Everywhere Relations

Citation

Permanent link

Terms of Use

Share Your Story

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of South Asian Studies

have examined a dissertation entitled

Everywhere Relations

presented by **Allison Aitken**

candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.

Signature _____

Typed name: Prof. Parimal G. Patil (Co-chair)

Signature _____

Typed name: Prof. Leonard W.J. van der Kuijp (Co-chair)

Signature _____
Jeffrey McDonough (Jun 15, 2020 08:15 EDT)

Typed name: Prof. Jeffrey McDonough

Signature _____

Typed name: Prof. Alison Simmons

Date: June 12, 2020

EVERYWHERE RELATIONS

A dissertation presented by

Allison Aitken

to

the Department of South Asian Studies

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the subjects of South Asian Studies and Philosophy

Harvard University

Cambridge, Massachusetts

June 2020

Dissertation Advisors: Parimal Patil and Leonard van der Kuijp   Allison Aitken

EVERYWHERE RELATIONS

Abstract

This dissertation explores non-standard theories of relations and dependence structures within two historical contexts—medieval South Asia and Early Modern Europe. It includes selected papers from two separate but conceptually linked projects: the first is a defense of Madhyamaka Buddhist metaphysical indefinitism; the second is an interpretation of the Lockean person as a relation.

Part I explores the possibility that *everything* is ontologically dependent on something else. Mādhyamika Buddhist philosophers claim just that. I analyze the anti-foundationalist "neither-one-nor-many argument" of the ca. seventh/eighth century Indian Mādhyamika Śrīgupta in his *Commentary on the Introduction to Reality* (*Tattvāvatāravṛtti*). As I show, this argument rejects the possibility of ontologically independent entities by rejecting the possibility of mereological simples, both material and mental. I argue that Mādhyamikas like Śrīgupta are committed to a position I call "metaphysical indefinitism," and I make a case for its internal consistency and identify its virtues.

In Part II, I examine Locke's unintuitive yet philosophically promising account of relations, which has been neglected thus far, despite the fact that it bears on nearly all aspects of his system, including his influential theory of personal identity. I argue that for Locke the person is, metaphysical speaking, a relation. With this account, I shed light on a historically overlooked distinction between the Lockean self and the Lockean person. I

further show how a relation-interpretation of the Lockean person yields significant

metaphysical and epistemological payoffs.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

This dissertation would not have been possible without the support, guidance, and care of so many individuals—my mentors, teachers, colleagues, friends, and family members. I am fortunate to have benefitted from the guidance of an incredible team of committee members. Parimal Patil and Leonard van der Kuijp co-advised me with a consummate combination of expertise and kindness throughout the past decade, from an undergraduate on through to the completion of my doctoral studies. I thank Parimal for his always open door, for innumerable hours guiding me through Sanskrit texts and mapping arguments, and for forging the path for me to pursue the study of Buddhist philosophy *as* philosophy in a program bridging the Departments of South Asian Studies and Philosophy. I thank Leonard for his generosity with his encyclopedic knowledge, for countless hours guiding me through Tibetan texts, for modeling and expecting the highest standard of philological rigor, and for the special care he took in mentoring and encouraging me. I had the great good fortune to benefit from the guidance of a second pair of dream-team mentors, Alison Simmons and Jeff McDonough, meetings with whom I could always count on for greater clarity and new insights. I thank Alison for igniting in me a love of Early Modern European philosophy and for investing in me and inspiring me as a philosopher, teacher, and writer. I thank Jeff for his always thoughtful guidance and advice, for his transformative feedback on my writing, and for helping me to see old puzzles in a new light.

I am grateful to have received advice and comments on sections of this dissertation from a number of esteemed scholars. I am particularly indebted to Jay Garfield for his many incisive comments on multiple chapter drafts as well as his invaluable mentorship. Selections of this project have also benefitted from comments at various stages of development from Donald Ainslie, Katherine Dunlop, Michael Della Rocca, Gideon Rosen, and Jan Westerhoff, each of whom is owed my sincere appreciation. Some of the material of this dissertation was improved on the basis of feedback received at presentations at the Berlin-Groningen-Harvard-Toronto Workshop on Medieval and Early

Buddhist philosophy, without which this dissertation would not be possible. I am also deeply thankful to H. E. Zurmang Gharwang Rinpoche for his illuminating teachings and for his immense generosity in providing me a home in Sikkim, India, as well as the opportunity to study at the Zurmang Shedra, the Institute for Higher Buddhist Studies, where I especially benefitted from classes with Khenpo Karma Rigzin. And I am forever grateful to Geshe Acharya Thubten Loden, a guiding light as a scholar-practitioner who urged me to return to academia and pursue a PhD.

I am profoundly grateful to my parents for their unconditional and selfless love and support, and for encouraging me to pursue a meaningful life, whatever unconventional places that led to. I am thankful as well to my brother, Adam, for modelling the curiosity and open-mindedness that prompted me to wander off the well-trodden path for sources of insight. I also wish to express my gratitude for the shining example of my late grandmother, Lois Gatchell, who, with her singular blend of wit and wisdom, inspired me as a writer, communicator, and human being.

Finally, my greatest debt of gratitude is owed to my husband, Daniel, for his unwavering love and support and his unending patience and encouragement, without which I surely would never have crossed the finish line. I thank him for innumerable insight-filled discussions, for holding my hand through the peaks and the valleys, for the greater meaning of our shared journey, and for happiness.

INTRODUCTION

It's a safe bet that most of us—prereflectively at least—take ourselves to be living in a world of *things*, some animate and others inanimate. If you survey the room for *things*, chances are you'll immediately identify a whole host of them. You might find a panoply of possessions and persons—maybe furniture and books, friends or family members. If, however, I ask you to survey your surroundings for *relations*, you may have a harder time. Perhaps you'll notice that the table is *bigger than* the chair or your friend is *taller than* your dog. But proportional relations such as these only scratch the surface when it comes to the variety of relations that structure our world.

In fact, John Locke observes that "when attentively considered," *all* of our ideas of things—and ultimately the things themselves—include some kind of relation (II.xxi.3).[1] Consider, for instance, your dining table. It involves a mereological relation of whole in respect of its parts (a top and some legs). And the table bears a constitution relation with the wood that makes it up. It is also the result in a causal relation in respect of the agent that produced it—perhaps a factory or a carpenter. The table is also a member of a correlative pair with the chair needed to sit at it. The dining table also stands in some kind of functional relation with the meal you cooked and laid out for supper. And it even stands

---

[1] John Locke, *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch (Oxford: Oxford University Press, 1975). References to the *Essay* are by book, chapter, and section.

in a relation with the human minds that agreed to treat a certain arrangement of wood as a <dining table> using this specific conceptual-linguistic convention.

Not only are ordinary objects *involved in* a vast number of relations but, according to Locke, more of our ideas than we might have supposed actually *are* ideas of relations. And by implication, the referents of these ideas out in the world are themselves relations. As Locke has it, an idea of a relation is any idea that has direction, that draws our mind beyond the immediate referent to some secondary referent in respect of some ground for consideration. The idea of my niece, for instance, implicates not only the girl who is the immediate referent but secondarily implicates my sister too. Likewise, the idea of a friend or enemy secondarily implicates the person who holds that individual with affection or animosity. And the idea of a lid implicates not only the piece of metal containing the steam rising from my boiling soup but also secondarily implicates the pot on which it sits. In fact, any object that I think of as my <possession> not only implicates some hunk of matter but also secondarily implicates me as its possessor.

One of Locke's most important and overlooked insights is the pervasive, though often tacit, directionality of our ideas. For the most part, we do not simply experience things in and of themselves; we experience things in such a way that other things are implicated. In fact, for Locke, all objects of knowledge are relations, for knowledge consists in the agreement or disagreement of ideas, which corresponds to his definitions of two of the most fundamental relations: identity and diversity.

Locke is not alone in recognizing the ubiquitous role of relations in our mental lives and in the world itself. Going a step even further than Locke's already surprising suggestion,

Mādhyamika Buddhist philosophers observe that, upon inspection, everything stands in an *indefinite* number of relations. Moreover, these philosophers share Locke's well-known suspicion of the ontological category of substance, which is somewhat ironic given Locke's reference to the "poor *Indian* Philosopher" who, had he "but thought of the word *Substance*, he needed not to have been at the trouble to find an Elephant to support [the earth], and a Tortoise to support his Elephant" (I.ii.19). Here, Locke gestures sarcastically to the futility of merely stipulating the category of substance to serve as a metaphysical foundation.[2] Indeed, Mādhyamika philosophers take skepticism about substances as their starting point and develop an entire anti-foundationalist metaphysics out of it.

Mādhyamikas observe that, not only do we ordinarily take ourselves to live in a world populated by *things*, but, on a kind of naïve metaphysical interpretation, we take these things to be substantial, independent *unities*. My dining table certainly seems to be *one* distinct and independent object. And surely, *I* am *one* distinct and independent individual. These philosophers agree with the intuition that true unity and ontological independence go hand-in-hand. But they point out that anything that has parts is ontologically dependent on those parts, just as a table is on some slabs of wood and a molecule on some atoms. And, as it happens, *everything* has parts. Even those things that seem to be indubitably independent, fundamental, and simple—like some minute quanta of matter or our phenomenally unified conscious experience—are always analyzable into further parts standing in yet further relations to one another, or so Mādhyamikas argue. And the claim is not simply that

---

[2] Mādhyamikas would instead be in agreement with Locke's "intelligent *American*" who, upon enquiring "into the Nature of Things, would scarce take it for a satisfactory Account, if desiring to learn our Architecture, he should be told, That a Pillar was a thing supported by a *Basis*, and a *Basis* something that supported a Pillar" (I.ii.20).

everything is involved in an indefinite number of *trivial* relations; rather, everything depends for its very existence on some parts, which in turn depend for their own existence on some parts, *ad indefinitum*. In the final analysis, we never arrive at any independent unities at bedrock. As it turns out, then, it's relations (not tortoises) all the way down.

This dissertation is comprised of a collection of papers that explore non-standard theories of relations and dependence structures in the two historical contexts mentioned above: medieval South Asia and Early Modern Europe. The two papers comprising part I take as their starting point the anti-foundationalist "neither-one-nor-many argument" (*ekānekaviyogahetu*) of the Indian Madhyamaka Buddhist philosopher Śrīgupta (ca. seventh/eighth century) and explore the implications of the argument for the metaphysical dependence structure of the material world and the mind. I argue that this structure is best characterized as a form of "metaphysical indefinitism." The two papers comprising Part II explore Locke's account of the unusual structure of relations, specifically as it bears on his influential theory of personhood. Here I argue that, according to Locke, a person is—metaphysically speaking—a relation. The papers within each part are complementary but free-standing, so there is some repetition in the set-up of the neither-one-nor-many argument in the papers of Part I and the relation-interpretation of persons in the papers of Part II, but each approaches the set-up from a different angle. And although the two parts of this dissertation represent distinct studies, they are conceptually linked by their shared emphasis on the significant and underappreciated role that relations play in our world.

It's easy to see how certain things depend for their existence on other things: an army can't exist without soldiers, nor a molecule without atoms. But what would the world be like if *everything* depended for its existence on something else? Is an unending structure of ontological dependence relations even coherent? Historically, Western philosophical traditions have predominantly endorsed metaphysical foundationalist positions, yet Buddhist philosophy provides a rich source of anti-foundationalist arguments that merit broader attention. Part I of this dissertation centers on one such anti-foundationalist argument, the "neither-one-nor-many argument," as formulated by the Indian Buddhist philosopher Śrīgupta (ca. seventh/eighth century) in his *Introduction to Reality (Tattvāvatāra)*.

As a Mādhyamika Buddhist, Śrīgupta's central commitment is that nothing has ontologically independent being (*niḥsvabhāvatā*), and by implication, then, everything depends for its existence on something else. He recognizes a tight connection between ontological dependence and mereological dependence: a thing is ontologically independent only if it is mereologically simple. That's because anything that has proper parts depends for its existence on the existence of those parts. Śrīgupta argues that nothing is ontologically independent because there are no simples. While someone like Leibniz rejects material simples only to appeal to mind-like simple substances to ground the multitude, Śrīgupta insists that the existence of any kind of simple—whether material or immaterial—is metaphysically impossible.

