and vice versa. Applying this to the Chan debate reveals a crucial point. If one accepts that the nature of enlightenment requires the conditioning of one's causal stream (which the preceding sections have established is necessary to avoid antinomianism), then mere propositional knowledge of emptiness will not automatically manifest this dispositional state. Huineng and Shenxiu thus address two distinct epistemological categories of the same soteriological goal. Huineng's verse concerns the propositional knowledge of the nature of the self in relation to the universe, while Shenxiu is espousing the pedagogy of the dispositional skill of living according to this ontological reality. The error of the methodological monism promoted by the traditional interpretation is the belief that Knowing-That (the sudden realisation of emptiness) automatically generates Knowing-How (the dispositional capacity to live without attachment). This is demonstrably false. A person who wishes to be healthy may fully grasp and believe the proposition that "smoking is unhealthy" (Knowing-That) while completely lacking the dispositional capacity to actually quit (Knowing-How).

Similarly, a practitioner may realise that the ego is an illusion yet still react to the world with ingrained habits of anger and greed because their psychological stream—or the causal momentum of Relation R—has not been properly conditioned. Thus, Shenxiu's gradual method is the necessary bridge between metaphysical theory and ethical reality. Only the gradual re-habituation of the causal stream can remove the "dust" of conflicting desires. The Northern method of gradual cultivation is therefore necessary to attain the dispositional nature of enlightenment.

5. The Pluralistic Synthesis: Narrative Logic and the Necessity of Insight

If Shenxiu's method of gradual cultivation is strictly necessary to prevent antinomianism and achieve enlightenment, one might ask: why does the *Platform Sutra* dramatise his apparent defeat? Why did the Fifth Patriarch prioritize Huineng's verse, award him the title of the Sixth Patriarch, and seemingly dismiss Shenxiu?

The answer lies in the epistemological sequence of enlightenment. While Shenxiu provides the necessary method (Knowing-How), Huineng provides the strictly necessary metaphysical target (Knowing-That). If a practitioner engages in moral cultivation without first realising the propositional truth of emptiness, their practice is spiritually blind because it lacks a definitive soteriological goal. Without the orienting framework of the reductionist-monist metaphysics—which explicitly directs the practitioner toward absolute detachment from worldly desire—the advice to cultivate one's moral character degrades into mere conventional morality. This is why without Huineng's ontological assertion, Shenxiu's mandate to "diligently polish" the mind is incomplete: it reads as generic ethical advice on how to be a "good" person within the illusion, rather than a radical method for escaping it. Because such practitioners are still attached to the concept of a persistent self achieving a worldly goal, they remain fundamentally trapped within *samsara*.

This clarifies the narrative logic of the *Platform Sutra* and justifies the Fifth Patriarch's prioritisation of Huineng. Huineng's verse was chosen not because it represented the entirety of enlightenment, but because it represented the indispensable metaphysical target. His sudden, propositional destruction of the "Mirror Stand" acts as a vital safeguard. It ensures that when a practitioner engages in cultivation, they are regulating a causal stream (Relation R) rather than feeding an ego.

It is here that we can expose the root of the traditional interpretation's error. The orthodox commentators saw the Fifth Patriarch rightfully crowning Huineng because his ontological insight was strictly necessary to transcend conventional morality. However, they made a fatal logical leap: they conflated this necessity with sufficiency. They weaponised Huineng's narrative victory to assume that because sudden insight was the mandatory first step, the subsequent gradual cultivation of the causal stream was entirely obsolete.

However, the text of the *Platform Sutra* itself resists this claim of sufficiency, demonstrating that once the metaphysical target is acquired, Shenxiu's gradualist ideas are quietly adopted. As discussed in Section 2, Huineng actively confers the "Formless Repentance" upon his disciples, demanding a continuous, ongoing vigilance to reform past and future karmic defilements. Furthermore, after receiving "sudden insight," the Sutra depicts Huineng living in obscurity among hunters for over a decade as a means of cultivating his dispositions.

Ultimately, the *Platform Sutra* should not be read as the victory of one mutually exclusive set of doctrines over another, but as espousing a pluralistic theory of enlightenment. The nature of enlightenment must be defined as both propositional insight into emptiness and the dispositional purity to live according to it. Consequently, the method of attainment strictly requires both sudden awakening (to destroy the ego and secure the metaphysical target) and gradual cultivation (to purify the causal stream).

5.1. **Historical Precedents for Methodological Pluralism**

The necessity of methodological pluralism was explicitly recognised by the Tang Dynasty scholar-monk Guifeng Zongmi. As a patriarch of both the Huayan school and the Heze lineage of Southern Chan, Zongmi was uniquely positioned to diagnose the sectarian failures of his era. He observed firsthand that the strict methodological monism of the radical Southern School was actively leading to the exact antinomianism warned about in this paper.

To resolve this crisis, Zongmi argued for a definitive synthesis of Huineng and Shenxiu's views, formalising the doctrine of "Sudden Awakening followed by Gradual Cultivation" (dùn wù jiàn xiū, 顿悟渐修). Zongmi recognised that sudden insight into the ontological reality of the self (Huineng's monism) is not the culmination of the Buddhist path, but merely its prerequisite.

To illustrate how these two seemingly opposed methodologies must seamlessly integrate, Zongmi utilised a series of powerful phenomenological metaphors in his *Preface to the Collected Writings on the Source of Chan*. His most famous case study compares enlightenment to the birth and maturation of a child:

> "It is like a child being born all at once. At the moment of birth, its limbs and faculties are completely formed, no different from those of an adult. However, its physical strength is not yet mature. It requires many months and years of gradual nourishment before it can eventually become a fully functioning adult."[13]

This metaphor directly maps onto our previous analysis. The "sudden birth" represents Huineng's propositional realisation of the Metaphysical Subject: the sudden, occurrent understanding that one's fundamental nature is entirely identical to the Buddha-nature. Ontologically, the infant is fully human, just as the practitioner is fully Buddha.

However, Zongmi explicitly warns that this ontological completeness does not translate into immediate practical capacity. The "gradual nourishment" represents Shenxiu's pedagogy. Even after the "Further Fact" of the self is destroyed conceptually, the practitioner's psychological continuum including their behavioural dispositions and cognitive habits remains intact. Zongmi utilises a second metaphor to explain this lingering causality:

> "It is like the wind blowing across a pond. The wind may suddenly cease, but the waves subside only gradually. The wind of ignorance may suddenly be extinguished by awakening, but the waves of habitual defilements can only be eliminated through gradual cultivation."[14]

In Parfitian terms, Zongmi is illustrating the independence of Relation R. The sudden cessation of the "wind" is the conceptual destruction of the Cartesian Ego. But the "waves," which represent the causally linked mental states of anger, greed or selfishness, will continue to linger. If practitioners do not try to actively calm the waves after the cessation of the wind, they fall victim to antinomianism.

By collapsing the false narrative of mutual exclusivity, we can therefore conclude that a Chan conception of enlightenment is neither exclusively an occurrent realisation nor merely a dispositional state; it is the necessary, pluralistic integration of both. It requires both the sudden, conceptual destruction of the "Self" and the gradual regulation of one's causal stream.

[13] Peter N. Gregory, *Tsung-mi and the Sinification of Buddhism* (Princeton: Princeton University Press, 1991), 228.

[14] Gregory, *Tsung-mi and the Sinification of Buddhism*, 228.

6. Conclusion

The traditional interpretation of the *Platform Sutra* commits a fundamental categorical error: it conflates the necessity of Huineng's metaphysical insight with soteriological sufficiency. By weaponising his narrative victory, the orthodoxy dismisses Shenxiu's teachings as an elementary ontological mistake, thereby enforcing a dogmatic methodological monism. As this paper has demonstrated, this orthodox view is philosophically untenable. Regarding the nature of enlightenment, treating propositional knowledge of emptiness as sufficient invariably invites antinomianism. Consequently, dispositional purity must be recognised as a necessary constituent of enlightenment. Furthermore, through the lens of Parfitian reductionism, Shenxiu's verse is exonerated from accusations of non-reductionism; his mandate to "polish the mirror" is not the reification of a Cartesian Ego, but the necessary pedagogical stewardship of the psychological continuum (Relation R).

Regarding the method of enlightenment, Ryle's epistemology reveals that sudden, propositional insight (Knowing-That) cannot independently generate dispositional purity (Knowing-How). Therefore, Shenxiu's method of gradual cultivation remains necessary to convert propositional understanding into ethical reality. Ultimately, a tenable Chan soteriology based on the *Platform Sutra* requires the pluralistic synthesis historically proposed by Guifeng Zongmi. The *Platform Sutra* should not be read as a contest between mutually exclusive doctrines, but as an exposition of a plurality of doctrines required for enlightenment. By rejecting the traditional interpretation's false dichotomy, this analysis reveals that Chan Buddhism is not merely a mysticism of sudden insight, but a sophisticated framework of reductionist virtue ethics.

Bibliography

Gregory, Peter N. *Tsung-mi and the Sinification of Buddhism*. Princeton: Princeton University Press, 1991.

Hume, David. *A Treatise of Human Nature*. Oxford: Clarendon Press, 1888.

McRae, John R. *The Northern School and the Formation of Early Ch'an Buddhism*. Honolulu: University of Hawaii Press, 1986.

McRae, John R. *Seeing through Zen: Encounter, Transformation, and Genealogy in Chinese Chan Buddhism*. Berkeley: University of California Press, 2003.

Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.

Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson, 1949.

Siderits, Mark. *Buddhism as Philosophy: An Introduction*. Indianapolis: Hackett Publishing, 2007.

Yampolsky, Philip B., trans. *The Platform Sutra of the Sixth Patriarch*. New York: Columbia University Press, 1967.

Zongmi, Guifeng. *Chanyuan zhuquanji duxu* [Preface to the Collected Writings on the Source of Chan]. Translated and quoted in Peter N. Gregory, *Tsung-mi and the Sinification of Buddhism*. Princeton: Princeton University Press, 1991.

Culpability Without Ability: A Hard Determinist Proposal

Yuyang Ma, *Duke University*

This essay defends a coherent form of hard determinism against the charge that, if agents are never able to do otherwise, our practices of moral advice and blame lose their ground. I formulate this pressure as the Culpability Dilemma. On the one hand, powerful incompatibilist arguments deny any genuine ability to do otherwise, whether determinism or indeterminism is true; on the other hand, ordinary moral life still trades in judgments like 'you should have...' and 'it's your fault...', judgments that seem to presuppose exactly that ability. To get out of this dilemma, I challenge van Inwagen's Moral Advice Argument by proposing that second-person 'ought' claims are not primarily reports of an antecedent normative fact that presupposes present categorical ability but are practical interventions aimed at influencing what an agent will do. On a deterministic picture, correct advice need only satisfy a weaker condition: if the advice were given and taken up, it would causally issue in the advised action. I then revisit Broad's 'ought implies can' by distinguishing comparative, deliberative, and retrospective uses of 'ought'. These revisions allow the hard determinist to preserve the intelligibility of advice and much of the grammar of fault and blame while rejecting the claim that culpability requires the ability to do otherwise. The essential move is the reinterpretation of Broad's retrospective "ought": showing that past-tense moral verdicts reduce to counterfactual claims about correct prior guidance and therefore do not presuppose categorical ability.

1. **Introduction**

We often blame people for certain states of affairs. We say a careless driver is at fault for an accident, or that a dishonest politician is to blame for a scandal. When we make such judgments, a trivial underlying assumption seems to be that: there must be an occasion on which the person we blame was able to do otherwise. We can more formally capture this thought down in the following Culpability Thesis:

> If a bad state of affairs S is truly the fault of a person x, then there is some earlier time t and some course of action A such that:
>
> (1) At t, x was choosing between two or more incompatible courses of action, one of which is A;
>
> (2) at t, x ought to perform A instead of what x actually did;
>
> (3) if x had performed A, S would not have occurred;
>
> (4) and at t, x was able to perform A.