The two papers that make up Part I explore the implications of Śrīgupta's neither-one-nor-many argument on different domains. In the first paper, "No Unity, No Problem:

Madhyamaka Metaphysical Indefinitism" (Chapter 1), I provide a general overview of the argument and take up the question of what a world devoid of fundamentalia would look like. With the help of Śrīgupta's influential account of conventional reality (*saṃvṛtisatya*), I argue that Mādhyamikas like Śrīgupta are committed to a kind of metaphysical indefinitism, and I make a case for its internal consistency and identify its virtues.

I clarify that Madhyamaka dependence chains are indefinite rather than straightforwardly infinite inasmuch as they are potentially and mind-dependently infinite, rather than actually and mind-independently infinite. I delineate the structural properties of the Madhyamaka metaphysical dependence relation, showing how it deviates from standard accounts of the ontological dependence relation, proper parthood relation, and metaphysical grounding relation, each of which are commonly characterized as asymmetric, irreflexive, and transitive. The Madhyamaka metaphysical dependence relation is indeed irreflexive. After all, something being self-grounding is tantamount to having ontological independence. But since, for Mādhyamikas like Śrīgupta, nothing is ontologically independent, nothing is ontologically prior to or more fundamental than anything else. Thus, Madhyamaka dependence relations do not honor the strict asymmetrical metaphysical priority of one relatum to the other, which—given irreflexivity—also prevents them from honoring transitivity. On a second level of analysis, however, I show how Mādhyamikas like Śrīgupta can accommodate a contextualist form of asymmetry, which, together with his revisable theory of conventional truth, will deliver significant payoffs for the view including its capacity to accommodate developments in scientific explanations without compromising his final ontology.

Metaphysical indefinitism structures not only to the material but also to the mental world. In the second paper, "Can a Mind Have Parts? Śrīgupta on Mental Mereology" (Chapter 2), I reconstruct and assess the lengthiest and most complex subargument of Śrīgupta's neither-one-nor-many argument, which is devoted to his refutation of mental simples and thus the rejection of ontologically independent mental entities. This argument turns on an analysis of the relation between the mind and mental content, and it results in one of the most radical skeptical positions on the unity of consciousness in the history of philosophy. There is a common line of thought that says we are entitled to infer the metaphysical unity of the mind from the phenomenal unity of our conscious experience. But Śrīgupta's sustained attack on the unity of even a momentary mental state calls into question the warrant for taking the simplicity of the mind for granted.

To be sure, the rejection of fundamental true unities, together with the endorsement of ontological dependence relations *ad indefinitum*, paints an unintuitive picture of the structure of reality. But engaging with non-standard metaphysical models of this kind may be instructive for challenging commonly held assumptions and for revealing possible unexplored avenues within gridlocked problem spaces of longstanding questions in metaphysics.

II. *Locke on Relations and Persons*

Part II comprises two papers that center on Locke's account of the metaphysics of relations, persons, selves, and consciousness. Locke's unintuitive, yet philosophically promising, account of relations has been neglected, despite the fact that it bears on nearly

all aspects of his system, including his influential theory of personal identity. I argue that, according to Locke, the person *is*—metaphysically speaking—a relation.

Commentators are currently engaged in a lively debate about the ontological status of the Lockean person: is it a substance or a mode? I argue for the neglected alternative: the Lockean person is a relation. There's a reason that this interpretation has been overlooked. It sounds odd to us today to say that a person is a relation. But Locke does not mean the same thing by "relation" that we do. And the structure of Lockean relations has not been well understood. Many of Locke's stock examples of ideas of relations—like friend, enemy, father, and son—appear to contemporary readers not to be ideas of relations at all, but of *things* that stand in relations to other *things*. For instance, most of us take a father to stand in the relation of fatherhood to his child, without taking fathers themselves to be relations. Locke, however, claims just that: fathers *are* relations. That's because, according to Locke, an idea of a relation consists in the consideration of one thing vis-à-vis some other thing in respect of some ground for comparison. Accordingly, the idea of a father is an idea of a relation because it consists in the consideration of a man vis-à-vis a child in respect of generation.

A proper analysis of Lockean relations helps us see that—against the tide of past scholarship—a person for Locke is best categorized not as a substance (like a body or a soul) or as a mode (like a number or an activity), but as a relation (like a father or a friend). A relation-interpretation of the Lockean person not only makes the best sense of the text but also yields an account of persons that delivers significant metaphysical and epistemological advantages.

The first paper in Part II, "Locke's Relational Account of Persons" (Chapter 3), begins with an analysis of the unusual structure of Lockean relations. I then present my relation-interpretation of Lockean persons, focusing first on how the idea of a person conforms to the structure of an idea of a relation, and next explaining how to understand a person as a relation ontologically. So what exactly does it mean to say that an idea of a person is an idea of a relation? What are the <person>'s relata and ground for comparison? Filling out this picture requires recognizing the distinction between the Lockean <self> and the Lockean <person>, which have standardly been understood to be first- and third-personal equivalents. Indeed, the conflation of the Lockean self and person has led to confusion on the part of interpreters who have sought to condemn, revise, or rehabilitate Locke's seemingly circular account of personal identity. But the <self> and the <person> are not strictly equivalent. For Locke, "person" is famously a forensic term, and it carries out the role in his system of linking a present subject of judgment with a past agent for the purposes of identifying the morally responsible party. It is the person—not the self—that is a diachronic forensic entity tracking moral accountability, and it is the self—not the person—that is the synchronic object of knowledge of the cogito. The idea of a person as an idea of a relation is (i) a way of considering the present self vis-à-vis some past self/selves in respect of diachronic identity as well as (ii) a way of considering the present self vis-à-vis some past action in respect of agency. And if the *idea* of a person is an idea of a relation, then by implication the person *is*—metaphysically speaking—a relation.

Understanding persons as relations also delivers important payoffs for Locke. I demonstrate how this metaphysically thin account of persons not only honors the non-

substantialist spirit of Locke's account of personal identity but also sidesteps Reid's failure of transitivity objection. That's because on the relation-interpretation, (i) persons *are* transitive relations, and (ii) they are relativized to the first-person, present perspective from which one makes judgments about one's own personhood. Moreover, I show how sameness of consciousness serves as a grounding criterion for persons without entailing circularity or resorting to a substantialist account of consciousness. Finally, I show how the relation-interpretation claims the principal advantages of the competing interpretations: (i) like the mode-interpretation, ideas of persons as ideas of relations are always real and adequate, making them suitable for Locke's demonstrative science of ethics; and (ii) like the substance-interpretation, persons as relations can possess agential power.

The metaphysics of persons has epistemic consequences, which I address in the second paper of Part II, "The Epistemological Payoffs of a Relation-Interpretation of the Lockean Person" (Chapter 4). Locke famously endorses a strong form of privileged access when it comes to first-personal judgments concerning one's own personal identity. He ought not, however, to endorse infallibility. Many interpretations of Lockean personal identity, including the popular "Memory Theory," commit him to infallibility. I argue that understanding the Lockean person as a relation (rather than a substance or a mode) affords an account of *qualified* privileged access in first-personal judgments of personal identity.

There are two metaphysical features of Lockean relations that jointly yield some significant epistemological payoffs for the view in terms of its capacity to support an account of *qualified* privileged access when it comes to knowledge of our own personhood. First, (i) Lockean relations are in some sense mind-dependent insofar as they involve the mental act

of comparing or considering one thing vis-à-vis some other thing in respect of some ground. Yet (ii) relations are also beholden to the existence of particular substances denominated by their relata, which act as truthmakers for our ideas of and judgments about relations and which also serve as metaphysical foundations for relations. (i) That persons as relations are mind-dependent honors the epistemic privilege of the first-person stance in personal identity judgments. Yet (ii) the fact that persons as relations must answer to the existence of substances as their foundations provides a basis for excluding infallibility. In this way, the relation-interpretation provides a non-circular basis for excluding infallibility, thereby preventing false memories from constituting persons. A relation-interpretation, I conclude, yields the strongest reading of Locke's influential account of personhood.

PART I

ŚRĪGUPTA ON UNITY AND BEING

In reality, all that exists within and without is devoid of independent being,
since all things are neither one nor many, like a reflection.

~ Śrīgupta, TA 1

If there is nothing truly one, *then every* true thing *will be eliminated.*

~ Leibniz, GP II 251/AG 176

NO UNITY, NO PROBLEM:

MADHYAMAKA METAPHYSICAL INDEFINITISM

It's easy to see how certain things seem to depend for their existence on other, more fundamental things, like a molecule depends on its constituent atoms or a gaggle on its geese. These cases illustrate the Hierarchy Thesis, which says that irreflexive, asymmetric, and transitive metaphysical dependence relations hierarchically structure reality. [3] Standard accounts of the ontological dependence relation, the metaphysical grounding relation, and the proper parthood relation all conform to the Hierarchy Thesis and honor this structure. Commitment to the rather intuitive Hierarchy Thesis is usually accompanied by commitment to another rather intuitive claim: the Fundamentality Thesis, which simply says that there must be something fundamental. In other words, hierarchical metaphysical dependence chains don't just go on forever; they terminate in something basic, something that is itself ontologically independent.

We can push on these two theses in various ways to arrive at different pictures of the structure of reality. For instance, if I hold the Fundamentality Thesis but reject the

---

[3] Bliss and Priest (2018a, 2) identify four theses to which the "standard view" of metaphysical foundationalism in the contemporary literature is committed: (1) the Hierarchy Thesis, (2) the Fundamentality Thesis, (3) the Contingency Thesis, which says that fundamentalia are merely contingently existent, and (4) the Consistency Thesis, which says that the dependence structure has consistent structural properties.

Hierarchy Thesis, I could still maintain a familiar metaphysical foundationalist view, which says that the structure of reality is flat and, strictly speaking, only fundamentalia exist (e.g., some forms of atomism or mereological nihilism). Alternatively, if I reject the Fundamentality Thesis but maintain the Hierarchy Thesis, then I wind up with an infinitist picture on which the world is structured by unending unidirectional dependence chains. In the absence of fundamentalia, the infinitist honors the extra structural property of extendability,[4] which says that everything depends on something else. If I reject both the Hierarchy Thesis and the Fundamentality Thesis, I might endorse a form of metaphysical coherentism, on which things stand in mutual, or symmetrical, relations forming dependence loops or circles of one kind or another.

The endless dependence chains and dependence circles that respectively populate infinitist and coherentist structures are commonly dismissed on suspicion of entailing a vicious regress or vicious circularity. Indeed, when it comes to the structure of reality, metaphysical foundationalist intuitions have dominated throughout the history of philosophy, and that's particularly true in the history of Western philosophy. Madhyamaka Buddhist philosophy, however, offers an arsenal of anti-foundationalist arguments that may be useful for at least calling into question the warrant for taking the pervasive metaphysical foundationalist intuition for granted. But I suggest that an even more promising potential payoff of taking Madhyamaka anti-foundationalist arguments seriously stands to be gained from an analysis of the metaphysical dependence structure that follows from such

---

[4] I follow Bliss and Priest in using the term "extendability" to describe the structural property that everything depends on something else. Although the term may seem to suggest that this is a modal property, that is not intended here. My thanks to Christopher Peacocke for this clarification.

arguments, which reveals that the three standard categories of metaphysical foundationalism, infinitism, and coherentism do not exhaust the possibilities for the structure of reality.

According to Mādhyamikas, the ultimate truth (*paramārthasatya*) says that all things lack ontological independence, and by implication, all things depend for their existence on something else. We might say, then, that the Madhyamaka central commitment is equivalent to the affirmation of the structural property of extendability, which we saw in the infinitist framework, and which said that everything depends on something else. Clearly, extendability rules out foundationalism, which demands some independent, ungrounded element(s), yet Mādhyamikas do not affirm a straightforward infinitism or coherentism either. Instead, I'll argue that they are committed to an alternative model that breaks the mold of this standard set of three categories—a structure I call "metaphysical indefinitism."

I begin by presenting a reconstruction and analysis of a Madhyamaka anti-foundationalist argument known as the "neither-one-nor-many argument" (*ekānekaviyogahetu*), as formulated by the Indian Mādhyamika philosopher Śrīgupta[5] in his *Commentary on the Introduction to Reality* (*Tattvāvatāravṛtti*, hereafter *Introduction to Reality*).[6] This

---

[5] The standard Tibetan account of the sequence of the Mūlasarvāstivāda preceptor lineage for monastic ordination that entered Tibet is as follows: Bhāviveka (ca. sixth century) → Śrīgupta → Jñānagarbha (early eighth century) → Śāntarakṣita (eighth century) → Kamalaśīla (late eighth century). See, for instance, Bu ston Rin chen grub's (1290–1364) *History of Buddhism*, *Chos 'byung gsung rab rin po che'i gter mdzod* (1989, 141b). See also 'Gos lo tsā ba Gzhon nu dpal's (1392–1481) *Blue Annals*, *Deb gter sngon po* (1971, 17a). I thus follow Seyfort Ruegg (1981, 67) in placing Śrīgupta roughly in the seventh century. This relative chronology has been challenged by several contemporary scholars, including Matsumoto (1978), Kobayashi (1992, 37–42), and Akahane (2003, 127), who argue that Śrīgupta postdates Śāntarakṣita. A treatment of this issue lies beyond the scope of this dissertation, though I argue elsewhere (forthcoming) that there is insufficient evidence to upend the Tibetan relative chronology.

[6] The original Sanskrit of the *Tattvāvatāra* (TA) and Śrīgupta's autocommentary, the *Tattvāvatāravṛtti* (TAV), are lost, and the root text survives only as embedded in the autocommentary, which is extant

argument constitutes the negative phase of the dialectic, which consists in the rejection of fundamentalia by way of the rejection of true unities. This first phase, then, works to demonstrate that the Fundamentality Thesis does not obtain.

I will then turn to the positive phase, the project of fleshing out the picture that follows from Śrīgupta's argument when taken together with his threefold criterion for conventional reality (*saṃvṛtisatya*). Here, I will present a two-stage model, first showing how Śrīgupta would reject the Hierarchy Thesis, and next showing how he could endorse it in a qualified form. Mādhyamikas like Śrīgupta are metaphysical egalitarians of a sort: in the final analysis, nothing is ontologically independent, and so, strictly speaking, nothing is metaphysically prior to or more fundamental than anything else. Thus, Madhyamaka dependence relations do not honor the strict asymmetrical metaphysical priority of one relatum to the other as demanded by the Hierarchy Thesis, which, given irreflexivity, also precludes transitivity. The resulting structure, characterized by irreflexivity and extendability, admits of dependence loops of greater than zero length (owing to irreflexivity) as well as dependence chains of indefinite length. And I underscore *indefinite* here by contrast with infinite, where indefinite signifies a potential, mind-dependent infinite, rather than actual, mind-independent infinite.[7] On a second level of analysis, however, I will show

only in Tibetan. See Ejima (1980) for a Japanese translation of the root verses of the TA, and see Kobayashi (1992, 1994) for a Japanese translation of the TAV. All citations of the text refer by page number to the Bstan 'gyur Dpe bsdur ma edition (PD), vol. 116, text no. 3121, and all verse numbering follows my forthcoming critical edition and annotated translation of the TAV.

[7] Bohn (2018, 178 n. 38) argues that what Bliss and others term "infinitism" is more properly "indefinitism." I instead draw a distinction between two views where "infinitism" stands for a dependence structure that allows for dependence chains that are actually and mind-independently infinite and "indefinitism" allows for dependence chains that are potentially and mind-dependently

how the flexible ontology of Śrīgupta's Madhyamaka can support a contextualist form of the Hierarchy Thesis, which, together with his revisable theory of conventional truth, will deliver significant payoffs for the view, including its capacity to accommodate developments in scientific explanations.

1. The Negative Phase: Śrīgupta's Case against Foundationalism

In the history of Buddhist philosophy, versions of the neither-one-nor-many argument have been deployed to refute the existence of a variety of (purported) pseudo-entities, from atoms to universals. Śrīgupta grants the neither-one-nor-many argument a universal scope of application in his *Introduction to Reality*,[8] cementing its place in the Madhyamaka tradition, for which it came to be popularized as one of the five great arguments for establishing the Madhyamaka ultimate truth, which says that, in reality, nothing has ontologically independent being (*svabhāva*).[9]

---

infinite. Moreover, the indefinitism I propose as characterizing Madhyamaka is neither asymmetrical nor transitive, though it is irreflexive and extendable.

[8] Śrīgupta presents what appears to be the earliest extant fully developed formulation of the Madhyamaka iteration of the neither-one-nor-many argument. Śāntarakṣita's influential *Madhyamakālaṃkāra* (MA) is likely an expansion of Śrīgupta's TAV, with Śrīgupta taken by the Tibetan tradition to be the teacher of Śāntarakṣita's teacher, Jñānagarbha. Śrīgupta's neither-one-nor-many argument is prefigured in the writings of Nāgārjuna (ca. second century), the progenitor of the Madhyamaka philosophical tradition; see, for example, Nāgārjuna's RĀ 1.71 and ŚS 32ab. Nāgārjuna runs a loosely related argument against real wholes and parts in VP 33–39. For instance, RĀ 1.71abc: *naiko 'nekapradeśatvān nāpradeśaś ca kaś cana / vinaikam api nāneko . . .* (Hahn 1982, 30); "Due to having distinct parts, *x* is not a unity. There is nothing that is partless. In the absence of a unity, neither is there a multitude." (All translations are my own.) See also Āryadeva's (third century) CŚ 14.19: *tasya tasyaikatā nāsti yo yo bhāvaḥ parīkṣyate / na santi tenāneke 'pi yenaiko 'pi na vidyante //* (Lang 1986, 132); "Whatever object one examines, none has unity. Given that there is no unity, neither is there a multitude."

[9] According to Tibetan traditions, the five great Madhyamaka arguments for emptiness are: (1) the diamond sliver argument (Skt. *vajrakaṇahetu*, Tib. *rdo rje gzegs ma'i gtan tshigs*), which establishes the emptiness of independent being through rejecting four possible manners of production, i.e., production from self, other, both self and other, or without a cause, (2) the argument rejecting the production of

1.1. Independent Being and Mereological Dependence

Svabhāva, which I translate here as "independent being," literally means "own-being" or "being per se." [10] We might also characterize *svabhāva* as a kind of essential independence,[11] that is, a form of ontological self-sufficiency that belongs to something by

---

existents and nonexistents (*sadasadutpādapratiṣedhahetu, yod med skye 'gog gi gtan tshigs*), which is an analysis of whether or not results exist at the time of their causes, (3) the argument rejecting production from the four alternatives (*catuṣkoṭyutpādapratiṣedhahetu, mu bzhi skye 'gog gi gtan tshigs*), which rejects independent being through an analysis of both causes and results, showing that multiple results cannot arise from a single cause, nor a single result from multiple causes, nor multiple results from multiple causes, nor a single result from a single cause, (4) the so-called "king of reasonings," the argument from interdependent origination (*pratītyasamutpādahetu, rigs pa'i rgyal po, rten cing 'brel ba'i gtan tshigs*), which demonstrates that whatever originates in dependence on anything else lacks independent being, and (5) the neither-one-nor-many argument (*ekānekaviyogahetu, gcig dang du bral gyi gtan tshigs*). Only the argument rejecting production from the four alternatives is not prefigured in some way in Nāgārjuna's works. This listing of arguments is not without precedent in Buddhist India. Kamalaśīla presents five arguments in his *Illumination of the Middle Way* (*Madhyamakāloka*, MĀ), which accord with later Tibetan listings, although he does not explicitly refer to them as a systematized list of five great arguments. Atiśa (982–1054) describes a set of "four great arguments" (*gtan tshigs chen po bzhi*), including the neither-one-nor-many argument, in his *Lamp for the Path to Enlightenment* (*Bodhipathapradīpa*), excluding the argument rejecting the production of existents and nonexistents.

[10] Unless otherwise indicated, I translate *svabhāva* as "independent being" in order to highlight that the negandum in Śrīgupta's argument is best understood as something ontologically independent and fundamental. *Svabhāva*, however, has a broad and complex semantic range, as evinced by the range of English terms contemporary scholars have used to translate it; for example, "aseity," "essence," "intrinsic nature," "substance," "self-nature," "essential nature," "inherent existence," etc. As Nāgārjuna defines it in his MMK 15.2cd, *akṛtrimaḥ svabhāvo hi nirapekṣaḥ paratra ca /* (Ye 2011, 236); "Independent being is [i] non-accidental and [ii] does not depend on anything else." On this stanza, see also Garfield (1995, 221) and Siderits and Katsura (2013, 155). The first property of *svabhāva*, (i) *akṛtrimaḥ*, translated here as "non-accidental," conveys both that *svabhāva* is a kind of intrinsic nature and also that it is not a conceptual fabrication, viz. it is not merely conceptually constructed, as in the case of the fictional unity superimposed on an aggregate; it is not a mere being of reason or imagination. (ii) Nor does *svabhāva* depend on anything else for its reality, as in the case of an aggregate whose reality is derivative, or "borrowed" from its constituents. See Westerhoff's (2009, 19–52) discussion of the ontological, cognitive, and semantic dimensions of *svabhāva*; in the ontological dimension, Westerhoff distinguishes three senses of *svabhāva* based on Candrakīrti's (seventh century) *Prasannapadā*: essence-*svabhāva*, substance-*svabhāva*, and absolute *svabhāva*.

[11] As Tahko and Lowe (2016) explain it, if *x* is essentially independent, then it is part of the essence of *x* to be ontologically self-sufficient, i.e., *x* by its very nature does not depend for its existence on anything else.

its very nature.[12] There are commonly two necessary and only jointly sufficient criteria that the metaphysical foundationalist demands of foundations: (i) ontological independence and (ii) constituting a complete basis on which all other things depend. The Madhyamaka ultimate truth, then, as the universal negation of ontological independence is, in effect, the universal negation of foundations. To demonstrate that nothing is ontologically independent is to reject foundationalism.

In making their case against ontological independence, Mādhyamikas regularly make use of three kinds of dependence relations: causal, conceptual, and mereological. In his neither-one-nor-many argument, Śrīgupta focuses on the mereological dependence relation: his argument against ontological independence—and against foundations—turns on the universal negation of mereological independence. Śrīgupta takes up his foundationalist opponents' picture of the world, which is populated by composites and parts. Each part stands in a proper parthood relation with some composite, where *x* stands in a proper parthood relation with *y* iff *x* is a part of *y* and *x* is not equal to *y*. Each composite is mereologically dependent on its proper parts such that a composite exists only if its proper parts exist. Śrīgupta thus takes mereological dependence to be a species of ontological dependence. The mereological dependence relation is also plural inasmuch as one proper part alone is insufficient to constitute a composite; one goose does not make a gaggle.

---

[12] Insofar as the category of substance is commonly regarded as something that is unitary, independent, self-sufficient, and persisting through change, substantial being can be helpfully thought of as a correlate to independent being. However, given that substance (*dravya*) (along with God, universals, inherence relations, etc.) is one among many ontological categories taken by certain of Śrīgupta's philosophical opponents to have independent being, substantial being is more properly a subcategory of independent being.

Critical for his argument, Śrīgupta understands the mereological dependence relation to be "topic-neutral," that is, it applies to all kinds of things.[13] The composite-part relation is not limited to material things; anything that can be physically or even conceptually divided is partite—be it a molecule, a mongoose, or even a mind. To be sure, the claim that parthood is topic-neutral is not uncontroversial, but it should not seem so strange either.[14] After all, just because we might not be capable of physically dividing some minute bit of matter doesn't prevent us from identifying its parts (e.g., left side, right side, etc.). Likewise, a four-dimensionalist will find the conceptual division of a perduring object into its temporal proper parts to be perfectly reasonable, despite our inability to physically divide objects into temporal parts. And the proper parts of an abstract object like a Euclidean triangle may include its three sides and three angles. So too, Śrīgupta would argue, the proper parts of a mental representation of a chair, for instance, may include the represented seat and represented legs.[15]

It's important to keep in mind here that conceptual divisibility is not equivalent to conceptual distinction. For instance, one might think that a single thing called by two

---

[13] On issues concerning the topic-neutrality of mereology, see Johnston (2005), Varzi (2010), Donnelly (2011), and Johansson (2015).

[14] This claim is by no means unique to Śrīgupta; rather, he is engaging with a supposition common among his interlocutors from competing Buddhist schools of thought that whatever can be either physically broken down or conceptually analyzable into discrete parts is not ultimately real (*paramārthasat*), viz. does not have independent being. For instance, in AKB *ad* k. 6.4, Vasubandhu explains that the mark of something that is merely conventionally real (*saṃvṛtisat*), viz. exists by conceptual designation (*prajñaptisat*), as opposed to something that is ultimately or substantially real (*dravyasat*), is that the object in question is no longer cognized (i) once it has been either actually or conceptually divided into parts or (ii) once it has been conceptually abstracted from other properties.