The Culpability Thesis interacts in a crucial way with the following three familiar philosophical theses:

> Ability Thesis: On at least some occasions, when an agent is deciding between two or more incompatible courses of action, the agent is able to perform each of them.
>
> Thesis of Determinism: The past and the laws of nature determine a unique future.
>
> Thesis of Indeterminism: The past and the laws of nature do not determine a unique future.

Both determinists and indeterminists can offer powerful arguments that their respective views rule out the kind of ability posited by the Ability Thesis. For the hard determinist (one who holds that the Thesis of Determinism is true and that it is incompatible with the Ability Thesis), Peter van Inwagen's Consequence Argument purports to show that if the past and the laws of nature jointly determine every future event, then no agent is ever able to do otherwise than what she actually does; the ability to do otherwise would require either changing the past or violating a law of nature, neither of which is within anyone's power. For the hard indeterminist (one who holds that the Thesis of Indeterminism is true and that it is incompatible with the Ability Thesis), the Mind Argument powerfully shows that indeterminism fares no better: if an agent's action is not determined by her prior mental states, then whether she performs it is a matter of luck or chance, and an action that occurs by chance is not one the agent has the ability to perform or to refrain from performing in the relevant sense.[1] Given that determinism and indeterminism are mutually exclusive and jointly exhaustive, if neither is compatible with the Ability Thesis, it follows that the Ability Thesis is false. Yet the problem remains that in ordinary life, we continue to speak and judge as if people are sometimes at fault. When we think about Nazi officers who oversaw mass killings, we say: "These crimes were their fault" or "they ought not to have done this," which, together with the Culpability Thesis, appears to require exactly the sort of ability that both sides have just ruled out.

C. D. Broad found himself in a similar situation in his inaugural lecture at Cambridge University, *Determinism, Indeterminism, and Libertarianism*, where he argued a libertarian view (one that affirms the Ability Thesis) is hard to maintain given the incompatibilist arguments. At this point, people like Broad, who deny the Ability Thesis on the strength of these incompatibilist arguments face a dilemma with two primary paths forward: (i) defend that "nothing is ever anyone's fault" is not merely counterintuitive but actually true, and then explain the nature of this powerful illusion that sustains ordinary fault and blame attributions or (ii) argue that the Culpability Thesis is false, i.e., moral responsibility does not require any ability to do otherwise. This dilemma is what I will call the Culpability Dilemma. Broad still believes the Culpability Thesis to be self-evident. His reasoning is that whenever we make a retrospective judgment of obligation ("you ought not to have done X"), we universally admit as relevant the defense "I could not help doing X" or "I could not possibly have done Y." The fact that this is treated as a legitimate rebuttal shows, for Broad, that there is some sense of "could" in which "ought" and "ought not" entail "could"; an action is, in his words, "obligable if and only if it is, in a certain sense, substitutable."[2] So (ii) is not really an option for him.

Peter van Inwagen also builds his own case to defend the Culpability Thesis and block path (ii) in *Free Will and Moral Responsibility: Once More Unto the Breach*, by connecting the concepts of fault, moral advice, and the ability to act, in what I shall call the Moral Advice Argument.[3] His argument is built on the core claim that: if "you ought to do X" is the correct answer to the question "What should I do?", then the agent is able to do X.

But does the Moral Advice Argument successfully establish the link between culpability and ability or "ought" and "can", such that one cannot coherently give advice or make fault and blame attributions if they reject the Ability Thesis? This is the question I wish to explore in what follows. Van Inwagen's core claim about advice, as I will show, is a specific application of the Broadian principle: that "ought" and "ought not" can apply to an action only if the agent could have done otherwise; so to establish culpability without ability requires showing that the retrospective "ought" (the "ought" of blame and moral verdict) does not commit us to the categorical sense that Broad takes to be essential.

My strategy proceeds in stages. I will first (§2) offer a way around van Inwagen's Moral Advice Argument by reconceiving advice as a causal intervention, which weakens the link between correct advice and present categorical ability. I will then (§3) turn to the deeper Broadian principle itself and argue that once we distinguish the comparative, deliberative, and retrospective uses of "ought," the retrospective "ought" can be reinterpreted as a counterfactual claim about correct prior guidance rather than as a marker of cate-

[1] Peter van Inwagen, *An Essay on Free Will*, rev. ed. (Oxford: Oxford University Press, 1986); and Peter van Inwagen, *Thinking about Free Will* (Cambridge: Cambridge University Press, 2017), chap. 7, 90–110, https://doi.org/10.1017/9781316711101.

[2] C. D. Broad, "Determinism, Indeterminism, and Libertarianism," in *Ethics and the History of Philosophy: Selected Essays* (London: Routledge, 2014), 197, https://doi.org/10.4324/9781315823058.

[3] Peter van Inwagen, "Free Will and Moral Responsibility: Once More unto the Breach," Belgrade Philosophical Annual 38, no. 1 (2025): 13, https://doi.org/10.5937/BPA2501007I.

gorical ability. This reinterpretation is the essential move to show that the hard determinist can retain the full grammar of "you ought to have done otherwise" without accepting Broad's claim that such judgments presuppose the ability to have done otherwise. Building on these two strands, I will (§4) demonstrate how the hard determinist can make sense of fault and blame ascriptions and dissolve the Culpability Dilemma via path (ii).

2. **The Moral Advice Argument**

Van Inwagen's Moral Advice Argument seeks to demonstrate that assigning fault for a state of affairs logically implies that the agent in question was, at some point, able to act otherwise. The core of the argument can be formulated in three steps. Step one is a simple and reasonable principle of fault, which states that:

> "If a certain state of affairs is someone's fault, there is something that person did not do and should have done (or did and should not have done)."[4]

For example, if it's Henry's fault that the cat died, then Henry should have fed the cat, or he should not have performed a cat-disabling experiment (but he did). The object of blame is typically a state of affairs (the cat's death), which is traced back to a failure to meet an obligation (a thing Henry should have done). Step two is a trivially true proposition:

> "If an agent is deciding between *A* and *B*, and they ought to do *A*, then the correct answer to the question 'what should I do?' is 'you ought to do *A*.' "[5]

This step simply blocks the objection that the "ought" of advice might differ in content from the moral "ought": if there is a fact of the matter about what one should do, then stating that fact constitutes correct advice. When an advice columnist in "The Ethicist" gives correct advice, the advice is correct precisely because it aligns with what the person morally ought to do. Step three argues that:

> "If 'you ought to do *x*' is the correct answer to an agent's deliberative question 'what should I do?', then the agent is able to do *x*."[6]

This is the critical step that establishes the 'ought implies can' principle. This step is, in effect, a particular application of a principle that Broad articulated in his treatment of obligability: that "ought" and "ought not" can only apply to an action that is, in the relevant sense, substitutable—one that the agent could have performed or could have left undone.[7] Van Inwagen's Step 3 applies this principle to the specific context of advice. The disagreement in what follows will therefore concern not only van Inwagen's particular formulation, but the deeper Broadian commitment that stands behind it, which I will elaborate on in §3.

To illustrate why advising someone to perform an action that is beyond their ability is senseless, van Inwagen used the following example:

> Suppose Planner is giving his friend advice on where to take Alison for dinner. Planner advises his friend to take Alison to Auberge du Vieux Puits in Paris, despite knowing full well that his friend cannot afford the trip and could not get there in time.

To see how each steps play together, we can apply them to the Henry feeding the cat example again.

1. Suppose the cat's death is Henry's fault (starting premise).
2. By step one, Henry should have fed the cat.
3. By step two, if Henry deliberates whether to feed the cat or go to Camp David, the correct answer to "What should l do?" is: You ought to feed the cat.

[4] van Inwagen, "Free Will and Moral Responsibility," 15.
[5] van Inwagen, "Free Will and Moral Responsibility," 15.
[6] van Inwagen, "Free Will and Moral Responsibility," 16.
[7] Broad, "Determinism, Indeterminism, and Libertarianism," 195-197.

4. By step three, for that advice to be correct, Henry must have been able to feed the cat.
5. Therefore, if the cat's death was Henry's fault, Henry was able to feed the cat.

2.1. Moral Advice as Attempt to Influence

The problem of this argument lies in step three which seems to presuppose a hidden principle:

> (CA.α) In order for X to give Y the correct advice, X must believe that Y is able to do the thing advised.

Planner's advice, "you ought to take Alison to Auberge du Vieux Puits for dinner" is a wrong answer to his friend's question "where should I take her to dinner?", because Planner does not believe that his friend is able to do the thing advised, violating CA.α.

If we build CA.α into step three, we get:

> Strong Step Three (SS.3): If "You ought to do x" is the correct answer to an agent's deliberative question "What should I do?", then the agent is (at the time of advice) able to do x.

Following SS.3, van Inwagen's point is that anyone who rejects the Ability Thesis simply cannot give out any moral advice that is consistent with their belief in that rejection; to the question "what should I do?", they cannot correctly answer "you ought to do x." To illustrate his reasoning, suppose that person X is a hard determinist. If Y asks X: "What should I do, A or B?", from X's epistemically limited perspective, X knows that Y will only do what Y does; since A and B mutually exclusive, X must believe that Y is not able to do at least one of them, but X does not know which. Therefore, on that ground, X would have to say to Y: "I cannot give you any advice, because I don't know what you're able to do."

However, this is not necessarily the case. In this scenario, what X could say is that: "once I gave Y the advice, then, Y would be able to do the thing because he would follow my advice, so it's worth a try." To elucidate, it helps to distinguish two senses in which giving advice might be "worth a try." At the metaphysical level, under hard determinism, whether X's advice will causally influence Y is already settled: the prior state of the world and the laws of nature determine whether the advice will be taken up, just as they determine everything else. In this sense, it is not "worth a try"—the outcome is fixed. But at the epistemic level, X does not know what Y is determined to do. X does not know whether the advice will be part of the causal chain that leads Y to act, or whether it will fall on deaf ears. From X's epistemically limited perspective, giving advice is worth a try in the sense that, for all X knows, the advice might be the very causal factor that determines Y's action.

This reframes what X can say when Y asks, "What should I do, A or B?" X need not claim to know that Y is presently able to do either. Rather, X can reason as follows: "I do not know, at this moment, whether Y will be able to do what I advise. But I do know that, if I give the advice and it is taken up, it will causally contribute to Y's doing the advised action." Since X cannot know in advance whether the advice will be efficacious, the rational course is to advise as if it might be—to treat the giving of advice as a causal intervention whose success is epistemically open, even if metaphysically settled.

The "able to do A" condition need not hold at the time of asking but can be satisfied at the time of action after the advice. In this sense, we can now replace CA.α with a more nuanced principle:

> (CA.β) In order for X to give Y the correct advice, X has to believe that Y is able to do the thing advised after receiving the advice.

Going back to step three, notice that the strong version secretly builds in the requirement of the agent's ability at the moment of advice. It states you may give it as correct advice only if the agent, independently of your giving it, already has the power to comply. A hard determinist can reject this as the wrong picture of advice. In a deterministic picture, when Y comes to X and asks, "What should I do, A or B?", X is not standing outside the causal order trying to read off, like a fortune teller, whose actions are already in Y's ability set. X is trying to influence Y's decision by adding a cause or a determining factor, namely, the advice.

Under this conception of advice, a hard determinist can then build CA.β into step three and have a weaker version:

> Weak Step Three (WS.3): If "You ought to do x" is the correct answer to an agent's deliberative question "What should I do?", then the agent is (after the advice) able to do x when the time for action comes.

We can see how the revised WS.3 can change the Moral Advice Argument by applying it to the Henry's cat case. Steps 1–3 remains the same, but the revised steps 4 and 5 now reads:

4. If "You ought to feed the cat" is the correct answer, then Henry must have been able to feed the cat (at the time of action) as a direct result of the advice itself.
5. Therefore, if the cat's death was Henry's fault, then there is a correct piece of advice such that, had it been given, Henry would have been able to feed the cat at the time of action.

The cat's death being Henry's fault only proves that the correct advice "feed the cat" would have been a causally efficacious factor in bringing about that action. It does not prove that Henry had any ability to choose that path independently of being caused to do so by the advice.