[15] See, for instance, Leech (2016) on taking seriously (rather than just metaphorically) the mereological structure of Kantian representations.

different names (e.g., the same woman considered as a "sister" and "friend"), or a thing

and its essence (e.g., matter and extension), are conceptually distinct, though ontologically

identical. But neither the sister and friend nor matter and extension are conceptually

divisible from one another in the mereological sense; extension is not a conceptually

separable proper part of matter, nor vice versa.[16] And neither is conceptual divisibility

inclusive of the conceptual distinguishability of a formal aspect, as in, for instance, the

distinguishability of a mouth from its smile. Rather, x is conceptually divisible in the

mereological sense just in case there are conceptually isolatable proper parts ys that

compose x, such that x is the sum of the ys.

If whatever has proper parts mereologically depends on those parts, then only

something that lacks proper parts, viz. a mereological simple, can claim mereological

independence. And on Śrīgupta's definition, only something that lacks proper parts counts

as a true unity. Since mereological dependence is a species of ontological dependence,

mereological independence, viz. simplicity, is a necessary condition for ontological

independence and for fundamentality. In other words, true unity, or unity per se, is a

necessary condition for independent being, or being per se.

But if there *are* no mereological simples, or true unities, then there are no candidates

for independent being or fundamentality. As Leibniz puts it, "if there is nothing *truly one*,

---

[16] Someone like Descartes would, of course, maintain that the mind and thought are conceptually distinct, but not conceptually divisible, insofar as thought is the principal attribute, or essence, of the mind. Neither Śrīgupta nor his principal interlocutors would agree with this account of the relation between the mind and thought. A common account of the essence of mind in Śrīgupta's intellectual milieu would instead be reflexive awareness (*svasaṃvedana/svasaṃvitti*).

then every *true thing* will be eliminated."[17] Śrīgupta's neither-one-nor-many argument turns

on this very premise, for he argues that it is precisely because there are no true unities that

nothing has independent being. Nevertheless, it does not follow that there is *nothing* at all.

Śrīgupta affirms the *conventional* reality of all things. In section 2, we will return to the

question of what precisely it means for being and unity to have a merely conventional status,

but in order to first see how Śrīgupta arrives at his anti-foundationalist conclusion, we turn

to his neither-one-nor-many argument.

1.2. Śrīgupta's Neither-one-nor-many Argument

In the opening stanza of his *Introduction to Reality*, Śrīgupta lays out the central

inference of the neither-one-nor-many argument as follows:

In reality, [*s*] all that exists externally and internally [*p*] lacks independent being,
[*r*] due to lacking independent being that is either one or many, [*e*] like a reflection.[18]

Here, [*s*] denotes the subject of the inference (*pakṣa/dharmin*), [*p*] denotes the predicate

(*sādhya*), or property to be proved, [*r*] denotes the reason (*hetu*),[19] and [*e*] denotes the example

---

[17] GP II 251/AG 176.

[18] TA 1: *phyi rol nang na gnas 'di kun // yang dag tu ni rang bzhin med // gcig dang du ma'i rang bzhin nyid // bral ba'i phyir na gzugs brnyan bzhin //* (PD 3121, 101). Cf. MA 1: *bdag dang gzhan smra'i dngos 'di dag // yang dag tu na cig pa dang // du ma'i rang bzhin bral ba'i phyir // rang bzhin med de gzugs brnyan bzhin //* (Ichigō 1989, 190). MA 1 is preserved in Sanskrit in Prajñākaramati's BCAP: *niḥsvabhāvā amī bhāvās tattvataḥ svaparoditāḥ / ekānekasvabhāvena viyogāt pratibimbavat //* (Vaidya 1960, 173). For an English translation of MA 1, see Ichigō (1989, 191).

[19] Although Śrīgupta does not explicitly classify the neither-one-nor-many reason, it is commonly regarded as a reason from the non-observation of the predicate (*sādhya*), or literally, the non-observation of the pervader (*vyāpakānupalabdhi*), e.g., Kamalaśīla's (ca. 740–795) MAP *ad* k.1 (Ichigō 1985, 23) and corresponding Sanskrit in Haribhadra's (late eighth century) AAA (Wogihara 1932–1935, 624.5–7); see also Kamalaśīla's MĀ (Keira 2004, 235). This refers to the fact that the reason property is *not* observed wherever the predicate is *not* present. In other words, the property of being neither unitary nor non-unitary is *not* observed in any subject that does *not* lack independent being (that is, in any subject that *has*

23

(*dṛṣṭānta*) in which the entailment relation between the reason property and predicate is instantiated. The operator "in reality" (*yang dag tu, *tattvatas*) signals that this inferential argument involves analysis of the subject's ultimate nature and does not bear on its status as a conventional being or a conventional unity/multiplicity.[20]

There are two points to keep in mind about the subject of the argument, <all that exists externally and internally>. First, Śrīgupta takes this subject to be all-inclusive, signifying a universal domain. And it is worth noting that, although external and internal are understood relative to some cognitive agent (the agent of the inference in this case), the external-internal dichotomy should not be confused for mapping precisely onto the material-mental dichotomy. In fact, unlike the external-internal dichotomy, the material-mental dichotomy is not all-inclusive according to the Abhidharma ontology that is accepted by many of Śrīgupta's interlocutors. On Abhidharma presentations, some (but not all) mental things belong to the internal category, some (but not all) material things belong to the external category, and some things (like space) are neither mental nor material.[21]

---

independent being). The neither-one-nor-many reason is referred to as a "pervader" here since it "pervades," or exhausts all possible ways in which something could exist with independent being, viz. as a simple or as a composite of simples.

[20] The use of a qualifier in Madhyamaka argumentation is, of course, a contentious and much debated issue, which Tibetan doxographers subsequently used as a differentiating criterion for assigning Mādhyamikas to the so-called *Svātantrika and *Prāsaṅgika subschools. On this issue, see Dreyfus and McClintock, eds. (2003).

[21] Similarly, according to Vaiśeṣika ontology, a non-Buddhist view targeted by Śrīgupta, things like universals fall outside the mental-material dichotomy but *are* included in the internal-external dichotomy. The internal-external division looks more phenomenological than ontological, with objective features of experience designated as "external" and subjective features of experience designated "internal." In the descriptive ontology treatises of the Abhidharma genre, when categorizing the eighteen constituents (*dhātu*) that comprise the world into internal and external, the material sense organs (*indriya*) are

Second, "exists" (*gnas*) as part of the articulation of the subject is not ontologically loaded; it does not signify any particular mode of existence, realist or otherwise. It simply means "whatever there is." I will thus use "all things" as shorthand for the subject.

Formulated in the standard three-part inference of classical Buddhist logic, the argument runs as follows:

Thesis (*pratijñā*):

All things lack independent being.

Major premise, statement of the entailment between the reason property and predicate (*vyāpti*):

Whatever is neither one nor many does not have independent being.

Minor premise, predication of the reason property of the subject (*pakṣadharmatā*):

All things are neither one nor many.

This argument hinges on the reason property: neither one nor many. Śrīgupta, in effect, poses a destructive dilemma, what I will call the "One-or-many Dilemma": *if anything has independent being, then it is either one or many*.[22] He argues that nothing that can satisfy either

---

standardly counted together with the varieties of consciousness as "internal" (classified as "internal material sense bases," *ādhyātmikāni rūpīṇyāyatana*). Conversely, mental objects (*dharmadhātu*) are counted together with the five kinds of sense objects as "external"; see, for instance, Vasubandhu's AKB *ad* 1.20ab and Asaṅga's AS (Pradhan 1950, 71). In addition, a grab-bag variety of things are classified as neither-mental-nor-material (*cittaviprayuktasaṃskāra*). It is worth noting that, although Śrīgupta does not invoke this pair of categories, <material and immaterial> *would* denominate a universal domain, given its <F and not-F> structure. Atiśa, for instance, uses this pair (*gzugs can dang gzugs can ma yin pa*) as the subject of the neither-one-nor-many argument in his MU (PD 3148, 283).

[22] Notice that this dilemma is the contrapositive of the entailment relation (*vyāpti*), which is discussed below.

disjunct of the consequent, and therefore, by modus tollens, that nothing can satisfy the antecedent.

So what are the conditions for predicating one or many of a given subject? As noted, the operator or qualifier, "in reality," in TA 1 indicates that the terms "one" and "many" in these premises should be understood as *true* unity and *true* multiplicity, where to be a true unity is to lack proper parts, viz. to be a mereological simple,[23] and to be a true multiplicity is to have proper parts, the most basic of which are themselves true unities.[24] With these definitions, Śrīgupta stipulates a foundationalist structure on which the world bottoms out in mereological simples. We can thus revise the One-or-many Dilemma to state [One-or-many Dilemma*]: *if anything has independent being, then it is either one simple or many simples.*

Two further points about the one-many pair in this argument merit noting. First, the terms *eka* and *aneka* translated here as "one" and "many" conform to the logical, grammatical, and conceptual structure <F> and <not-F>, which is more precisely conveyed by the translation "unity" and "non-unity." Since this pair of predicates is mutually exclusive and exhaustive, <*eka* or *aneka*> is an exclusive disjunction. If *x* has independent being, then, on pain of violating the law of excluded middle, it is either a unity or a non-unity. As Śrīgupta states, "Since [unity and non-unity] are contradictory, existing

---

[23] This definition, which is implicitly operative in Śrīgupta's argument, is made explicit by Kamalaśīla in his subcommentary on Śāntarakṣita's MA, where he defines "unity" in the context of this argument as follows, MAP *ad* k. 1: *cig pa zhes bya ba ni cha med pa nyid do //* (Ichigō 1985, 23); "'Unity' refers to something that lacks parts."

[24] As Kamalaśīla states in his MAP: *cig shos zhes bya ba ni du ma nyid de tha dad pa nyid ces bya ba'i tha tshig go //* (Ichigō 1985, 23); "The alternative member of the pair is non-unity (*anekatva*), which is synonymous with 'consisting in discrete parts' (*bhedatva*)."

[with independent being] in any other manner is surely untenable" (TA 2cd).[25] In English, <many> is a vague predicate. The point at which a quantity becomes <many> is indeterminate (think the Sorites Paradox). <Many> is also context-dependent; what counts as <many> in some contexts might be <few> in another. Three shoes are too many for a human to wear at once and too few for a centipede. But in this argument, even a pair counts as <many>. That's because *aneka* signifies any determinate number greater than one. These definitions thus exclude an indeterminate status, as well as that of a merely conventional status, like the conventional unity of an aggregate, such as an army or a flock.

Nevertheless, the translation "one or many" has the virtue of more naturally reflecting a second feature of this disjunctive pair. When understood as the *true* unity and

---

[25] *'gal ba'i phyir ni rnam pa gzhan // yod par yang ni mi 'thad do //* (PD 3121, 101). To the same effect, Śāntarakṣita states in MA 62: *gcig dang du ma ma gtogs par // rnam par gzhan dang ldan pa yi // dngos po mi rung 'di gnyis ni // phan tshun spangs te gnas phyir ro //* (Ichigō 1989, 210); "Aside from unity and non-unity, an object's having some other classification is impossible, since it is established that these two are mutually exclusive." Śāntarakṣita elaborates on this point in MAV *ad* k. 1: *rang bzhin zhig yod par gyur na ni gcig pa'am cig shos las mi 'da'o // de dag ni phan tshun spangs te gnas pa'i mtshan nyid yin pas phung po gzhan sal bar byed do /* (Ichigō 1985, 22); "If something existed with independent being, there is no option apart from its being either a unity or the alternative [member of this disjunctive pair, viz. a non-unity]. Since these two [predicates of unity and non-unity] are definitionally mutually exclusive, existing [with independent being] in any other manner is excluded." Similarly, Kamalaśīla states in his MĀ (PD 3116, 1323): *gcig dang du ma nyid dag gis ni rnam pa thams cad la khyab ste / de dag ni phan tshun spangs te gnas pa'i mtshan nyid yin pa'i phyir ro //;* "Unity and non-unity encompass all alternatives, because that pair is characterized by being mutually exclusive." See also MAP *ad* k. 1: *gang gi ngo bo rnam par bcad pa med na yongs su gcod pa med pa gang yin pa de ni de spangs te gnas pa'i mtshan nyid yin te / dper na dngos po med pa rnam par bcad pa med na yongs su gcod pa med pa ltar gcig pa nyid dang du ma nyid gnyis kyang phan tshun gyi ngo bo rnam par bcad pa med na yongs su gcod pa med pa dang yin te / tha snyad kyis nges par byed do /de lta bas na phan tshun spangs te gnas pa'i mtshan nyid dag yin no //* (Ichigō 1985: 23); Sanskrit in AAA: *yad rūpavyavacchedanāntarīyaka-paricchedanaṃ hi yat tat tat parasparaparihārasthitalakṣaṇaṁ, tadyathā bhāvo 'bhāvavyavacchedanāntarīyakaparicchedaḥ. paraspararūpavya-vacchedanānatarīyakapariccheda evaikatvānekatve tasmāt parasparaparihārasthitalakṣaṇa iti* (Wogihara 1932–1935, 635.15–18); "Whatever [pair of properties] is such that F is necessarily determined (*nāntarīyakaparicchedana*) when G is excluded is characterized as mutually exclusive. For instance, the <existence> of *x* is necessarily determined whenever the <non-existence> of *x* is excluded. With respect to <unity> and <non-unity>, it is indeed the case that when one is excluded the other is necessarily determined. Thus, [this predicate pair] is characterized as mutually exclusive."

true multiplicity of a foundationalist structure, *eka* and *aneka* are a peculiar contradictory pair in that they share not only a conceptual priority relation but also a metaphysical priority relation: the *existence* of a non-unity presupposes the *existence* of some unities. A plurality presupposes singular things as its building blocks. Many and one, thus, stand in both an ontological dependence relation as well as a mereological dependence relation of a composite on its proper parts. These features of the relation between unity and non-unity set it apart from most other contradictory predicate pairs: while a concept like <non-blue>, for example, conceptually presupposes <blue>, the existence of something that is non-blue—say, a red apple—certainly does not require the existence of something blue.