Those are different modal claims. SS.3 licenses an inference from "correct advice" to alternative possibility: holding fixed the actual past and the laws, Henry had both options genuinely open. WS.3 only licenses an inference from "correct advice" to counterfactual, advice-enabled possibility: in a nearby possible world where the past includes the advice, Henry feeds the cat. A natural objection is: what if the advice fails? What if Henry receives the advice and, being determined by his prior psychological states and circumstances, does not take it up? This is a real possibility, and the hard determinist should concede it since WS.3 does not guarantee that advice will always be efficacious. What it guarantees is weaker: that the correctness of advice does not presuppose any ability on the agent's part at the time of asking. Whether the advice succeeds is a further empirical question, one whose answer is metaphysically determined but epistemically open to the adviser. The hard determinist gives advice under the same epistemic conditions as anyone else—not knowing for certain whether it will work, but having reason to think it might, and recognizing that the giving of advice is itself a potential causal factor in the outcome.

2.2. **Moral Advice and Non-Cognitivism**

The account of advice I just sketched relies on a simple thought: when someone asks, "What should I do, A or B?", they are not merely asking for a description of how the moral universe already stands; they are asking for guidance that is meant to influence what they will do.

This way of thinking sits comfortably with many forms of non-cognitive or moral subjective views of morality. Suppose a person thinks that there are moral truths about things, then moral advice is just supposed to be a report about what the moral universe says. Asking, "what should I do, A or B?" is like asking, "which number is bigger, 7 to the 3 or 3 to the 7?" There's an answer to this question, and answering it is to report the truth, not to guide someone's arithmetic. That's the point of cognitivist morality, that my account of moral advice is inconsistent with.

This non-cognitive aspect is a useful note to register explicitly how my conception of advice differs from van Inwagen's. In his view, second-person 'ought'-statements are primarily informative or reportarial. When an agent asks, "What should I do, A or B?", van Inwagen understands this as a request for a judgment about an objective normative fact; the adviser's role is to provide an answer that accurately reports how things stand: what the right act is. What the advisee then does with that information is, strictly speaking, up to them and lies outside the content of the advice itself. In this picture, the adviser does not "have a dog in the fight" between the agent's competing inclinations; the adviser is not trying to strengthen one desire or create a new one, but to tell the truth about which option is right. By contrast, the hard-determinist reinterpretation treats second-person "ought"-claims as practical interventions in a causal process. When X says to Y, "ABC is the best restaurant," the primary role of this utterance is not to neutrally report a normative fact for Y to do with as Y pleases, but to strengthen or create an inclination in Y and thereby help determine what Y will in fact do. The content of the utterance is inseparable from its role in trying to steer behavior down a certain path. This is not to say that the advice lacks evaluable content: X can still be wrong about which restaurant is best, and if X is wrong, then X has given bad advice. The point is that the function of the utterance in the deliberative context is interventionist rather than reportorial. Whether the advice is correct (by whatever evaluative standards one endorses) and whether it is causally efficacious are two separate questions. The hard determinist can accept that advice is assessable for correctness while maintaining that its role in the practice of advising is to intervene in a causal process whose outcome is metaphysically determined but epistemically open to both adviser and advisee.

The Moral Advice Argument implicitly presupposes the more "objective-report" conception. My proposal is that a hard determinist can instead adopt the "interventionist" conception, and once that shift is made, the move from "correct advice" to present categorical ability in SS.3 no longer follows.

2.3. **Moral Advice and Hard Indeterminism**

Another important note is that this conception of moral advice does not work for the hard indeterminists and only works under a specific conception of hard determinism.

To see why, it helps to state more precisely what CA.β requires. On the closest-world semantics for counterfactuals, CA.β is satisfied when the following nested counterfactual is true:

> Had the agent been advised to do A, then, had the agent tried to do A, the agent would have done A.

The outer counterfactual shifts us to the nearest possible world in which the agent receives the advice; the inner counterfactual, evaluated at that world, asks whether trying to comply would have resulted in compliance. Crucially, this is a claim about what counterfactuals are in fact true—a matter of the modal structure of the world—and not a claim about what X knows. X gives advice under epistemic uncertainty about whether these counterfactuals hold, just as a doctor prescribes a treatment without certainty that it will work. What matters is whether the counterfactuals are true, not whether X can verify them.

With this in hand, we can see why the same proposal fails for the hard indeterminist. Suppose X is a hard indeterminist. If Y asks X, "What should I do, A or B?", X does not know, prior to giving advice, what Y is able to do. But crucially, even after giving the advice, X cannot be sure that Y will be able to do the action advised—and this is not merely an epistemic limitation. In X's picture of the world, the relevant counterfactual may lack a determinate truth-value: even in the closest possible world in which Y receives and tries to follow the advice, indeterministic factors may prevent compliance. The problem is the metaphysical level, not epistemic. Under determinism, if the advice enters the causal history, the closest world in which Y receives and tries to follow it is one in which Y succeeds (given causal stability). Under indeterminism, no such assurance is available even at the level of counterfactual truth, because chance may intervene between trying and succeeding.

This contrast also reveals a tacit presupposition in the deterministic case. My account of advice assumes a deterministic world with relatively low modal fragility at the level of ordinary deliberation: small variations in how advice is delivered do not produce wildly divergent outcomes within the timeframes relevant to deliberation. This is a claim about the modal structure of the world, not about what the adviser knows. It does not require infallible modal knowledge on X's part; it requires only that, as a matter of fact, the closest worlds in which the same advice is given with minor variations (delivered in 19 seconds rather than 20 or phrased slightly differently) are worlds in which the same outcome obtains. If the world were so chaotic that this were not the case, then the relevant counterfactuals would be false, and CA.β would not be satisfied—but this would be a failure of the world to cooperate, not a failure of X's knowledge.

My argument therefore supposes that, within the timeframes and levels of description relevant to ordinary deliberation, advice functions as a reasonably stable causal influence. This supposition goes beyond standard hard determinism, but it is precisely the sort of stability that our everyday advisory practices already take for granted.

3. Does 'Ought' Imply 'Can'?

So far, I have focused on van Inwagen's Moral Advice Argument and how a hard determinist can reinterpret the link between advice and ability. The deeper question remains: how can the hard determinist coherently use the language of 'ought' at all, or in other words, does all 'ought' imply 'can'? Having loosened van Inwagen's particular formulation, in this section, I turn to Broad's treatment of this deeper question in his inaugural lecture, *Determinism, Indeterminism, and Libertarianism*. I will argue that Broad's conception of 'obligability' is tied to a particular retrospective use of 'ought', and he does not clearly distinguish this from other, forward-looking uses of 'ought' as advice. On that basis, I reconstruct his two senses of 'ought' in a way that a hard determinist can retain.

3.1. Broad on 'Ought Implies Can'

Broad starts from a familiar kind of retrospective moral judgment about a past action. We say things like: "You ought not to have done *A*; you ought instead to have done *B*." He then notes that if the person wants to resist this judgment, they have two options. Either they say: (a) "I could have done *A* instead of *B*, but you are wrong about what I ought to have done; in fact, *A* was what I ought to have done," or they say: (b) "I could not help doing *A*," or "I could not possibly have done *B*." In (a), the agent grants that 'ought' or 'ought not' applies, but disputes which action is the one they ought to have done. In (b), the agent denies that 'ought' or 'ought not' applies at all, because he or she was not able to do otherwise. Broad takes (b) to rest on something like a shared background assumption: if an agent literally was not able to do otherwise, then neither 'ought' nor 'ought not' applies to that action. This is what leads him to the thesis that 'obligability entails substitutability'.

In Broad's terminology, an action is "obligable" if it is the kind of thing to which 'ought' or 'ought not' can be applied. An action is 'substitutable" if it was done but was able to have been left undone, or it was left undone but was able to have been done.[8] He then tries to make this more precise by clarifying that the kind of substitutability entailed by obligability is a 'categorical substitutability', one that does not share the logical form: "I would have been able to do otherwise if I had willed otherwise" or "I would have willed otherwise if I had willed differently in the past." Broad's view can hence be summarized in the following thesis:

> " 'Ought' or 'ought not' can only apply to an action if the action was done but was able to have been left undone, or it was left undone but was able to have been done."[9]

Broad also holds that there are two general senses of 'ought'. The first he attributes to Spinoza and calls the 'comparative ought'; This 'ought' can be applied to both humans and non-humans alike which simply states that one state of affairs ranks higher than its competitors. For example, suppose a person says, "This soup ought to have less salt."; what the person meant is that its flavor is worse than that of a properly seasoned soup. Similarly, when applied to human conduct, "You ought not to cheat at poker" may mean that cheating falls below the standard of an average decent person. Such claims are entirely independent of whether the agent is able to meet that standard; this sense, when applied to humans, is perfectly consistent with an absence of categorical substitutability or the denial of the Ability Thesis.

The second sense of 'ought' is what Broad called, the 'categorical ought'. When a person feels guilty and thinks, "I ought not to have done that," the person is not just saying, "my character is worse than average"; the person is implying that he or she was able to have and should have chosen differently, independent of his or her pre-existing character. This is the sense of 'ought' that implies categorical ability and hence making the Culpability Thesis more believable.

[8] Broad, "Determinism, Indeterminism, and Libertarianism." In the actual paper, Broad used the classical phrase "could have" as opposed to "was able to". However, given that "could have" is ambiguous as mentioned in 1.1, I assume what he meant by "could have" is really just "was able to", as opposed to "might have".

[9] Broad, "Determinism, Indeterminism, and Libertarianism," 197.

3.2. The Grammars of 'Ought'

One important note for Broad's arguments is that he never distinguishes between the retrospective 'ought' and the moral advice 'ought'. In fact, Broad's examples and formulations in the lecture almost all concern past-tense judgments:

> "You ought not to have done X.",
>
> "You ought to have done Y instead.",
>
> "He ought not to have let himself become a slave to morphine," etc.

For Broad, the grammatical form here is that we look back at what someone did (or failed to do) and pronounce a verdict. Broad calls these retrospective judgments of obligation about past actions, and his whole discussion of 'obligability' and 'substitutability' is framed in terms of such judgments.

What he does not do is to distinguish this retrospective 'ought' from other very common uses of the word we discussed in §2, namely: the future-directed 'ought' of moral advice ("You ought to tell the truth tomorrow"). Instead, Broad slides from his careful discussion of past-tense 'ought' into a more general-sounding principle of 'ought implies can'. It is then natural for readers to take this as a constraint on all uses of 'ought'.

But this slide can be resisted. We can distinguish between different roles that 'ought' plays in our language and re-interpret this retrospective 'ought' in a way that aligns with the future-directed 'ought'.

Drawing on the earlier discussion of moral advice, we can now distinguish at least three useful senses or roles of 'ought':

(i) First, we still have the 'comparative ought'. This Spinozan sense of 'ought', as sketched out by Broad, is certainly compatible with hard determinism.

(ii) Second, we have the 'deliberative ought'. This is the sense that appears when X asks, "What should I do, A or B?" and Y answers, "You ought to do A." Given the picture developed in §2, this 'ought' is tied to the practice of advice as an attempt to influence or guide action and hence can also be used consistently under hard determinism.

(iii) Third, we have the 'retrospective ought'. This is the past-tense 'ought' that Broad focuses on when one looks back and makes a normative verdict about a past action. Based on Broad's view, this is a 'categorical ought' that is rendered false if the agent was not able to do otherwise.

The problem is how can hard determinists interpret the 'retrospective ought' in a way that does not commit them to the categorical sense. We can use the insight from WS.3 in §2; a hard determinist can analyze retrospective judgments of the form "X ought (not) to have done Φ" by reducing them to counterfactual claims about correct prior advice. More precisely:

(R1) "X ought (not) to have done Φ at t" $\Leftrightarrow$ "If, at some suitable time t^*, prior to t, X had asked 'What should I do?', then the correct answer would have been: 'You ought (not) to do Φ.' "

For a hard determinist, to make a retrospective judgment and say "you ought to have done or ought not to have done something" is like saying, "the advice I would have given you if you'd asked beforehand is x." On this analysis, a retrospective 'ought' is a claim about what the correct deliberative 'ought' would have been, had X sought advice or guidance beforehand. The truth-conditions of the retrospective "You ought to have done Φ" are therefore identical to the truth-conditions of the corresponding advisory claim "You ought to do Φ" evaluated at an earlier time t^*. In this way, the retrospective 'ought' inherits the interpretation of the deliberative 'ought' the hard determinist adopts in CA.β, without introducing any further commitment to categorical ability.