Śrīgupta points to this metaphysical priority relation between one and many, or a unity and a multiplicity, stating, "Given that [a multiplicity] consists of many unities, if one [viz. a unity] does not exist, the other [viz. a multiplicity] is also impossible."[26] Since unity is metaphysically prior to multiplicity, if unity is rejected, then multiplicity is *ipso facto* precluded. Just as a forest cannot exist without trees, a composite of simples cannot exist without simples. And since a multiplicity depends for its existence on some unities, it is not a proper candidate for ontologically independent being after all. The One-or-many

---

[26] TAV *ad* k. 2b: *gcig mang po'i ngo bo yin pas de med na 'di yang mi srid pa . . .* (PD 3121, 102). Cf. MA 61: *dngos po gang gang rnam dpyad pa // de dang de la gcig nyid med // gang la gcig nyid yod min pa // de la du ma nyid kyang med //* (Ichigō 1989, 210); "Whatever object one analyzes, none has unity. Given that there is no unity, neither is there a non-unity." Here, Śāntarakṣita closely glosses Āryadeva's CŚ 14.19. See also MAV *ad* k. 61: *'di ltar du ma ni gcig bsags pa'i mtshan nyid do / gcig med na de yang med de / shing la sogs pa med na nags tshal la sogs pa med pa bzhin no //* (Ichigō 1985, 172); "Thus, 'non-unity' is defined as a composite of unities. If no unity exists, then neither does that [composite of unities] exist, just like if no trees exist, neither does a forest exist." Cf. also TS 1995: *tad evam sarvapakṣeṣu naivaikātmā sa yujyate / ekānispattito 'nekasvabhāvo 'pi na sambhavī //* (Saccone 2018, 171); "Thus, a unitary nature is inadmissible on all accounts. Since unitary [independent being] is not established, non-unitary independent being is also impossible." See Saccone (2018, 253) for an alternative English translation.

Dilemma is thus further simplified as follows [One-or-many Dilemma**]: *if anything has independent being, then it is a simple.* All Śrīgupta needs to do to establish that nothing has independent being, then, is to demonstrate that there are no simples. The argument, thus, reduces to a refutation of true unity.

At this point, let's pause to rephrase the argument in more explicitly anti-foundationalist terms. To be sure, Śrīgupta's definition of true unity is a strong one, but it is not at all controversial to suppose that a foundationalist would maintain that foundations are in a strong sense, well, *fundamental*, i.e., basic, primitive, ontologically independent, and metaphysically ungrounded. And it is not far-fetched to suppose that being fundamental in these senses is incompatible with being partite. Yet Śrīgupta's argument does not even require his foundationalist interlocutors to accept the convertibility of simplicity and fundamentality—only that they accept that there exist(s) *some* true unity/unities at bedrock. Śrīgupta's argument can be structured in anti-foundationalist terms as follows:

P1 If there are any foundations, then necessarily they either *are* true unities or bottom out in true unities.

P2 There are no true unities.

∴ C1 There are no foundations.

P3 A foundationalist metaphysical structure is possible only if there is some foundation.

∴ C2 A foundationalist metaphysical structure is not possible.

The real heavy lifting is, of course, done by P2, the rejection of true unities. To establish this premise, Śrīgupta's strategy is to tackle the domain of his universal subject by way of

three jointly exhaustive categories. He divides the world into material things, mental things, and the grab-bag category of whatever is neither material nor mental (e.g., abstract entities).

Śrīgupta presents a systematic and exhaustive argument in his *Introduction to Reality*, which is represented in condensed form in the reconstruction below. P1 is the contrapositive of the entailment relation between the predicate and reason property, and the remainder of the premises work toward establishing the predication of the reason property of the all-inclusive subject (C2 and C3).

Table 1: Neither-one-nor-many Argument Reconstruction

P1 If there is anything that has independent being, then it is either a true unity or a true non-unity.	*Contrapositive of the entailment relation between the reason property and the predicate*
P2 There is a true non-unity only if there are true unities. P3 Something is a true unity if and only if it is a simple, viz. has no proper parts. P4 Whatever exists is material or mental or neither-material-nor-mental. ∴ C1 If there is a true unity, it is either a mental simple or a material simple or a neither-material-nor-mental simple. (from P3, P4) P5 There are no material simples. P6 There are no mental simples. P7 There are no neither-material-nor-mental simples. ∴ C2 There are no true unities. (from C1, P5, P6, P7) ∴ C3 There are no true non-unities. (from P2, C2)	*Proof of the predication of the reason property of the subject*
∴ C4 There is nothing that has independent being. (from P1, C2, C3)	*Thesis*

As reconstructed here, the argument rests on the subarguments in support of P5, P6, and P7, which collectively reject the existence of any kind of simple. Although a thorough

treatment of this argument is beyond the scope of this chapter,[27] I will touch on Śrīgupta's

subarguments against true unities in each of these three categories.

Śrīgupta first targets material simples by taking up fundamental particles and posing

the following dilemma: if matter is constituted by fundamental particles, those particles are

either extended or unextended. (i) If extended, then fundamental particles could not be

simple, because whatever is extended is (at least conceptually) divisible (into, say, a right

side and a left side, etc.), and whatever is divisible has proper parts. But in that case,

fundamental particles would be composites themselves and could not be fundamental after

all.[28] (ii) On the other hand, Śrīgupta argues, if fundamental particles were unextended,

they could never constitute an extended hunk of matter. Since unextended particles could

not have spatially discrete sides at which to conjoin with neighboring particles, he reasons,

the entire composite would absurdly collapse into a single unextended point. Śrīgupta

summarizes the argument as follows:

> A fundamental particle could not be a [true] unity because an [extended] composite [of unextended particles] is impossible. This is because if they were unitary in nature, then facing [particles] would [absurdly] occupy a single location. Nor is it the case that fundamental particles possessed of some other kind of [extended] nature could face with one another, since in that case it would absurdly follow that [each fundamental particle] would be a manifold.[29]

---

[27] For a detailed reconstruction and analysis of Śrīgupta's neither-one-nor-many argument, see my forthcoming translation, critical edition, and philosophical introduction to the *Tattvāvatāravṛtti*.

[28] For contemporary arguments defending the coherence of extended simples, see Markosian (1998, 2004a, 2004b) and McDaniel (2007). See McDaniel (2003) for an argument against extended simples.

[29] TAV *ad* k. 2a: *rdul phra rab ni gcig pa nyid ma yin te / rang bzhin gcig pu de la mngon par phyogs par yul gcig na gnas pas na bsags* [D, C: *bstsags*] *pa mi rung pa'i phyir ro / rang bzhin gzhan gyis mngon du phyogs pa yang ma yin te / du ma nyid du thal bar 'gyur ba'i phyir ro //* (PD 3121, 102). Here, Śrīgupta follows in a long tradition of Buddhist anti-atomist arguments utilizing a similar strategy. For instance, Āryadeva argues in his CŚ 9.15: *gang la shar gyi phyogs yod pa // de la shar gyi cha yang yod // gang gi rdul la phyogs yod pa // des rdul rdul phran min par bsnyad //* (Lang 1986, 92; cf. Lang's translation 93); "Whatever has an eastern side has an eastern part. Whoever accepts that a particle has sides [viz. directional parts] must admit that it is not a

31

Śrīgupta concludes that, since there is no coherent story to be told about how material bodies are founded in either extended or unextended fundamental particles, there are no material simples.

Śrīgupta next argues that since fundamental particles have been rejected, whatever is purportedly founded in them—whether directly or indirectly—is thereby precluded:

> Accordingly, since [purportedly] fundamental particles in fact lack independent being, whatever is [held to be] constituted either directly or indirectly by them, such as substances that possess [particles as their] parts, as well as whatever is ontologically dependent on substances, like properties, action, universals, and so forth, are indeed soundly rejected.[30]

Once the foundations of a materialist foundationalist structure are rejected, Śrīgupta argues, the rug has been pulled out of the entire ontological framework.[31]

---

fundamental particle after all." Śrīgupta's argument here also follows Vasubandhu's line of reasoning in the Vś 12–13: *ṣaṭkena yugapadyogāt paramāṇoḥ ṣaḍaṁśatā / ṣaṇṇāṁ samānadeśatvāt piṇḍaḥ syād aṇumātrakaḥ // paramāṇor asaṁyoge tatsaṁghāte 'sti kasya saḥ / na cānavayavatvena tatsaṁyogo na sidhyati //* (Silk 2016, 15–17; see here also for an alternative English translation); "Since an atom conjoins simultaneously with six [other surrounding atoms], it would have six parts. [Or else,] since the six [atoms] would be co-located, the collection [of atoms] would be [reduced to the size] of a mere atom. [Vś 12] If atoms do *not* conjoin [with one another], then when there is a composite of them, what is that the [conjoining] of? And it is not the case that their conjoining [i.e., the conjoining of composites] cannot be proved by virtue of the fact that they are partless, [for composites have parts]. [Vś 13]" For an alternative translation of these stanzas together with a detailed reconstruction and analysis of the argument in Vś 11–15, see Kapstein (2001, chapter 7); on this argument, see also Oetke (1992), Arnold (2008), Kellner and Taber (2014), and Kellner (2017a). This section of Vasubandhu's argument, which targets the Vaiśeṣika account of atomism, also appears in AKB *ad* k. 43d2.

[30] TAV *ad* k. 3ab: *de ltar rtsom byed med pa'i phyir // rdzas la sogs pa thams cad bsal* [N: *gsal*] *// de lta bur rdul phran rang bzhin med pa nyid yin pas na de mngon sum dang / gzhan du brtsams pa yan lag can gyi rdzas dang de la brten pa dang / yon tan dang / las dang / spyi la sogs pa'ang ring du spangs pa kho na'o //* (PD 3121, 102). Cf. MA 14–15: *rdul phran rang bzhin med grub pa // de phyir mig dang rdzas la sogs // bdag dang gzhan smras mang po dag // rang bzhin med par mngon pa yin // de yi rang bzhin des brtsams dang // de yi yon tan de las bdag // de yi spyi dang khyad par yang // de dag de dang 'du ba can //* (Ichigō 1989, 194).

[31] Here, Śrīgupta references the Vaiśeṣika ontological categories (*padārtha*) of substances (*dravya*)—which claim fundamental particles as their basic parts—as well as properties (*guṇa*), action (*karma*), universals (*sāmānya*), particulars (*viśeṣa*), and the inherence relation (*samavāya*) between a substance and its properties,

Taking himself at this point in the dialectic to have rejected the possibility of material simples, and thus material foundations, Śrīgupta next targets a range of idealist foundationalist positions belonging to his fellow Buddhists from the Yogācāra tradition, by rejecting the possibility of a truly unitary mind or mental state. Śrīgupta's sustained attack on mental simples comprises his lengthiest subargument. In his rejection of material simples, Śrīgupta follows earlier Mādhyamikas like Āryadeva as well as Yogācāra Buddhists like Vasubandhu.[32] It is with the refutation of mental simples that Śrīgupta makes a unique philosophical contribution.