Therefore, when X says to Y, "you ought to have done A or ought not to have done B," X might be doing at least two things:

(1) Evaluating *Y*'s past action relative to a standard ("*Y*'s action falls below the standard for a decent person"), which is a comparative 'ought',
or

(2) Indicating what the correct advice would have been and, simultaneously, issuing a forward-looking recommendation ("In similar situations, you ought to do *A* and not do *B*"), which is effectively a deliberative 'ought'.

Under this interpretation, there are still essentially two senses of 'ought', both of which are compatible with hard determinism, undermining the 'ought implies can' principle as argued by Broad.

4. Out of the Culpability Dilemma

So far, I have argued that a hard determinist can make sense of second-person moral advice by treating "You ought to do *A*" as a causal intervention rather than as a neutral report about an independent normative fact. A natural next question is whether the same strategy can be extended to fault and blame.

The hard determinist can answer this by extending the same causal interventionist account from advice to blame. To say "It is your fault that *S* happened" has at least two components. First, there is a descriptive, cognitive component: it attributes to the agent a certain causal and normative relation to *S*; it notes that an action or omission by the agent played a role in the deterministic history leading to *S* and evaluate that action as falling below some accepted moral standard. In this sense, the hard determinist can say "You are at fault," to mean "Your behavior was one of the causes of *S* and manifested a defect relative to the standard we endorse." No appeal to categorical abilities is required for that much.

Secondly, there is a forward-looking, non-cognitive component. In ordinary life, we use "You are to blame" to produce feelings of guilt and remorse to shape the agent's future disposition or to reinforce a shared norm in the community. In the interventionist picture, ascriptions of fault and blame are themselves attempts to influence the future causal chain in much the way that advice is.

This allows the hard determinist to reasonably reject the ability requirement found in clause (3) of the Culpability Thesis and instead substitute a weaker condition, parallel to the Weak Step Three from §2.1. Accordingly, we can construct a hard deterministic reformulation of the Culpability Thesis as follows:

> A bad state of affairs S is truly the fault of a person *x* if and only if (1*) *x*'s action or omission was a part of the causal history of *S*, (2*) that action or omission falls below a standard that we are entitled to enforce, and (3*) there is some correct pattern of guidance or advice such that, had *x* been appropriately exposed to and responsive to it earlier, *x* would have behaved differently and *S* would not have occurred.

On this view, fault-ascriptions, like advice, are part of the machinery by which a deterministic system regulates itself.

We can now see how this returns us to the Culpability Dilemma. Given that the practice of any fault and blame ascriptions require only causal connection to bad outcomes, deviation from standards we are prepared to enforce, and a forward-looking role for blame and advice as instruments for influencing behavior, we can consistently deny 'ought implies can' under hard determinism, and path (ii) remains open to get out of the Culpability Dilemma.

Bibliography

Broad, C. D. "Determinism, Indeterminism, and Libertarianism." *Ethics and the History of Philosophy: Selected Essays*. Routledge, 2014. https://doi.org/10.4324/9781315823058.

van Inwagen, Peter. *An Essay on Free Will*. Rev. ed. Oxford: Oxford University Press, 1986.

van Inwagen, Peter. *Thinking about Free Will*. Cambridge: Cambridge University Press, 2017. https://doi.org/10.1017/9781316711101.

van Inwagen, Peter. "Free Will and Moral Responsibility: Once More unto the Breach." *Belgrade Philosophical Annual* 38, no. 1 (2025): 7–19. https://doi.org/10.5937/BPA2501007I.

van Inwagen, Peter. "Does 'Ought' Imply 'Can'?" *Free Will: Historical and Analytic Perspectives*, edited by Marco Hausmann and Jörg Noller, 313–33. Cham: Springer, 2021.

Do we need to prove the existence of the 'external' world?

Haseeb Ali, *University of Kent*

In this paper, I defend Wittgenstein's insight that Cartesian-style skeptical arguments against the existence of the external world are not false, but linguistic confusions that violate the constitutive rules of our epistemic practices. To do this, I advance a resolutely grammatical interpretation of his later notebooks, posthumously published as *On Certainty*, where so-called 'hinge-propositions' function as rules of grammar and logically exclude the possibility of 'global' doubt. I then defend this reading from two notable objections by Crispin Wright, namely that (1) if 'hinges' are not semantically truth-apt propositions, they cannot function as premises in logical inferences, and (2) the groundless nature of grammatical propositions renders the foundation of our language-game arbitrary, leaving us unable to distinguish a rational world-picture from mere dogma.

Introduction

Kant famously regarded the 'scandal of philosophy' to be its inability to present a sound proof for the existence of the world that could withstand any doubt[1]. While G.E. Moore argued that his proof accomplishes that task, Wittgenstein maintained that the demand for a proof itself rested on an incorrect understanding of language that, upon correction, would render a proof unnecessary and logically exclude the possibility of doubt. This paper proceeds in three main parts. In Section 1, I present Wittgenstein's criticism of Moore's proof that it incorrectly uses the term "I know", and argue that hinges function as grammatical propositions. I also present a unique criticism of what I term P.M.S. Hacker's 'intermediate view' on the status of hinges. In Section 2, I present Crispin Wright's theory of entitlement and warrant transmission as a rival strategy before outlining two central objections he raises against the grammatical reading. Finally, in Section 3, I develop a novel defense against Wright's position and his specific objections. I argue that treating hinges as rules of grammar does not leave our epistemic practices at risk of arbitrariness or leaching off of unearned warrant, which is a more robust dissolution of the problem of skepticism than Wright's strategy. Furthermore, I establish that Wright's entitlement theory functions as an epistemically inert pragmatism that is excessively concessionary to the skeptic.

1. Wittgenstein's Engagement with Moore

1.1. Moore's Proof and the Limits of Sense

Wittgenstein's reflections respond primarily to G.E. Moore's two seminal papers, 'A Defence of Common Sense' (1925) and 'Proof of an External World' (1939).[2] Moore sought to rebut idealism with an ostensibly basic proof by declaring, while raising each hand in turn, "Here is one hand", and "Here is another". He insisted that this demonstration *ipso facto* proved an external world by meeting the formal requirements for proof, namely, having epistemic certainty of the premises and logically inferring the conclusion. Acknowledging the sceptic's objection regarding unproven premises, Moore rejected the requirement for an infinite regress of proof; instead, he maintained that certain common-sense truths were known directly and had to be accepted to make further reasoning possible.[3] Moore's strategy, later known as the "Moorean shift",[4] was to reverse the sceptic's logic. While the sceptic used *modus tollens* to argue that a lack of knowledge

[1] Kant, *Critique of Pure Reason*, Bxxxix.
[2] G. E. Moore, *Philosophical Papers* (1959).
[3] Moore, *Philosophical Papers*, 146–150.
[4] See Rowe, *"The Problem of Evil"* (1979), 335–341.

regarding skeptical hypotheses precludes knowledge of the external world, Moore employed *modus ponens* to argue that because he knew he had hands, the skeptical hypotheses must be false.

Wittgenstein denied that Moore could claim knowledge of his premises, arguing that the error lay in misusing the term "I know." Central to this critique is the logical requirement, tracing back to the *Tractatus,* that empirical propositions must be bipolar, i.e. either true or false, to possess 'sense' (TLP 4.022–4.023). Moore's truisms lack this bipolar character, prompting Wittgenstein to label as "obviously nonsensical" any attempt to raise doubts where no questions can be asked (TLP 6.51). This principle reappears later in *On Certainty,* where Wittgenstein maintains that knowledge-claims are only intelligible where doubt is possible (OC 58). Because Moore's assertions, unlike ordinary cases, cannot be justified by citing evidence, they carry an illicit "metaphysical emphasis" on account of their groundless nature (OC 482–484). They are more akin to a child's acceptance of a "tree" through ostensive teaching (OC 480). Consequently, these propositions function not as empirical descriptions but as "hinges" (OC 341) that serve as the "axis" of all inquiry (OC 152–153). Just as a door requires fixed hinges to move, our empirical inquiries depend on certain propositions to remain exempt from doubt so that we can evaluate other claims. These propositions are termed hinges to illustrate their critical role in our language-game. This marks an anti-foundationalist shift that dissolves the infinite regress of justification by recognizing the "groundlessness of our believing" (OC 166). If these hinges are not empirical, they must be understood as constitutive rules of linguistic practice.

1.2. Hinges as Grammatical Propositions

In his transition from the *Tractatus,* Wittgenstein reconceives the status of foundational propositions, moving away from a strictly picture theory of language. This meant that words and propositions only possessed a sense if they functioned like a logical map that mirrored a corresponding fact in the world. Language, then, was a descriptive enterprise to 'map' out all the facts of the world and form a picture. However, in his later transitionary period after returning to Cambridge from 1929 to 1936, Wittgenstein's understanding of language begins to shift towards a rule-governed activity, as seen in his *Philosophical Grammar.* Wittgenstein recognizes that before any 'picturing' can occur, there must be a background of rules and conventions that function as the conditions of picturing. This marks a shift from a logic of representation to one of linguistic practice. While maintaining that tautologies, contradictions, and logical rules lack bipolarity (TLP 4.466), he departs from the standard logico-semantic usage that reserves the term 'proposition' exclusively for truth-evaluable claims. This later shift marks a widening in his understanding of grammar. Here, grammar is no longer merely a system of syntax but encompasses "all the conditions (the method) necessary for comparing propositions with reality" (PG I, 45).

Hinges, despite carrying the outward syntactic form of empirical propositions, satisfy this deeper grammatical role. For example, the sentence 'nothing can be red and blue all over at the same time' isn't discovered by examining thousands of objects to find a counter-example. Instead, it defines how we use color-words. If someone were to claim they found an object that was red and blue all over, we would question if they understood the meaning of the words 'red' and 'blue' rather than attributing a scientific discovery to their name. They function as logical enabling conditions that make our epistemic practices possible (OC 401-402). Consequently, I retain his technical use of 'grammatical proposition' [*grammatischer Satz*] to capture their unique logical function despite possessing the appearance of a descriptive fact about the world. They act as 'norms of representation' i.e. rules that "describe the structure of language; describe its possibilities" (BT, 149e). Because hinges are propositions that have been 'hardened' from what were once empirical observations (OC 96, 167), they are removed from the realm of verification. They can no longer be tested like hypotheses as they are the very criteria for what it means to 'test' anything at all. Acting as the framework of our language-game, they are the 'limiting cases' (TLP 4.466; cf. OC 54) that give our notions of 'truth' and 'falsehood' their sense.

This constitutive role connects the establishment of a grammatical rule directly to the adoption of a unit of measurement. By fixing the 'unit', grammar enables the use of empirical propositions to describe the world; as Wittgenstein notes, while the rule itself establishes the metric, the empirical proposition "says how long an object is" only by making use of that established rule (BT, 189e). This illustrates how hinges, as logical enabling conditions, do not describe facts but rather provide the very possibility of factual description. Consequently, Wittgenstein argues that grammar itself is autonomous and, being unaccountable to any reality, is ultimately arbitrary (PG I, 133). Attempting to justify the rules of grammar is a category mistake; it is nonsensical because they define the terms of justification itself. As he notes: "the conventions of grammar can't be justified by a description of what is represented... [because] any description of that kind already presupposes the rules of grammar" (BT, 188e). Justification, therefore, must come to an end. What lies at the bottom of the language-game is not a self-evident truth that we perceive, but our way of acting. Hence, grammatical propositions lack truth-aptness, as "the ground is not true, nor yet false" (OC 204-205).

Because these foundational grounds are exempt from truth-evaluations, their adoption is not compelled by objective facts but by our shared practices. The 'arbitrariness' of grammar is thus understood in this strict sense of choosing a metric (BT, 186e). Just as the standard meter bar in Paris (PI 50) is used to measure lengths but cannot itself be measured against the scale it creates, hinges are used to 'measure' truth but cannot be subjected to those same evaluations. The logical exclusion of doubt is absolute. To doubt a hinge is not to question a fact, but to break the tool that makes its expression possible. Such 'global' doubt amounts to an "annihilation of all yardsticks" (OC 492). We recognize these hinges as the constitutive rules of our language-game because of their peculiar role in logically excluding the possibility of doubt (OC 194). This normative function of regulating the boundaries of language, despite being 'arbitrarily' chosen, is how hinges ensure the stability of the entire framework of inquiry.

Given the unrefined nature of Wittgenstein's notebooks[5], the categorical status of hinges remains contested. Two dominant interpretations emerge from this ambiguity[6]. G.H. von Wright and Danièle Moyal-Sharrock emphasize the 'animal' nature of these certainties, situating them as non-propositional forms of praxis, which refers to the unreflective and habitual ways of acting that underpin our language-games.[7] Conversely, P.M.S. Hacker argues that they occupy a 'curious intermediate position,' neither wholly empirical nor strictly grammatical.[8] He contends that they lack the normative function required to determine concepts or inference rules.[9] However, I argue that Hacker's refusal to grant hinges grammatical status is difficult to reconcile with his own admission that empirical propositions can be transformed into 'norms of representation'.[10] It is important to note that while Genia Schönbaumsfeld primarily defends the grammatical reading as a rejoinder to Moyal-Sharrock's non-propositional account,[11] she does not explicitly extend this criticism to Hacker's 'intermediate' stance. Rather, Schönbaumsfeld's only mention of Hacker is agreeing with him that only empirical propositions are bipolar while grammatical remarks are not. She does not seem to note Hacker's position is yet different, despite this brief congruence. I argue this only further indicates that Hacker's rejection of hinges as grammatical propositions is a dogmatic position, since he is already keen on granting the possibility of other empirical propositions 'hardening' into rules, particularly in mathematics, that have the same normative rule-like function as hinges. As such, I extend Schönbaumsfeld's previously focused argument against Moyal-Sharrock into a broader view that also extends to other competing interpretations of *On Certainty* and attempt to resolve previous contradictions into a clearly defined position.

1.3. Grammatical Analysis of the 'External' World

A consequence of the grammatical reading is the exposure of the sceptic's error in bifurcating reality into 'internal' and 'external' domains. Since linguistic meaning is tethered to a shared world, any expression of doubt necessarily concedes the world's existence as a condition of its own intelligibility. The sceptic's position is therefore semantically nonsensical, as it attempts to negate the very framework that makes their speech possible. As Wittgenstein notes, "If you are not certain of any fact, you cannot be certain of the meaning of your words either" (OC 114). On the reading I propose, hinges function as logical enabling

[5] See OC Preface, vii.

[6] This account is not exhaustive. See Williams, *Unnatural Doubts* (1991); Coliva, *Moore and Wittgenstein*, (2010); Pritchard, *Epistemic Angst* (2015).

[7] See von Wright, *Wittgenstein* (1982, 178); Danièle Moyal-Sharrock, *Understanding Wittgenstein* (2004).

[8] Hacker and Baker, *Wittgenstein: Rules, Grammar and Necessity* (2009, 258).

[9] Hacker, *Wittgenstein's Place in 20th Century Analytic Philosophy* (1996, 215).

[10] Hacker, *Wittgenstein: Comparisons and Context* (2013, 161). For a detailed critique of Hacker's rejection of hinges as grammatical propositions, see Moyal-Sharrock, "Beyond Hacker's Wittgenstein" (2013).

[11] Schönbaumsfeld, "Hinge Propositions and the "Logical Exclusion of Doubt" " (2017, 104).

conditions that regulate the language-game rather than describing a reality beyond it. While they enable the sensible application of truth and falsehood, they are not themselves truth-apt in an empirical sense. This reading, however, faces two significant challenges from Crispin Wright's account of 'epistemic entitlement': (1) that the non-truth-aptness of hinges precludes them from serving as premises in valid logical inferences, thereby blocking the warrant-transmission required to justify our ordinary beliefs, and (2) that the autonomy of grammar poses an epistemic challenge of arbitrariness, threatening the rationality of our foundational commitments.

2. Hinges without Warrant?

2.1. Wright's Diagnosis: Transmission Failure

Wright develops his critique by reconstructing Moore's argument with a tripartite schematic of Type-I, Type-II, and Type-III propositions. Here, (I) the subjective premise 'I seem to have hands', provides warrant for (II) the empirical claim that 'a hand exists', which in turn entails (III) the metaphysical conclusion that 'an external world exists'. Type-I propositions describe sensory experiences that justify empirical Type-II propositions about the external world, such as Moore's truisms. However, warrant can only transmit from Type-I to Type-II propositions if we already possess antecedent warrant for Type-III propositions. He defines these 'cornerstones' as foundational commitments such that, if we lack warrant for them, we cannot claim justification for the set of empirical propositions, or as Wright puts it, a *'region of thought'*, that depends upon them for their validity. [12].

Moore's premises, then, are "information-dependent" on Type-III propositions because sensory experiences can only ground empirical claims if we already grant causal interaction with the world. Since belief in an external world lacks antecedent evidential warrant, skeptical scenarios remain live possibilities that threaten to undercut the evidence of our senses. As such, Wright concludes that Moore's proof is epistemically circular.[13]

Wright tentatively draws support for his position from Wittgenstein's own observations that the grounds for common-sense propositions are never "as certain as the very thing they were supposed to be grounds for". Since the grounds are no surer than the assertions themselves, such propositions are instead affirmed "without special testing" (OC 136, 243, 307). Nevertheless, Wright fundamentally diverges from reading hinges as non-truth apt grammatical propositions by reintroducing a need for rational justification that he holds can be acquired non-evidentially.

2.2. Epistemic Entitlement

Wright correctly recognizes that Moore's proof cannot avoid transmission failure if evidential justification is the only way to acquire warrant for cornerstones. To avoid this impasse, Wright proposes a "unified strategy" that expands the definition and scope of acquiring warrant to disjunctively encompass evidential justification, based on proof, and entitlement, based on trust. In other words, Wright now argues that a cornerstone can be rationally held even in the absence of evidence, provided it meets the criteria for entitlement (WN 177, 208).

Wright introduces the concept of 'entitlement' to *accept*, rather than *believe*, any proposition *P* despite lacking evidential warrant. He concedes that the sceptic is correct to claim that we lack justified *belief* in cornerstones but argues this does not prevent our *acceptance* of them (WN 174-176). Whereas belief is a doxastic attitude requiring evidence and conviction of truth, acceptance is the practical attitude of acting on the assumption that *P*. Unlike justified belief, entitlement is a rational trust in accepting *P*. Wright divides this non-evidential warrant into various categories. First, Wright constructs a dominant strategy he terms 'Strategic Entitlement' and analogizes it to the decision-theoretic model in Hans Reichenbach's defense of induction. Reichenbach agreed with Hume that *epistemic* justification for induction was impossible to

[12] Wright, "Warrant for Nothing", 167; henceforth WN. See also Davies, "Epistemic Entitlement," 219n2, who observes a nuanced distinction between not having rational warrant and not being able to claim it.

[13] Wright, "Wittgensteinian Certainties," 23, 26–28; henceforth WC. Wright, WN, 171–172.

prove through both *a priori* and *a posteriori* means but argued that we are rationally entitled to using it because it is the only decision that offers a possibility of success, whereas skepticism guarantees failure (WN 178-180). Likewise, Wright's strategy reframes Reichenbach's pragmatic argument to a form of epistemic consequentialism. It justifies the acceptance of cornerstones based on the epistemic consequence of securing the possibility of cognitive inquiry. Second, Wright identifies the 'Entitlement to Cognitive Project' as an entitlement to presuppositions that stave off cognitive paralysis from an infinite regress of doubt.

Wright insists that this appeal to entitlement does not constitute a mere dogmatic dismissal of the sceptic's challenge[14] but is grounded in the *epistemic* necessity of trust. He notes that while these entitlements are effective against Cartesian and Humean forms of skepticism, they do not readily extend to ontological cornerstones (WN 188, 194–197). Nevertheless, because our lives as rational agents depend on trusting Type-III propositions, Wright argues that avoiding them makes life impossible since rational agency is a "not an optional aspect of our lives". As such, we are justified in presupposing their truth without needing to earn a warrant (WC 53).

Wright immediately identifies a serious problem that threatens to undermine his project. Since Wright affirms that cornerstones are held by trust, they are beyond evidential justification and introduce a degree of epistemic risk. This unearned warrant might 'leach up' and contaminate the warrants for otherwise ordinary empirical claims, such as visually perceiving one's hand. By highlighting the "leaching problem", the skeptic can now argue that ordinary justified beliefs are taken up on mere faith. To resolve this, Wright proceeds by stratifying knowledge claims and restricting warrant to second-order claims. We cannot claim to know that cornerstones are true, but we can still possess first-order knowledge of them (WN 209). On this view, possession of knowledge is "liberal", requiring no antecedent warrant, while claiming it is "conservative".[15] Wright claims that leaching would thus be confined to second-order warrants, i.e. claims of possession, leaving our first-order warrants untroubled.

2.3. **The Semantic and Epistemic Challenge**

Wright's notion of epistemic entitlement presents a semantic challenge to the grammatical reading of hinges. Namely, Wright questions how hinges can enable us to distinguish between true and false beliefs if their own semantic content is not truth-conditional. Wright maintains that because empirical inquiry has an external objective, the "divination of what is true and the avoidance of what is false" (WC 43), our cornerstones must be answerable to this truth-seeking goal. The rules of arithmetic, like '13 + 7 = 20', must not force us to "misassess the truth-values" of the specific things we are investigating. Wright presents an example of a person counting 13 satsumas and 7 bananas but mistakenly arriving at a count of 21 fruits, compelling the person to conclude that they miscalculated. Here, the rule overrides the sensory evidence because its certainty is ranked higher. However, if hinges were merely groundless norms of representation, they would have no authority to dismiss the physical possibility that there really are 21 fruits. This is because, for Wright, unless a rule like $13+7 = 20$ is truth-apt, it cannot determine or describe what empirically exists in the world. Wright concludes that if a rule is to reliably guide us to the truth, it must be a "substantial proposition, apt to be true or false" (WC 34, 44–46). Denying this would leave cornerstones incapable of functioning as premises in valid inferences, like Moore's proof, and 'saves' skepticism only by expelling them from the realm of logic entirely.

Even if cornerstones, such as the Earth having existed for many years past (OC 138, 401), serve as rules, Wright argues this is not mutually exclusive with their being truth-evaluable. Interpreting cornerstones as *hors de combat* i.e., incapable of intelligent mismatch with the world, is hasty and does not banish "the spectre of profound and sweeping error" (WC 44).

In addition to the semantic challenge, Wright identifies a deeper epistemic threat of arbitrariness that looms over the grammatical reading. If grammar is autonomous and, by Wittgenstein's own admission (PG 133), arbitrary, it's difficult to see how we can secure a rational standing for epistemic practices. Although Wright acknowledges cornerstones are themselves unfounded [*unbegründet*] by calling them presuppositions (2004b, 48–49) he still maintains the need for a rational basis in affirming and acting upon them.

[14]Cf. OC 71, 155, 498, 558, 559. These passages may imply dogmatism if hinges are treated as empirical propositions.
[15]Wright, "On Epistemic Entitlement (II)," 217–218; henceforth WSE.

Wright terms the "Problem of Demarcation" as the danger of being unable to distinguish cornerstones from "prejudices, mere assumptions, and idées fixes" (WSE 245). If our cornerstones lack rational warrant, we risk succumbing to a pernicious relativism where we cannot distinguish a legitimate constitutive rule and a mere subjective whim.