Since the subargument against material simples (i) targets a monadic subject, e.g., one purportedly fundamental particle, and (ii) relied on pairs of monadic properties <unitary> and <non-unitary>, and <extended> and <unextended>, we can think of that subargument as the "monadic phase" of the argument. The subargument against mental simples instead takes up a dyadic subject, the mind and mental content, or awareness (*jñāna*) and its mental representations (*ākāra*), and poses two dilemmas using two additional property pairs:

etc., all of which are indirectly founded in fundamental particles by virtue of ontologically depending on substances in one way or another.

[32] See Āryadeva's CŚ 9.15–17. Vasubandhu uses a version of this argument in Vś 11–15 in support of his thesis that everything consists in cognition-only (*vijñaptimātratā*). The work done by this section of his argument is, however, much disputed. For instance, Oetke (1992) argues that the argument applies only to objects of experience, and that it leaves open the possibility that material objects exist; Arnold (2008) contends that this subargument is intended to establish metaphysical idealism; Kellner (2017b) instead argues that this section must be understood within the argumentative context of the entire text, which represents an *argumentum ad ignorantiam*, and that Vś 11–15 falls under the section in which scriptural testimony (*āgama*) is precluded from serving as a means by which we can reliably gain knowledge of the existence of external objects.

1. Real/unreal (*satya/alīka*) Dilemma: mental content is either real in the same way that the mind is supposed to be or it is an unreal figment.

2. Distinct/non-distinct (*bheda/abheda*) Dilemma: the mind and mental content taken together are either one identical thing or distinct things.

Figure 1: Range of Views on the Mind and Mental Content
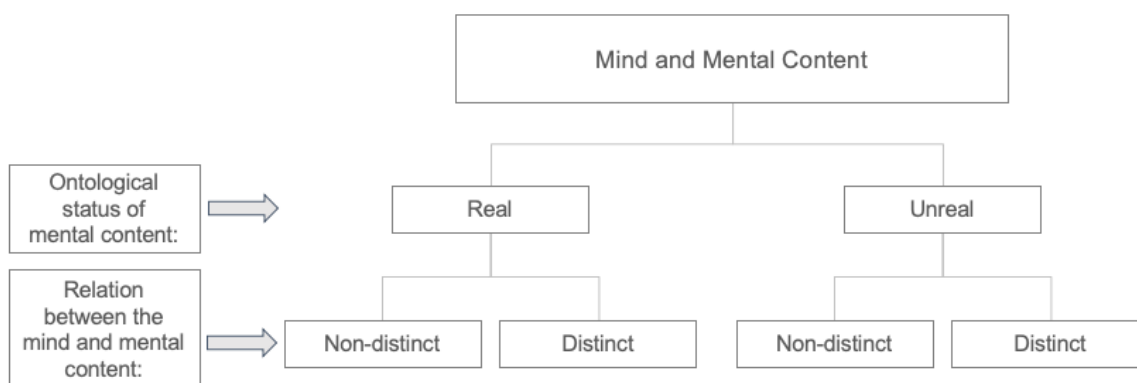
Śrīgupta runs a multitiered argument from dilemma relying heavily on a version of the law of non-contradiction (LNC), according to which contradictory properties cannot be predicated of the same subject (*viruddhadharmādhyāsa*). Using these two dilemmas, the logical space of views on the mind and mental content is as follows:

i.   Mental content is real and non-distinct from the mind.

ii. Mental content is real and distinct from the mind.

iii. Mental content is unreal and non-distinct from the mind.

iv. Mental content is unreal and distinct from the mind.

Śrīgupta rejects each option in turn, arguing that there is no coherent account of how the mind could exist as a true unity. To summarize:[33] (i) Śrīgupta's argument against the first

---

[33] This sequence of subarguments is found at TAV (PD 3121, 102–4).

view runs the One-or-many Dilemma on mental content. He reasons that *mental content that is real and non-distinct from the mind* is either unitary or non-unitary. This argument turns on the datum that mental content, like the perception we might have of this page, is phenomenally variegated (Skt. *citra*, Tib. *sna tshogs*). He takes it that whatever is phenomenally variegated, being always conceptually divisible into proper parts (like the perception of this word and the perception of that word), is not truly unitary. But if non-unitary mental content is non-distinct from the mind, then given the LNC, the mind too is non-unitary. (ii) But if *mental content is real and distinct from the mind*, a different problem crops up: if, as Śrīgupta's idealist foundationalist interlocutors would have it, mental content is metaphysically founded in the mind, how could it also be metaphysically distinct from (and thus metaphysically independent from) the mind? These idealist foundationalists are thus in agreement with Śrīgupta in discounting this second possible view.

(iii) Moving to the other horn, or the Real/unreal Dilemma, Śrīgupta next argues that if *mental content is unreal and non-distinct from the mind*, then given the LNC, the mind too would be unreal, which is obviously an unacceptable consequence for his idealist foundationalist opponents. (iv) Finally, in tackling the view that *mental content is unreal and distinct from the mind*, Śrīgupta first points out that if mental content does not stand in an identity relation with the mind, it must nonetheless stand in *some* kind of relation with it— perhaps a causal one—in order to account for our experience of it. But only real things can stand in relations with other real things. A dragon can't cause a real forest fire. Likewise, an unreal percept could not cause a real perception of it. Moreover, Śrīgupta adds, being a mere figment, unreal mental content could not account for the phenomenal determinacy

35

and consistency that is the default of our ordinary experience. With this, Śrīgupta dismisses the fourth and final possible way in which a mind or mental state might exist as a true unity.

Finally, in a series of additional subarguments targeting neither-material-nor-mental simples, Śrīgupta also rejects the simplicity of purportedly all-pervasive entities (*vyāpin*), like ether (*ākāśa*), space (*diś*), time (*kāla*), and the self/soul (*ātman*), all of which, he argues, are—in some manner—conceptually divisible into proper parts due to being connected with spatially or temporally discrete loci.

With the conclusion of this series of subarguments against material, mental, and neither-material-nor-mental simples, Śrīgupta takes himself to have exhausted the possibilities for how something could exist as a true unity. Since the existence of a true multitude presupposes the existence of true unities, he thereby establishes the minor premise that all things are neither one nor many. And given the major premise—that whatever is neither one nor many does not have independent being—Śrīgupta concludes that all things are devoid of independent being; there are no fundamentalia of any kind.

2. The Positive Phase: Conventional Reality and Metaphysical Indefinitism

If nothing that exists is ontologically independent, then whatever exists is ontologically *dependent*. But precisely what kind of metaphysical dependence structure do Mādhyamikas endorse? This question has received surprisingly little serious attention in the secondary literature beyond competing suggestions that Madhyamaka is a form of

coherentism or else a kind of infinitism.[34] Recently, Ricki Bliss and Graham Priest have put forward the most technical account of the Madhyamaka dependence structure to date, claiming that it is a form of metaphysical infinitism, characterized by extendability together with the structural properties demanded by the Hierarchy Thesis, viz. irreflexivity, asymmetry, and transitivity.[35] Though a helpful starting place, their characterization does not get at the heart of what is most radical about the Madhyamaka picture. In his neither-one-nor-many argument, Śrīgupta runs a reductio on his foundationalist opponents, supposing that a strict priority relation of the kind demanded by the Hierarchy Thesis obtains between true unities and the multitude that they constitute. But according to Śrīgupta, since there are no true unities to ground a strict, asymmetrical priority relation of this kind, neither can there exist such a relation. But if Śrīgupta's Madhyamaka is not a straightforward infinitism, what is it? The resources for beginning to flesh out this picture can be found in Śrīgupta's account of conventional reality.

---

[34] Some have claimed that Madhyamaka endorses a kind of metaphysical coherentism (e.g., Walser 2005, 243–44; Goodman 2016, 143), others have claimed that it endorses appearances all the way down (e.g., Sprung 1977, 264; Huntington 1983, 326; Cabezón 1994, 163; Arnold 2010, 375), and still others have suggested that both coherentism and infinitism are defensible accounts of Madhyamaka (Westerhoff 2016, 356). Claiming that Madhyamaka endorses a metaphysical dependence structure at all—or *any* metaphysical claims for that matter—is not uncontroversial, given that many interpreters take Nāgārjuna to be a skeptic, a mystic, or an anti-metaphysicalist, based in part on his famed and interpretively vexed statement in VV k. 29 that he has no thesis; see also YŚ k. 50.

[35] See Bliss and Priest (2018b, 70–71), where they claim that Nāgārjuna's Madhyamaka conforms to this infinitist dependence structure; a similar paper with this same claim appears as Priest (2018). In the contemporary space, the metaphysical possibility of metaphysical infinitism has been defended by Schaffer (2003), Bohn (2009, 2018), Bliss (2013), Tahko (2014), and Morganti (2014, 2015). On the logical consistency of infinitism and non-well-founded set theory, see Aczel (1998); on the application of non-well-founded set theory to Madhyamaka, see Priest (2009, 2014).

Upon concluding his neither-one-nor-many argument, Śrīgupta is quick to affirm the conventional reality of all things, the second of the Madhyamaka "two truths/realities" (*satyadvaya*),[36] in order to clarify that the rejection of independent being is not tantamount to nihilism. Instead, his view is intended to navigate a middle way between the realism of foundationalism and thoroughgoing nihilism:[37] the kind of being and unity that actually exist are merely conventional. But what precisely does it mean to characterize the being and unity of everything from an aardvark to an atom as "conventional"?

Śrīgupta sets out an influential threefold criterion for conventional reality,[38] according to which whatever exists conventionally (i) is satisfactory only when not analyzed

---

[36] The semantic range of the term *satya* includes "truth," "reality," "existence," and "being," and thus *satyadvaya* is commonly translated as "two truths" as well as "two realities." Perhaps *satya* is best understood as "reality" insofar as *satya* generally refers to that which is non-deceptive, trustworthy, or which exists and appears in the same way. "Truth," then, as a property of a proposition or a sentence, may simply be a special case of this non-deceptiveness. Indeed, *saṃvṛti*, which I translate here as "conventional" in accord with popular practice (though *vyavahāra* is more properly "conventional"), means "concealing," or "obscuring," and might more literally be understood as the provisional truth that obscures the actual truth, or the appearances that obscure reality. For the present purposes, I translate *saṃvṛtisatya* as "conventional existence/being/reality" since Śrīgupta's definition of the conventional describes the manner of existence of appearances. However, below I will also discuss the implications of Śrīgupta's account of conventional reality on conventional truth, understood as the truth-tracking claims we make about conventionally real things. I translate *paramārthasatya* throughout as "ultimate truth" since in the present discussion it refers to the claim, or thesis, of Śrīgupta's neither-one-nor-many argument, viz. the universal absence of ontologically independent being.

[37] For Nāgārjuna on Madhyamaka as the middle way, see for instance, MMK 15.2. Śrīgupta echoes this Madhyamaka refrain in TA 23: *sgro 'dogs pa dang skur pa'am // mtha' gnyis kyi ni rnam spangs* [D: *sngags*] *pa // 'di ni dbu ma'i lam yin par // seng ge'i nga ros bstan pa yin //* (PD 3121, 110); "That the rejection of the two extremes of reification and nihilism is the path of the middle way is what was taught by the lion's roar."

[38] Subsequent endorsements of this threefold criterion include, for instance, Jñānagarbha's SDV 8, 12, and 21; Śāntarakṣita's MA 64; Kamalaśīla's MAP *ad* 64; Haribhadra's AAA (Wogihara 1932–1935, 594.18–25); the ca. eighth century Bhāviveka's MAS 9–11 and MRP I.4; Atiśa's SDA 3. For instance, Śāntarakṣita MA 64 reads: *ma brtags gcig pu nyams dga' zhing // skye dang 'jig pa'i chos can pa // don byed pa dag nus rnams kyi // rang bzhin kun rdzob pa yin rtogs //* (Ichigō 1989, 202); "The conventional should be understood as something whose nature: [1] satisfies only when not analyzed, [2] has the properties of

(*avicāraramaṇīya* or *avicāramanohara*), (ii) is interdependently originated (*pratītyasamutpanna*), and (iii) has the capacity for causal, or pragmatic, efficacy (*arthakriyāśakti* or *arthakriyāsāmarthya*).[39] In other words, whatever is conventionally real (i) does not withstand the kind of analysis into its ultimate nature that seeks to uncover some ontologically independent entity, and yet it (ii) comes into being in dependence on other conventionally real things and (iii) fulfills our pragmatic expectations in accordance with how it appears.