To avoid running into an infinite regress of demands for justifications, Wright's tentative response is to reject what he calls the "Evidential Ideal", the notion that a rationally epistemic thinker should proportion all their beliefs to evidence and accept nothing without it. Wright argues that the ideal itself is incoherent because evidential reasons to believe can only occur within a framework held by trusting acceptances. This trust, far from being a flaw, is an essential aspect of rational inquiry without which we lose the ability to determine what counts as evidence (WSE 242–245). For example, determining why perceptual experiences relate to matters in the external world necessitates the acceptance of cornerstone that "There is an external material world, broadly manifest in normal sensory experience" (WSE 217). The grammatical reading, despite claiming logical immunity from skepticism, does not concord with the life of a rational agent precisely because it is free from the necessity of trusting cornerstones. As such, for Wright this reading fails because it cannot distinguish itself from the arbitrary acceptance of dogma.

3. Response to Wright

3.1. The Logical Barrier

Wright is correct to identify that if hinges are not truth-apt, then they cannot function as premises in any logical inferences, including Moore's proofs. Moreover, he is right to note that their non-descriptive nature, as norms of representation lacking sense, precludes them from being items of knowledge.

However, I reject that this has any epistemically catastrophic consequences. First, Wright's rejection of the grammatical reading rests on a subtle category mistake that conflates the rules of the language-game with the moves made within it. Because Wright treats the foundations of our rational system and the empirical claims made within it as belonging to the same class of propositions, namely empirical ones, he incorrectly demands that they be subject to the same epistemic evaluations. To follow Wittgenstein's analogy of language as a game of chess, Wright's approach is akin to claiming that the rule 'The bishop moves diagonally' is of the same kind as the report 'The white bishop moved from its starting square at c1 to f4'. The first sentence only expresses a rule that defines how the bishop functions in every game whereas the latter describes a specific instance of play. Wright overlooks the notion that the scope of rationality, so to speak, is 'local'; it can only take place against a backdrop of propositions, our world-picture as the inherited background, that we don't stand in any epistemic relation to. Hinges are, as Schönbaumfeld notes, '*logical enabling conditions*' that give our notions of truth and falsity their sense and without which we would lose the very meaning of our words. Their scaffolding status serves to clarify the correct use of language. Moore's very attempt to use hinges, along with his belief that proof was necessary, was mistaken precisely because he treated the constitutive rules of grammar as if they were empirical claims subject to evaluation.[16]

In contrast to Wright's assertion that hinges must be empirical propositions in need of warrant, a resolute grammatical reading maintains that this criterion is unnecessary because it is logically impossible to doubt hinges. Wittgenstein repeatedly analogizes the rules of grammar to the rules of chess (BT, 118e) to illustrate that questioning them is no more sensible than questioning why the bishop moves diagonally.

Wright's confusion lies in the fact that he views cornerstones to be descriptive facts about reality. But rather than describing empirical facts, the meaning of grammatical propositions lies in their use in language where their "explanation of meaning" (PG I, 23) serves a prescriptive function; to establish the rules of our language-game and 'police', so to speak, the correct method of using concepts. This is the key insight behind Wittgenstein's assertion that the proposition, "there are physical concepts", is senseless and cannot be formulated (OC 36). Because it is impossible to make sense of the proposition's denial, i.e. that there are only sense-experiences, it follows *ipso facto* that the assertion itself is mistaken. Such propositions find their only use as instructions. It does not follow, therefore, from the lack of truth-aptness of hinges, and their inadmissibility as premises that we cannot make any logical inferences with empirical propositions. Rather,

[16] Genia Schönbaumsfeld, " 'Logical' and 'Epistemic' Uses," 239, 244–245.

as I have already attempted to explain, Moore's mistake lay in using hinges in his proof, for the exact reason that Wright purports is allegedly a defect in the grammatical reading.

Having shown that a well-defined distinction exists between hinges as the rules of our language-game and the empirical propositions as 'moves' within it, I argue it becomes clear that Wright's concern about leaching is dissolved. At the bottom of our language-game lie hinges that do not provide warrant but logically enable the conditioning of sense. Since 'risk' is an epistemic concept, it cannot apply to the grammatical framework itself. Thus, leaching does not seep up, and any attempt at grounding this world-picture would only invite further grounding, *ad infinitum*. In a recent debate with Annalisa Coliva and Anil Gupta[17], despite amending his prior definition of entitlement by adding that it is the attitude towards a cornerstone proposition *P*, Wright's view is still at risk of such regress, for the rational trust he places in *P* remains exposed to the demand of entitlement towards *that* attitude as well. Wright's demand for hinges to be truth-apt is an attempt to repair a defect that does not exist if they are taken to constitute a logical barrier for our language-game. By conceding that cornerstones require justification, Wright leaves himself vulnerable to an infinite regress of justification in addition to leaching. I argue this conundrum is downstream of the category mistake I have identified in Wright's work so far. By treating empirical claims and cornerstones as the same kind of proposition, Wright renders the task of preventing leaching impossible.

3.2. **Contra Entitlement**

Wright explicitly rejects the label of pragmatist by insisting that cornerstones are essential for the possibility of our cognitive projects,[18] and that this constitutes a normative condition. Wright's decision-theoretic strategy contends that because these entitlements aim to enhance "expected epistemic utility" within a cognitive project, they are also epistemic reasons (WSE 239, 241–242). However, I argue this appeal to utility and the goal-dependent nature of the 'dominant strategy' to avoid intellectual stasis functionally represents a pragmatist solution despite Wright's insistence to the contrary. Wright shifts the justification of cornerstones from the evidence of their truth to the utility of their acceptance as the means of attaining warrant, drawing on an earlier contrast he makes between acceptance and beliefs (WN 175, 184–188). As Wright acknowledges, "there is no coherent conception of epistemic rationality which views unevidenced trusting as per se a lapse" (WSE 243). In other words, Wright's unified strategy remains an instrumental form of rationality. But, as Wittgenstein is keen to remind us, the sceptic's challenge is not one of practical doubt but a more radical 'global' form that lies behind it (OC 19). The sceptic would succeed in pinpointing the presuppositions of our epistemic practices, and consequently the very meaning of notions like "justification" and "warrant" that Wright is keen to employ.

Furthermore, Wright's stratification of warrant fails to dispel concerns of leaching because, as Aidan McGlynn argues, Wright's liberalism towards first-order warrants, such as having a justified belief in Moore's premise ("Here is a hand") in combination with the recognition that it entails belief in the cornerstone proposition ("There is an external world") leads him to adopting a form of neo-Mooreanism that delivers the antecedent warrant that Wright's conservatism at the second-order demands, resulting in his entitlement theory being rendered "an idle wheel".[19]

Finally, I argue that the notion of an entitlement is epistemically inert; it satisfies the structural requirement for cornerstone propositions without providing the truth-conducive grounding necessary to stop the leaching of unearned warrant. This results in a system of conditional rationality where the cognitive project of our empirical beliefs is internally consistent but anchored by a set of "blind" wagers that we are incapable of confirming. This underlying pragmatism is notable in Wright's rejection of the "Evidential Ideal" (WSE 243) on the grounds that it is an integral part of our rational reflective inquiry that we take some items of knowledge on unearned trust, because the alternative is "intellectual stasis". This, as I have attempted to show, does not sufficiently address the 'global' nature of the sceptic's doubt.

[17] See Coliva, Gupta, and Wright, "A Debate on Skepticism and Perceptual Belief," 361, 379.

[18] See Pritchard, "Wittgenstein's On Certainty,"189–224; Jenkins, "Entitlement and Rationality," 25–45. Cf. Wright, WSE, 213–248 for his defense against such "instrumentalist" interpretations.

[19] McGlynn, "Epistemic Entitlement and the Leaching Problem," 95–96, 98.

3.3. Hinges as Constitutive of Rationality

This brings us to the core of Wright's 'Problem of Demarcation': the concern that treating hinges as groundless rules renders our world-picture merely optional. I argue Wright's notion of 'Strategic Entitlement' treats cornerstones instrumentally as seen in the contrast Wittgenstein makes between the rules of cooking and grammar; cooking is defined by its end, whereas speaking a language isn't (PG I, 133). There is a categorical difference between instrumental activities and grammatical ones:

> "I believe the reason one isn't tempted to call cooking a game is this: …'cooking' doesn't refer to an activity that follows those rules—rather, to an activity that has a particular result…On the other hand "playing chess" doesn't mean an activity that has a particular result, but rather an activity that corresponds to such and such rules." (BT, 186e).

If someone were to change or not follow the rules of cooking (e.g., boiling an egg for five minutes instead of three), he might fail to achieve the intended result; but if we change the rules of grammar, we don't speak "falsely", rather we are playing a different game. The notion of not following a rule of grammar is nonsense. Wright treats hinges instrumentally in his decision-theoretic strategy, like a cooking recipe designed to achieve a desired outcome: truth or survival. Secondly, Wittgenstein's assertion that grammar is arbitrary does not imply hinges are mere prejudices, assumptions, or *idées* fixes, as Wright worries by identifying a "Problem of Demarcation" (WSE 245) but only that one shouldn't confuse a rule with the use of words or expect it to be answerable to reality in the way an empirical proposition is. This is because grammar is autonomous; it does not 'mirror' the world but rather provides the framework that makes any description of the world possible. The conventions of grammar cannot be justified by a description of what is being represented, as "any description of that kind already presupposes the rules of grammar" (BT, 188e).

As I established earlier, because hinges are logical enabling conditions like the constitutive rules of a game,[20] they cannot be random assumptions that we could conceivably make differently, owing to their contingent nature. The arbitrariness of grammatical rules stems from our observing their fixity in language before removing these propositions from the realm of contingency.

> Consider: "In language the only correlate to natural necessity is an arbitrary rule. It is the only thing one can remove from this necessity and put into a proposition." Refers to propositions such as $\sim\sim p = p$. (BT, 185e)

Just as the rule of double negation is not an empirical fact we discover, but a constitutive rule informing us of how the word "not" is used, likewise the arbitrariness of grammar consists in the fact that it is not responsible to any reality; it is not a description of facts, but the very framework that makes descriptions of facts possible. Furthermore, our world-picture of hinges are not axioms that provide the starting point for language, but the very medium in which rationality exists.

> "And this system is not a more or less arbitrary and doubtful point of departure for all our arguments: no, it belongs to the essence of what we call an argument. The system is not so much the point of departure, as the element in which arguments have their life." (OC 105)

Wright's mistake lies in treating cornerstones as these points of departure since they open us to question if alternate points of departure could be chosen or not, so to speak. By insisting on the empirical and contingent nature of cornerstones, Wright's entitlements are vulnerable to the same problem of demarcation he identified earlier.

4. Conclusion

The debate between Wright and the grammatical reading I have proposed turns on the nature of the ground of our epistemic practices. Wright accepts the sceptic's demand that cornerstones require warrant and offers

[20] Schönbaumsfeld, "'Logical' and 'Epistemic' Uses," 241.

'entitlement' as part of his unified strategy. Yet, as I have argued in this paper, Wright's strategy is epistemically inert because it doesn't adequately resolve the problem of leaching, leaving us with a system that is only conditionally rational. By conceding to the skeptic that our cornerstones require warrant, Wright leaves our entire epistemic practices vulnerable to the leaching problem. Because entitlement is mere trust, it cannot prevent the further contamination of our empirical claims. In contrast, by acknowledging hinges as grammatical propositions, I argue such doubts are logically excluded and that the arbitrariness of grammar is constitutive, functioning like the rules of a game that define the field of play rather than the moves made within it. I provide a more robust dissolution of the skeptical challenge by identifying Wright's category mistake through my analogies with chess games. I show that hinges are the constitutive grammatical rules that define the moves made within our language game. This is a significant shift towards an anti-foundationalist attempt at securing the rational integrity of our epistemic practices without having to justify, potentially infinitely, each cornerstone proposition.

List of Abbreviations

Abbreviation	Full Title & Edition
BT	The Big Typescript: TS 213, ed. and trans. C.G. Luckhardt and M.A.E. Aue (Oxford: Blackwell, 2005).
OC	On Certainty, ed. G.E.M. Anscombe and G.H. von Wright, trans. D. Paul and G.E.M. Anscombe (Oxford: Blackwell, 1969).
PG	Philosophical Grammar, ed. R. Rhees, trans. A. Kenny (Oxford: Blackwell, 1974).
PI	Philosophical Investigations, trans. G.E.M. Anscombe, P.M.S. Hacker, and J. Schulte, 4th ed. (Oxford: Wiley-Blackwell, 2009).
TLP	Tractatus Logico-Philosophicus, trans. D.F. Pears and B.F. McGuinness (London: Routledge & Kegan Paul, 1961).
WC	Wittgenstein and Skepticism, ed. Denis McManus (London: Routledge, 2004).
WN	"Warrant for Nothing (and Foundations for Free)?" The Aristotelian Society Supp. 78, no.1 (2004).
WSE	"On Epistemic Entitlement (II): Welfare State Epistemology." In Scepticism and Perceptual Justification (Oxford: OUP 2014).

Bibliography

Coliva, Annalisa. *Moore and Wittgenstein: Scepticism, Certainty and Common Sense.* Basingstoke, UK: Palgrave Macmillan, 2010.

Coliva, Annalisa, Anil Gupta, and Crispin Wright. "A Debate on Skepticism and Perceptual Belief." In *Empirical Reason and Sensory Experience*, edited by Ori Beck and Miloš Vuletić, 355–390. Cham, Switzerland: Springer, 2024 https://doi.org/10.1007/978-3-031-52231-4_45.

Davies, Martin. "Epistemic Entitlement, Warrant Transmission And Easy Knowledge." *Aristotelian Society Supplementary* 78, no. 1 (2004): 213–245. https://doi.org/10.1111/j.0309-7013.2004.00122.x.

Hacker, P. M. S. *Wittgenstein's Place in Twentieth-Century Analytic Philosophy.* Oxford: Blackwell, 1996.

Wittgenstein: Comparisons and Context. Oxford: Oxford University Press, 2013.

Hacker, P. M. S. , and G.P. Baker. *Wittgenstein: Rules, Grammar and Necessity: Essays and Exegesis of 185–242.* Oxford: Blackwell, 2009.

Jenkins, C. S. "Entitlement and rationality." *Synthese* 157, no. 1 (2007): 25–45. https://doi.org/10.1007/s11229-006-0012-2.

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

McGlynn, Aidan. "Epistemic Entitlement and the Leaching Problem." *Episteme* 14, no. 1 (2017): 89–102. https://doi.org/10.1017/epi.2015.63.

Moore, G.E. *Philosophical Papers.* London: Allen and Unwin, 1959.

Moyal-Sharrock, Danièle. *Understanding Wittgenstein.* New York: Palgrave-Macmillan, 2004.

"Beyond Hacker's Wittgenstein: Discussion of HACKER, Peter (2012) "Wittgenstein on Grammar, Theses and Dogmatism" Philosophical Investigations 35:1, January 2012, 1–17." *Philosophical Investigations* 36, no. 4 (2013): 355–380. https://doi.org/10.1111/phin.12021.

Pritchard, Duncan. "Wittgenstein's On Certainty and Contemporary Anti-skepticism." In *Readings of Wittgenstein's On Certainty*, edited by Danièle Moyal-Sharrock and William H. Brenner, 189–224. London: Palgrave Macmillan, 2005.

Epistemic Angst: Radical Skepticism and the Groundlessness of Our Believing. Princeton, NJ: Princeton University Press, 2015.

Rowe, William. "The Problem of Evil and Some Varieties of Atheism." *American Philosophical Quarterly* 16, no. 4 (1979): 335–341. https://www.jstor.org/stable/20009775.

Schönbaumsfeld, Genia. " 'Hinge Propositions' and the 'Logical' Exclusion of Doubt." In *Hinge Epistemology*, edited by Annalisa Coliva and Danièle Moyal-Sharrock, 94–109. Leiden, Netherlands: Brill, 2017. https://doi.org/10.1163/9789004332386_007.

"'Logical' and 'Epistemic' Uses of 'to Know' or 'Hinges' as Logical Enabling Conditions." In *Skeptical Invariantism Reconsidered*, by Genia Schönbaumsfeld, edited by Christos Kyriacou and Kevin Wallbridge, 235–251. London: Routledge, 2021.

Williams, Michael. *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism.* Oxford: Blackwell, 1991.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus.* Translated by D. F. Pears and B. F. McGuinness. London: Routledge & Kegan Paul, 1961.

On Certainty. Edited by G. E. M. Anscombe and G. H. von Wright. Translated by Denis Paul and G. E. M. Anscombe. Oxford: Blackwell, 1969.

Philosophical Grammar. Edited by Rush Rhees. Translated by Anthony Kenny. Oxford: Blackwell, 1974.

The Big Typescript: TS 213. Edited by C. Grant Luckhardt and Maximilian A. E. Aue. Translated by C. Grant Luckhardt and Maximilian A. E. Aue. Oxford: Blackwell, 2005.

Philosophical Investigations. 4th. Translated by G. E. M. Anscombe, P. M. S. Hacker and J. Schulte. Oxford: Wiley-Blackwell, 2009.

Wright, Crispin. "Warrant for Nothing (and Foundations for Free)?,." *The Aristotelian Society* 78, no. 1 (2004): 167–212. https://doi.org/10.1111/j.0309-7013.2004.00121.x.

"Wittgensteinian Certainties." In *Wittgenstein and Skepticism*, by Crispin Wright, edited by Denis McManus, 22–55. London: Routledge, 2004.

"On Epistemic Entitlement (II): Welfare State Epistemology." Chap. 10 in *Scepticism and Perceptual Justification*, edited by Elia Zardini and Dylan Dodd. Oxford: Oxford University Press, 2014.

Wright, Georg Henrik von. *Wittgenstein*. Minneapolis, MN: University of Minnesota Press, 1982.

Another Hello to the Third Man

Julian Torres, *University College London*

This essay presents some objections to a prominent solution to the Third-Man Argument (TMA) as it appears in the Parmenides due to Constance Meinwald (2022). According to this solution, the regress seemingly generated by the TMA can be blocked by distinguishing two kinds of predication: Ordinary Predication, which attributes properties to particulars, and Tree Predication, a special form of self-predication that specifies the nature or definition of a Form.

I argue that Meinwald's Tree Predication model ultimately fails. First, I argue that the distinction between Ordinary and Tree Predication does not successfully generalise to cases discussed in the Sophist, such as predications involving the Form of the Same, without reintroducing a version of the regress it was meant to avoid. Second, I argue the Tree Predication model lacks adequate explanatory power: when multiple Tree Predications are available, the view yields explanatory overdetermination, weakening the original motivation for the account. Once I suggest some ways Meinwald may respond to these objections, I argue her resulting position either reintroduces a regress at the level of Forms or deprives the Tree-Predication model of its explanatory power and philosophical motivation.

1. **Introduction**

In the Sophist, the Visitor introduces a seemingly absurd thesis that 'young people and old late-learners' hold: *only the good is good and the man is man* (Sph, 251c). Surprisingly, it seems uncontroversial to everyone involved in the discussion that predications of the form 'The Man is man' are intelligible. But understanding the way in which Plato thought of these instances of so-called Self-Predication has proven to be no easy task. It has also proven to be crucial to reconstructing his views on the Third Man Argument of the Parmenides.

By an examination of evidence from the Sophist and the Parmenides, I evaluate one prominent view in the literature by Constance Meinwald of Plato's views on Self-Predication in the Parmenides. First, I reconstruct the Third-Man Argument (TMA) as found in the *Parmenides* and assess the success of Meinwald's proposal in avoiding the TMA's conclusion. After this, I raise some objections to the proposed solution. I argue that, although it correctly identifies Self-Predication as the premise to qualify in order to avoid the infinite regress of Forms that gives the TMA its argumentative weight, the model of Tree Predication that Meinwald defends is ultimately unsuccessful as a solution to the TMA, and thus an inadequate reconstruction of Plato's own views. However, in order to get a clearer picture of the place of Meinwald's solution in the overall dialectic, some familiarity with Plato's theory and the challenge posed to it by the TMA is required.

Plato's forms play crucial roles in his metaphysics and epistemology. In general, they appear to be abstract entities, as opposed to sensible ones, which possess at least some subset of the following properties: they are causally responsible for and explanatory of the properties of sensible objects, they are the objects of knowledge, and grasp of them seems to lead to ethical improvements in our lives in some substantive sense (Harte, 2019, 421–422).

Now, the Third-Man argument, as traditionally construed, purports to undermine the coherence of this broad notion of the Theory of Forms. Thus, if successful, the TMA seriously compromises the status of one of the central pieces of Plato's broader philosophy. I will now reconstruct the argument, and then consider one suggested solution to it.

2. The Third Man

The Third-Man Argument is fairly short, and it goes as follows (*Prm*, 132a–132b):

> **P1. One-over-many**: Let $A_1, \ldots, A_n$ be some objects. If $A_1, \ldots, A_n$ are all F, then there is a thing, F-ness, such that F-ness is a unique Form in respect of which $A_1, \ldots A_n$ are all F (132a).
> **P2. Self-Predication**: For all Forms F-ness (Largeness, Smallness, etc.), 'F-ness is F' holds. (*Sph*. 258b9-c3).
> **P3. Non-Identity**: For all x, if x is F, then x is not identical with F-ness (This premise is supplied following Vlastos (1954, 325)).

Consider now the form of Largeness and a set of large things. From **One-over-many** and **Self-Predication**, it now follows that Largeness is large and that there is a unique Form of *Largeness* in respect of which Largeness and the members of the set are large. But from **Non-Identity** it follows that Largeness is not identical to the Form of *Largeness*. So Largeness and other large things must be large with respect to a new form, Largeness_1. By iterating the argument *ad infinitum*, it follows that:

> **C.** There is an infinite set of Forms Largeness, $Largeness_1, \ldots Largeness_n$, where $Largeness_k$ explains the property of being large for all $n < k$.

Informally, we have an infinite regress of Forms of Largeness. The regress seems vicious in at least one sense. If F-ness is meant to explain the features of things that are F, then one ought to first understand how F-ness_1 explains the features of F-ness, and so on. Thus, F-ness seems to never fulfil its intended explanatory role.

One option to reject C is to note that it follows from a contradiction: Self-Predication and Non-Identity are jointly inconsistent (Vlastos 1954, 326) so one (or both) of them must be rejected or qualified to not contradict each other. In what follows, I present a modification of Self-Predication by Meinwald (2022) that distinguishes between Ordinary Predication and predication of the form 'Largeness is large'.

3. Self-Predication

Consider the following instance of Self-Predication (*Sph*, 251c):

> M: The Man is a man.

On a natural reading of the sentence, 'The Man' refers to the Form in virtue of which all men are men. By **One-Over-Many**, if Plato and Aristotle are both men, then there is a Form 'the Man', and so on.[1] But, if we follow the natural reading suggested, the predicate term which applies to the Form 'The Man' is left unclear. Quite plausibly, Plato is not committing himself to the view that the form of Man-ness walks around Athens, having conversations with other men. The Man must be a man in some way different from how Aristotle is a man.

3.1. Ordinary Predication

It is worth clarifying Meinwald's position further. In her view, Self-Predication stands in contrast with **Ordinary Predication**. In cases such as 'x is F', where F is some property and x is some particular thing, F simply stands for a particular feature or property which x has. In the TMA, 'x is large' is **ordinary** when it just means what it says: that x is a large thing.

3.2. Tree Predication

On one construal (Meinwald 2022 and 1992), Forms stand in predication with themselves in virtue of a relation that differs from the relation that obtains between a subject and its property. On this view, Plato uses the Third Man argument to motivate a distinction that will be important in the second half of the

[1] And we want to prevent the need for a 'Third-Man' to make sense of the predication!

Parmenides, namely that of predication 'in relation to itself' as distinct from predication 'in relation to others'. The former is called **Tree Predication (TP)**, and the latter **Ordinary Predication (OP)**. When in the predication 'x is F,' x is the property corresponding to F-ness, then F in 'is F' does not express a property of x but a part of the definition of what it is to be x. Given that instances of predication in which concrete particulars figure seem to express properties instead of definitions predication 'in relation to itself' only occurs when x is a term referring to a Form.[2]

Applying the distinction to the Third Man Argument, Meinwald's view amounts to a rejection of Non-Identity when x is a Form and 'is F' is predicated of x; no third form is needed because 'Largeness is large' just specifies the nature of largeness.