I will argue that these three criteria jointly yield a structure I call "metaphysical indefinitism," which involves dependence relations that are irreflexive and extendable, but not strictly asymmetric or transitive, admits of indefinite—but not straightforwardly infinite—dependence chains as well as dependence loops of non-zero length, and yet allows for a contextualist form of the Hierarchy Thesis that will bring the Madhyamaka ontological dependence relation closer in line with standard accounts of the metaphysical grounding and ontological dependence relations. The resulting picture supports a flexible

---

arising and disintegrating, and [3] has the capacity for causal efficacy." For alternate translations, see Ichigō (1989, 213) and Blumenthal (2004, 242).

[39] TA 11: *ma brtags gcig pu nyams dga' ste // de 'dra las byung de bzhin no // dngos po de dag de lta bu'i // don bya de dang de byed do //* (PD 3121, 105); "[1] Satisfactory only when not analyzed, [2] [things] arise from [causes] similar to themselves. [3] Those things enact their respective forms of causal efficacy." The TAV continues: *de lta bas na phyi rol dang nang na snang ba'i dngos po brtag pa'i spungs mi bzod pa rang dang mthun pa'i rgyus bskyed pa 'di dag ni gang las tha snyad 'dir 'gyur ba don bya ba ma brtags na nyams dga' ba nyid de dang der nye bar byed do //* (PD 3121, 105–6); "Thus, regarding these things that appear both externally and internally, which cannot withstand the pressure of analysis and which are produced from causes similar to themselves, based on which conventions (**vyavahāra*) then come into being—if one has not examined their causal efficacy, one will approach satisfaction here and there." As Eckel (2008, 25) points out, Śrīgupta's TAV appears to be the earliest extant text in which we find this threefold characterization of conventional reality, with the first criterion as listed above possibly adapted from Candrakīrti (e.g., MAv 6.35), the second inherited from Nāgārjuna, and the third a repurposing of Dharmakīrti's criterion for ultimately real particulars (*svalakṣaṇa*) in PV 3.3. On these three criteria, see also Eckel (1987, 137–38 n. 104).

ontology and revisable theory of conventional truth that will deliver important payoffs for the view, including its capacity to keep in step with the latest developments in scientific explanation. How, then, do Śrīgupta's three criteria for conventional reality deliver this picture?

2.1 The "Satisfies When Not Analyzed" Criterion and Infinitism vs. Indefinitism

On Śrīgupta's first criterion for conventional reality, whatever is conventionally real can satisfy our ordinary notions of existence and unity so long as it is not subjected to metaphysical analysis like that involved in the neither-one-nor-many argument. [40] Conventionally real things, Śrīgupta claims, dissolve under analysis of their ultimate nature: every object taken up for analysis is physically and/or conceptually divisible; nothing turns out to be a true unity, ontologically independent, or fundamental. Instead, <unity> and <being> are designated in dependence (*upādāya prajñapti*) upon some proper parts. [41] A gaggle is designated, or conceived in dependence on some geese, whereby it derives accidental being and accidental unity. Yet the being and unity of an individual goose is also

---

[40] One may be reminded here of Hume's claim that the distinct and continued existence that we attribute to material objects is the work of the imagination, and although this operation of the imagination is epistemologically unfounded (Śrīgupta's first criterion for conventional existence), it is nevertheless pragmatically efficacious (Śrīgupta's third criterion for conventional existence).

[41] See MMK 24.18–19. Here, Nāgārjuna identifies dependent origination with emptiness, which he in turn identifies as dependent designation. See Salvini (2011) for an argument based on grammatical analysis in support of reading Nāgārjuna as equating *upādāyaprajñapti* with *pratītyasamutpāda*, as Candrakīrti does in his PP *ad* MMK 24.18.

designated in dependence on its own proper parts.[42] And aardvarks and atoms are no different. And with no simples to be found upon analysis, it's aggregates—i.e., conventional unities and beings—all the way down.

Since conventional unity and conventional being are necessarily designated or conceived by some cognitive agent in dependence upon some parts, just like the unity and being of an army are designated in dependence upon some soldiers, whatever exists conventionally is in some sense mind-dependent.

In fact, in Śrīgupta's presentation of the conventional reality of all things, he glosses "things" as "things that *appear* externally and internally."[43] Whatever exists conventionally, then, is an appearance, which, by definition, exists in relation to some cognitive agent insofar as it necessarily appears *to* someone.[44] Importantly, the ontological status of dependence relations and dependence structures can be no different from that of the relata that populate the structure.[45] The unity and being of the structure too dissolve under

---

[42] This, of course, means that there is no deep, metaphysical difference between the unity of an aggregate like an army and the unity of something like a human organism, which also turns out to be an aggregate (of aggregates).

[43] See TAV *ad* 11, where Śrīgupta unpacks *dngos po* from TA 11 as *phyi rol dang nang na snang ba'i dngos po* (PD 3121, 105).

[44] Here, one might draw a parallel with Kantian appearance (*Erscheinung*) which is empirically real, though transcendentally ideal. However, Śrīgupta should not be read as suggesting that there may be some non-appearing thing akin to a Kantian thing in itself (*Ding an sich*), which might claim ontological independence, since Śrīgupta argues that ontological independence is metaphysically impossible.

[45] As Westerhoff (2017a, 288) points out, an ontological structural realism, such as that developed by Ladyman and Ross (2007) and French (2014), which "privileges structures over the individuals individuating the structures, and attempts to dispense completely with the notion of a fundamental level" is incompatible with the Madhyamaka denial of "ultimately real entities." For the Mādhyamika, relations are no more fundamental or ontologically independent than the relata that they structure. Westerhoff suggests, however, that the combination of such a position with another account, such as the rejection of absolutely general quantification or semantic contextualism, might be a promising strategy for characterizing the Madhyamaka view in contemporary terms. See Westerhoff's (2017a, 292–94)

analysis, are designated in dependence on some parts, and are mind-dependent. Thus, although dependence chains are endless insofar as they do not terminate in any ungrounded or self-grounding entity, they cannot be mind-independently infinite. They are, instead, only indefinite.

By "indefinite," I intend a kind of potential, mind-dependent infinite, as opposed to an actual, mind-independent infinite. Indefinite characterizes the relation among members in a series such that for any given member, there will always be a subsequent member; there will always be more than one may specify. In the case of an actual infinite, on the other hand, the quantity in question is put in one-to-one correspondence with the natural numbers.[46] For the Mādhyamika, then, although there is no mind-independent, actually infinite dependence chain, were one to analyze any given chain, one would never arrive at a limit. Accordingly, when it comes to mereological dependence as a subspecies of ontological dependence, any given hunk of matter—whether a canyon or a quark—is potentially indefinitely divisible, though not actually infinitely divided. In this way, Śrīgupta's metaphysical indefinitism is subjectivist and anti-realist,[47] differentiating it from

___

application of radical contextualist semantics to resolve two apparent problems for Mādhyamikas: (i) the apparent paradox of the Madhyamaka commitment to the claim that all conceptual hypostatization (*prapañca*) is false, together with the fact that this very claim involves conceptualization (and thus, much like the Liar Paradox, this claim is false if it is true); and (ii) recourse to intrinsic natures (*svabhāva*) is used in ordinary discourse to stop justification regresses, but the Mādhyamika rejects the existence of intrinsic natures. Below, I explore a related contextualist strategy.

[46] It is unclear whether or not Indian philosophers like Śrīgupta were ever actually working with the concept of a quantitative, or mathematical, infinite. On the other hand, concepts like limitless (*anantaka*) and immeasurable (*aparimāṇa*) were commonplace and, I suggest, conform to the structural notion of an indefinite infinite. Similarly, the Sanskrit term for an infinite regress, an endless series, *anavasthā*, is suggestive in its etymology of unfoundedness, ungroundedness, or of falling without stopping.

[47] For an account of Madhyamaka as a form of global anti-realism, see, for example, Siderits (1988, 1989) and Westerhoff (2011). Madhyamaka has been variously categorized as a kind of "nihilism,

realist forms of mereological infinitism that see the world constituted by mind-independent gunk.

The indefinitism that characterizes Madhyamaka dependence chains is a third alternative to the infinitism-finitism dichotomy, where finitism picks out a structure—whether foundationalist or coherentist—wherein a finite quantity of relata stand in a finite number of dependence relations. Since indefinitism follows from the mind-dependence of the structure and its members,[48] it is not simply a claim of epistemic or semantic indeterminacy. For instance, the mereological structure of a quark is not indefinite simply because it is impossible to know the dividedness of the quark in its entirety. Nor is it indefinite merely due to the limitations of our semantic or representational resources. Rather, the indefinitism of Madhyamaka dependence chains represents a kind of *metaphysical* indeterminacy; the reality of the dividedness of a given object is settled only insofar as we have (mentally or physically) carried out the division.[49] Epistemic and

---

monism, irrationalism, misology, agnosticism, skepticism, criticism, dialectic, mysticism, acosmism, absolutism, relativism, nominalism, and linguistic analysis with therapeutic value" (Seyfort Ruegg 1981, 2). To this, we may add panfictionalism (Matilal 1970), ontological deflationism (MacKenzie 2008), conceptualism (Spackman 2014), quietism (Tillemans 2016), and realist-antimetaphysicalism (Ferraro 2017). Similarly, Garfield and Samten (2006, xx) sum up the variety of characterizations of Nāgārjuna as follows: "an idealist (Murti 1960), a nihilist (Wood 1994), a skeptic (Garfield 1995), a pragmatist (Kalupahana 1986), and . . . a mystic (Streng 1967). He has been regarded as a critic of logic (Inada 1970), as a defender of classical logic (Hayes 1994), and as a pioneer of paraconsistent logic (Garfield and Priest 2003)."

[48] This only follows, of course, so long as the mind on which the structure depends is not itself actually infinite. My thanks to Gideon Rosen for this clarification.

[49] This account anticipates certain elements of the resolution of Kant's second antinomy, according to which composite substances are neither composed of simples nor are they actually infinitely divided. Instead, on Kant's transcendental idealism, since the world as a totality is not given in appearance, matter is indefinitely divisible without consisting of infinitely many parts, i.e., matter is only divided insofar as we have carried out that division. As he concludes in his *Metaphysical Foundations of Natural Science*, "Therefore, one can only say of appearances, whose division proceeds to infinity, that there are just so many parts in the appearance as we may provide, that is, so far as we may divide. For the parts,

semantic indeterminacy may follow from this, however indefinitism is foremost a metaphysical claim.

Now, one might think that this "indefinitism" is actually a veiled form of finitism along the following lines: If the dividedness of an object is mind-dependent, and if only a finite number of divisions have been made at t_1, then the structure is finite at t_1. And if at t_2 further divisions are made, then there will be *more* parts than at t_1, but the structure remains finite and perfectly definite nonetheless. Alternatively, one might take "indefinitism" to mean that, as things stand, prior to analysis, it is unsettled as to whether or not a given object has parts; that is, when uninspected, an object is neither simple nor complex.[50] Yet neither veiled finitism nor indefinitism about simplicity vs. complexity is what is intended by the indefinitism under discussion here. Instead, since everything is necessarily indefinitely divisible, and since divisibility is a sufficient criterion for being partite, the fact that any given object *x* has parts is settled *a priori*. Furthermore, the fact that *x* will, upon analysis, turn out to have more parts than one may specify, viz. indefinitely many parts, is settled *a priori*. What *is* unsettled prior to analysis is *which* parts get carved out and identified as the basis of imputation for *x*'s conventional unity and being.

---

as belonging to the existence of an appearance, exist only in thought, namely, in the division itself" (4:506–7; 2002, 218).

[50] My thanks to Gideon Rosen for raising these two points.