The view strikes me as plausible in one important way: it maintains Self-Predication while retaining the intuitive principle of Non-Identity for Ordinary Predication between a particular thing and its properties. Indeed, it is evident throughout the *Sophist* that *there is* in fact a difference between the features which things have 'in virtue of their own nature' or 'in virtue of sharing in some other (Form)' (255e; 250c). When 'the apple is red' holds, the apple is certainly not identical with the *Redness* in virtue of which it is red. But in 'Redness is red', the subject and the form *are identical*. This gives us reason to think that a different kind of relation between subject and predicate holds in each case. I shall assess the plausibility of Meinwald's distinction by applying it to the discussion of kinds in the *Sophist* (250a–258d). If the distinction cannot account for the various ways in which certain Forms are said to stand in relation to themselves and other Forms, then we will have reason to doubt the success of the distinction in dealing with the TMA.

4. The *Sophist*

So far, the proposed solution to the TMA distinguishes Tree from Ordinary Predication, thereby weakening the Non-Identity premise of the argument and avoiding the infinite regress of forms. But now consider the following predication implicit in the *Sophist* (256a)[3]:

> **S.** The Same is the same.

Following Meinwald, we may want to distinguish in this case both senses of predication present in **S**. In this case (noting that 'x is the same' means x is the same as itself, i.e., x is self-identical), Plato seems to be committed to:

> **Ordinary Predication (OP$_S$)**: The Same is the same.

Note that the weakened form of Non-Identity which emerged from Meinwald's solution of the TMA applied only to Ordinary Predication. If this is right, then one can immediately see how Meinwald's solution fails to avoid the infinite regress in cases like OP$_S$. One may run a version of the TMA where The Same and a collection of other things (anything whatsoever, because everything is self-identical) are said to be 'the same', and an infinite regress of Forms The Same$_1$…The Same$_n$ will emerge to explain the relevant instances of Ordinary Predication. If instances of self-predication are meant to capture a definitional kind of predication for forms, Meinwald must explain cases in which forms, as entities, possess properties and may self-predicate in the **OP** sense.

However, Meinwald considers cases like **OP$_S$** to be unproblematic (Meinwald, 2022, 424–425). In her view, the solution to this new bolstered version of the TMA is simple. No new Form of The Same apart from the original one will emerge because there is no explanatory need for it. When various things, including the Form of The Same itself, are said to have the property of being the same (or self-identical), then this feature is explained by what she calls *Platonic Anaxagoreanism* (Meinwald, 2022, 411). According to this view, particular things' properties are explained by the features of their metaphysical constituents or

[2]One can define concrete particulars in classical predicate logic in a suitable model, but these definitions do not take the form of self-predication, so my point stands.

[3]To be sure, one may offer a solution to the TMA argument as it arises in the Parmenides only. Such a solution need not, presumably, account for instances of self-predication in other dialogues. Thankfully, Meinwald (2022, p. 427) makes clear that her model of Tree-Predication is intended as a model for instances of self-predication in the Sophist too. Thus, by her own lights, failure to avoid the TMA as it arises in the Sophist bears on the plausibility of her solution in the Parmenides too.

ingredients. In particular, if 'x is the same (as itself)' is true, then this is explained by x having a share of The Same. But, so the story continues, it follows from Self-Predication in Meinwald's interpretation that:

> **Tree Predication ($\mathbf{TP}_S$)**: The Same is the same as itself.

And it is precisely $\mathbf{TP}_S$ that explains why having a share of The Same makes something self-identical (Meinwald, 2022, 419–420). In other words, no extra form of The Same is needed because $\mathbf{TP}_S$ explains $\mathbf{OP}_S$. The form of The Same has the property of being self-identical in virtue of itself.

Note that this amounts to a rejection of Non-identity for certain cases of Ordinary Predication. However, Meinwald (Meinwald, 2022, 424–425) thinks she has a principled reason for doing so: Forms have explanatory roles which no sensible particular can perform; it is fine for Forms to explain their own properties and those of other particulars because they are the only things which could play such roles.

But it seems to me that this response is not satisfactory. One could equally run a bolstered version of the argument where every subject of predication is itself a Form (for Forms are also self-identical). Consider the set of **OP** predications:

> **Ordinary:** The Same is the same, The Fast is the same, The Slow is the same.

In the case of **Ordinary**, the response from the unique explanatory and causal roles that forms perform is no longer available to Meinwald. Any of the Forms (The Fast, The Slow, or The Same) can play the relevant explanatory role if being a Form is the only requirement that matters. Indeed, if Meinwald is to assert that the Form of The Same has a distinct status *qua* Form to explain instances of the **OP** predicate: 'is the same', then more than merely meeting the requirement of being a form is needed.

Indeed, Meinwald does seem to think that instances of tree-predication for the Form of The Same provide the explanatory ground for the ordinary predication of objects which exhibit sameness (Meinwald, 2022, 421), and thus, she thinks that there is reason to think that the Form of The Same will take precedence in providing the relevant ground for the property of sameness which particulars exhibit.[4]

4.1. Explanatory Power

In any case, there are further arguments available to cast suspicion over Tree Predication's explanatory power. Consider the following cases of **TP** and the correlative cases of **OP**, all of which Meinwald accepts (1992, p.383):

> $\mathbf{TP}_B$: The Bear is omnivorous.
> $\mathbf{OP}_B$: Bear x is omnivorous.
> $\mathbf{TP}_X$: The Fox is omnivorous.
> $\mathbf{OP}_X$: Fox x is omnivorous.

Meinwald's view seems to be that each of the **OP** cases is explained by the relevant **TP** (assuming, of course, that Bear x has a share of bear-hood): The bears in the Yosemite Park eat plants and meat because part of what it is to be a bear is to eat those things, and the same goes for foxes. But it also seems to follow, by **One-Over-Many**, that there is a common property which both foxes and bears possess by having a share in some Form of Omnivorousness, and which in turn explains their feature. Now, Meinwald must explain which one of these Forms is meant to be the truly explanatory one of the properties the particulars have. She should explain which instance of Tree Predication grounds the sensible property:

> $\mathbf{TP}_X$, $\mathbf{TP}_B$

or

> $\mathbf{TP}_O$: Omnivorousness is omnivorous.

[4] I return to this issue in Section 5

Without a more complete story, Tree Predications seem to be too broad to give the precise explanatory ground required by Ordinary Predications. Instead, they result in explanatory overdetermination for sensible properties in cases where the grounding is clear—it seems plausible that Omnivorousness explains the diet of bears! More is needed to account for the explanatory priority of some forms over others when their Tree Predications overlap. Thus, I consider a way in which Meinwald may tweak her view to account for cases of overdetermination. The upshot of it is a much weaker case for accepting the Tree Predication solution.

5. Another regress?

Consider again the case above. In order to account for the priority of Omnivorousness in providing the explanatory ground for $\mathbf{OP}_X$ and $\mathbf{OP}_B$, Meinwald may ask us to run the following test. Whenever more than one Tree Predication seems a plausible candidate to explain the same instance of Ordinary Predication, then one must choose the most specific one to the nature of the respective form. Consider:

> $\mathbf{TP}_B$: The Bear is omnivorous.
> $\mathbf{TP}_O$: Omnivorousness is omnivorous.

In this case, we must choose $\mathbf{TP}_O$, for being omnivorous is the most specific **TP** possible for Omnivorousness, whereas it is not for The Bear. Being omnivorous fully specifies the nature of Omnivorousness, whereas many other **TP**s must be predicated of the Bear for a full definition. Bears share the feature of being omnivorous with other species distinct from them, and it seems plausible that a full definition of bear-hood ought to reflect that fact by providing distinctions that are fine-grained enough and do not assign the same definition to bears as it does to any other omnivorous animals. For instance, it seems that **TP**s specifying that The Bear is mammal, of the family *Ursidae*, and so on will belong to a full definition of The Bear via **TP**s. However, a simpler view is possible. By iterating the definitional process above, it seems one may reach a maximally specific instance of **TP** for the form of The Bear which no other animal form may share, namely:

> $\mathbf{TP}_{Bear}$: The Bear is a bear.

Let instances of **TP** that are maximally specific like $\mathbf{TP}_{Bear}$ be known as **Maximal Tree Predication (MTP)**. The upshot of this modified version of the view with **MTP**s may go as follows. The required explanatory ground for any **OP** of the form 'the x is F' is provided by the **MTP** of F-ness (i.e., the F is F). And so it seems we now obtain the required ground: Bear x is omnivorous in virtue of Omnivorousness, whose nature is just to be omnivorous.

Now, it seems that the source of the explanatory overdetermination in cases like the above is that 'The Bear is omnivorous' holds in virtue of a form distinct from it. Even though it is in the definition of The Bear to be omnivorous, this fact does not provide the required explanatory ground. We require a second form. To bring out the consequences of this version of Meinwald's view, consider a bolstered version of the Third Man argument where every subject of predication is a form. Let the subscript '*' stand for an instance of **TP**:

> **P1. One-over-many**$_*$: Let $A_1, \ldots, A_n$ be some forms. If $A_1, \ldots, A_n$ are$_*$ all F, then there is a form, F-ness, such that it is a unique Form in respect of which $A_1, \ldots\ A_n$ are$_*$ all F...
>
> And so on.

In order to stop this regress, a similar response is available to Meinwald: no Third Form arises from the argument because F-ness is $_*$ F for it is an instance of (**MTP**) in a different way to how X-ness is$_*$ F.

Now, this response may be successful in stopping the regress, but it appears to undermine the original motivation to accept Meinwald's account of Self-Predication. On her view, instances of **TP** ground the sensible properties of particulars because they specify the metaphysical ingredients of the Forms. Yet, the needed modifications to her view via **Maximal Tree Predications** do away with the explanatory power of any instance of **TP** that is not **MTP**. If Platonic Anaxagoreanism is right, it no longer seems that the metaphysical constituents of forms are specified by the **TP**-model, and the view that Self-Predication is Tree Predication loses much of its plausibility. Unlike in Meinwald's original proposal, on the **MTP**-model, instances of self-predication do not capture part of the definition of what it is to be some form, and so most instances of **TP** seem both philosophically unmotivated and explanatorily inert. If this is right, most of Meinwald's **Tree-Predication** model can be done away with.

6. Conclusion

In sum, I assessed a proposal to solve the TMA by Constance Meinwald in light of a passage of the *Sophist* that features instances of Ordinary Predication for forms which may be problematic for her solution. I offered some objections to her view and suggested ways in which her view may be modified to avoid them. I have argued that the modifications required undermine the original motivation for her solution and make much of the structure it required for her solution unnecessary.

Bibliography

Cooper, John, ed. 1997. *Plato: Complete Works*. Indianapolis: Hackett.

Harte, Verity. 2019. "Plato's Metaphysics." In *The Oxford Handbook of Plato*, edited by Gail Fine, 2nd ed. Oxford: Oxford University Press.

Meinwald, Constance. 1992. "Good-bye to the Third Man." In *The Cambridge Companion to Plato*, edited by Richard Kraut. Cambridge: Cambridge University Press.

Meinwald, Constance. 2022. "Another Goodbye to the Third Man." In *The Cambridge Companion to Plato*, edited by David Ebrey and Richard Kraut. Cambridge: Cambridge University Press.

Vlastos, Gregory. 1954. "The Third Man Argument in the *Parmenides*." *The Philosophical Review* 63 (3): 319–49. https://doi.org/10.2307/2182692.

Aporia

Undergraduate Journal of the St Andrews Philosophy Society

VOLUME XXVI

Aporia is funded by the University of St Andrews Philosophy Society, which receives funds from the University of St Andrews Department of Philosophy, the University of St Andrews Students' Association, and independent benefactors.
Aporia is published by The University of St Andrews Philosophy Society.

`aporia@st-andrews.ac.uk`

Visit `https://ojs.st-andrews.ac.uk/aporia` to learn more.

Aporia
School of Philosophy
Edgecliffe, The Scores
St Andrews, Fife
KY16 9AL
Scotland