2.2 The Interdependent Origination Criterion and the Structural Properties of Madhyamaka Dependence Relations

So far, I've argued that from Śrīgupta's first criterion for conventional reality—that the being and unity of a conventionally real thing satisfies only when not analyzed and is designated in dependence on its proper parts—it follows that Madhyamaka dependence chains are indefinite rather than infinite. To further flesh out the properties of the Madhyamaka dependence structure, let's look to Śrīgupta's second criterion for conventional reality, which says that whatever exists conventionally comes into being in dependence on something else. This is most literally a claim about causal dependence, implying the denial of self-causation as well as a first cause, and entailing causal dependence *ad indefinitum*. [51] Yet, this claim of universal dependent origination also applies to mereological dependence. After all, just as a sprout does not originate in the absence of a seed, neither does a gaggle originate in the absence of geese. On this criterion, however, neither does a seed originate in the absence of its own causes and conditions, nor a goose in the absence of its own proper parts—wings, beak, webbed feet, etc. To begin to pin down the structural properties of the Madhyamaka dependence relation that obtain in both these kinds of cases, it may be helpful to contrast it with the metaphysical grounding relation.

There are three commonly accepted features of metaphysical grounding that are *incompatible* with Madhyamaka metaphysical dependence: metaphysical grounding is standardly (i) a non-causal relation of metaphysical explanation, (ii) a relation that obtains

---

[51] This claim is not so strange given a conceptual context wherein time too has no beginning. Indeed, those upholding a beginning of time arguably take on a greater explanatory burden.

between facts and not between things, and (iii) a priority relation.[52] That is, if *x* is grounded in *y*, then *y* is prior to and more fundamental than *x*, and *y* metaphysically explains *x* (where *x* and *y* are facts). The dependence structure that Mādhyamikas like Śrīgupta affirm is broad ranging, making use of mereological, causal, and conceptual dependence relations, and it is not at all clear that any of them qualify as a (standardly characterized) metaphysical grounding relation. Madhyamaka dependence relations (i) are sometimes but not always causal and sometimes but not always involve metaphysical explanation, (ii) can obtain not only between facts but also (and perhaps more commonly) between things, and (iii) are not strict priority relations.

Let's take a closer look at the structural properties of the grounding relation that jointly enforce priority, which is the same set of three properties demanded by the Hierarchy Thesis and which yields a strict partial order: irreflexivity, asymmetry, and transitivity. To this, we can add extendability, as the infinitist would.[53] Mādhyamikas

---

[52] The features of metaphysical grounding are, of course, hotly debated, but I engage what I call a "standard account" characterized by these three features together with the three structural properties of irreflexivity, asymmetry, and transitivity, because this provides a clear picture against which to clarify Madhyamaka dependence relations. For arguments that the grounding relation is not necessarily irreflexive, see Fine (2010), Jenkins (2011), and S. Krämer (2013); for a challenge to its asymmetry, see Barnes (2018) and Thompson (2018); for a challenge to its transitivity, see Schaffer (2012), where he argues that transitivity can be restored by a contrastive account of grounding. It is a contested question whether grounding even picks out fundamentally one relation (Audi 2012, Rosen 2010, Schaffer 2009, Berker 2018) or whether it denominates a plurality of relations that include, for instance, metaphysical grounding, natural grounding, and normative grounding (Fine 2012, Wilson 2014). For surveys of disputed issues related to grounding see Correia and Schnieder (2012), Trogdon (2013), Raven (2015), and Bliss and Trogdon (2016).

[53] Bliss and Priest (2018b, 7) formalize these four structural properties as follows, where $x{\rightarrow}y$ represents "*x* depends on *y*": (1) anti-reflexivity: $\forall x \neg(x{\rightarrow}x)$; (2) anti-symmetry: $\forall x \forall y(x{\rightarrow}y \supset \neg y{\rightarrow}x)$; (3) transitivity: $\forall x \forall y \forall z((x{\rightarrow}y \wedge y{\rightarrow}z) \supset x{\rightarrow}z)$; (4) extendability: $\forall x \exists y(y{\neq}x \wedge x{\rightarrow}y)$.

unequivocally affirm irreflexivity.[54] That's because something's being self-grounding (or standing in a reflexive ontological dependence relation) is tantamount to having independent being. And extendability is demanded by Śrīgupta's second criterion for conventional reality, that everything originates in dependence on something else.[55]

But it is clear that Mādhyamikas *do* admit of instances of symmetrical dependence.[56] Indeed, the term for dependent origination, *pratītyasamutpāda*, with its *sam-* prefix, literally means dependent *co*-origination, implicating some kind of interdependence, or mutual dependence. Two things that are conceptually mutually dependent, like right and left, do not have an obvious priority relation. It's difficult to conceive of an ontology where right is more fundamental than left. At first blush, the flat, mutual dependence relation between right and left looks nothing like the hierarchical ontological priority relation that obtains between a molecule and its atoms. Yet mutual ontological dependence should not seem so

---

[54] For examples of Nāgārjuna's rejection of reflexivity, see MMK 3.4, 7.1, 7.8, and 7.28. These arguments can also be read as cases against the metaphysical possibility of fundamental entities that are self-grounding.

[55] Extendability *qua* the dependent origination of all things is the central claim of Nāgārjuna's MMK, as made explicit in the dedicatory stanza. For instance, extendability is endorsed in terms of causal dependence at MMK 4.2c2d: *na cāsty arthaḥ kaścid āhetukaḥ kvacit* (Ye 2011, 68); "There is nothing whatsoever that exists without a cause." On this stanza, see also Garfield (1995, 144) and Siderits and Katsura (2013, 53).

[56] Nāgārjuna affirms symmetrical dependence for conventionally existent things. See, for instance, MMK 8.12: *pratītya kārakaḥ karma taṃ pratītya ca kārakam / karma pravartate nânyat paśyāmaḥ siddhikāraṇam //* (Ye 2011, 142); "The agent exists in dependence on action, and action exists in dependence on that agent. We see no other means for establishing [them]." On this stanza, see also Garfield (1995, 181) and Siderits and Katsura (2013, 96–97). For a related point, see Nāgārjuna's *Śūnyatāsaptati* 13. At times, Nāgārjuna appears to reject symmetrical dependence, e.g., MMK 6.6, 7.6, 10.8–10, 11.5, 20.7. However, these arguments target symmetrical dependence as advanced by his realist opponents, who maintain that the relata that purportedly stand in symmetrical dependence relations have thick being (viz. ontologically independent being).

strange; consider, for instance, the mutual dependence of the north and south poles of a magnet.[57]

According to the Mādhyamika, the dependence between a part and a composite is more similar to that between the north and south poles than may be initially supposed. A part might just as easily be said to depend on the composite as the composite on the part.[58] For instance, a human organism depends on a heart, but the heart also depends on the human organism. Two conventionally real things might thus stand in a mutual dependence relation. This admission of symmetrical dependence taken together with the commitment to irreflexivity prevents Śrīgupta from honoring transitivity.[59] Thus, unlike standard accounts of the metaphysical grounding relation, the ontological dependence relation, and the proper parthood relation—all of which conform to the Hierarchy Thesis— Madhyamaka dependence relations are neither strictly asymmetrical nor transitive.

Bliss and Priest present a taxonomy of sixteen structures of reality derived from the range of combinations of the four structural properties of irreflexivity, asymmetry,

---

[57] I borrow this example from Bliss and Priest (2018a, 14).

[58] See, for instance, Candrakīrti's MAv 6.161ab, where he argues that, just as a whole cannot exist without parts, neither can parts exist without the whole, suggesting their mutual ontological dependence: *sattvaṃ rathasyāsti na cet tadānīṃ / vināṅgināṅgāny api santi nāsya //* (Li 2015, 24).

[59] To the best of my knowledge, transitivity was not a concept that Śrīgupta and his fellow Mādhyamikas were explicitly concerned with, so the claim that they do not strictly honor transitivity is not an independent condition, but derived from the fact that they are committed to irreflexivity but not to asymmetry. They could, however, endorse a limited transitivity, such that: $\forall x \forall y \forall z ((x{\rightarrow}y \wedge y{\rightarrow}z) \wedge (x{\neq}y \wedge y{\neq}z \wedge x{\neq}z)) \supset (x{\rightarrow}z)$; my thanks to Ginger Schultheis and Jan Westerhoff for raising this point.

transitivity, and extendability.[60] Although they assign Nāgārjuna's Madhyamaka to a kind of infinitism that accommodates all four of these properties (structure 1 in their list, an infinite partial ordering), the Madhyamaka of Nāgārjuna and Śrīgupta instead conforms most closely to structure 7 in their list, a kind of infinitism that honors irreflexivity and extendability but not asymmetry or transitivity. The rejection of asymmetry and transitivity undermines the priority ordering, and thus admits of dependence loops, or circles, but—given irreflexivity—only loops of greater than zero length. And given extendability, there are no fundamental elements. Bliss and Priest argue that structure 7 (together with the other non-standard permutations of these properties) is not only logically possible but also metaphysically possible.[61] Structure 7 meets Bliss and Priest's definition of infinitism on which "every element is ultimately ungrounded" ($\forall x \mathrm{UG} x$), where an element x is ultimately ungrounded (UG) if one never meets with a foundational element (FE) in x's dependence chain, $\forall y(x \rightarrow y \supset \neg \mathrm{FE} y)$, and extendability is entailed (2018b, 67). It does not meet their (strong) definition of coherentism, on which everything is dependent on everything else ($\forall x \forall y\ x \rightarrow y$).[62] But given that this structure permits symmetry and thus dependence loops, it may well satisfy other definitions of coherentism.

---

[60] As Bliss and Priest clarify, only ten of the sixteen combinations are logically possible (2018a, 7; 2018b, 65).

[61] See Bliss and Priest (2018a, 10ff). In the same volume, Barnes (2018) argues that ontological dependence is symmetrical, and Thompson (2018) argues that grounding is non-symmetric rather than asymmetric. See also Morganti (2018) for a recent case for metaphysical coherentism on which ontological dependence is symmetrical.

[62] Thus, for Bliss and Priest, coherentism obtains only in the case of a preorder, which honors reflexivity, symmetry, and transitivity, and may or may not be extendable.

To sum up the properties of the metaphysical dependence structure that follow from these first two criteria for conventional reality: (i) everything depends on something else (extendability), (ii) nothing depends on itself (irreflexivity), (iii) some elements may be (but need not be) symmetrically dependent, admitting dependence loops of >0 length, and (iv) dependence chains are indefinite, though not actually infinite, in length. As it turns out, then, the proponents of coherentist and infinitist interpretations of Madhyamaka each have it partly right; in a way, it's both.

This has been a sketch of the first level of analysis of the Madhyamaka metaphysical dependence structure, which highlights the way in which it falls outside the standard accounts of the three categories of foundationalism, infinitism, and coherentism, and also how it treads an alternative path to the infinitism-finitism dichotomy. On this picture, we might say that an organism like a human body ontologically depends on its heart, which depends on some cells, and so on, *ad indefinitum*; and yet the heart also ontologically depends on the organism.

Nevertheless, some Madhyamaka dependence relations might still be helpfully discussed in terms of metaphysical grounding. I will focus on the mereological dependence relation since that looks like the most promising candidate.[63] Suppose that the Mādhyamika agrees that mereological dependence is a kind of existential dependence that can be cashed out in terms of a relation of metaphysical explanation that obtains between facts. Still,

---

[63] See Cameron (2014) for an argument that the part-whole relation is an instance of the metaphysical grounding relation. See Wilson (2014) for an argument that the part-whole relation is an instance of a "small-g" grounding relation, which, together with a variety of other dependence relations, ought to be differentiated from the metaphysical explanation relation signified by the "big-g" Grounding relation. See Berker (2018) for a response to Wilson.

priority looks like a sticking point. I will next show how Śrīgupta's third criterion for conventional reality facilitates a second level of analysis that shows how this picture can support a qualified form of the Hierarchy Thesis and a certain degree of structural flexibility.

2.3 The Causal Efficacy Criterion and a Revisable Theory of Conventional Truth

Śrīgupta's third criterion for conventional reality says that whatever exists conventionally has the capacity for causal or pragmatic efficacy, which simply means that it can fulfill our pragmatic purposes in accordance with our expectations. If something fails to have the capacity for causal efficacy, like the apparent water of a mirage that fails to quench my thirst, then it is *not* conventionally real. Conventionally real things work, yet if we investigate their operations, we will find that they do not ultimately rest on foundations; their dependence relations do not terminate in ontologically independent beings.

Contemporary objections to Madhyamaka include the charge that its flat ontology is unable to accommodate developments in scientific explanation without resorting to a "dismal slough"[64] of "anything goes" relativism, and that Abhidharma Buddhism, for instance, read as a kind of reductionism, is more promising in this respect.[65] To the contrary,

---

[64] Tillemans (2011, 152) uses this expression to describe a relativist reading of Candrakīrtian Madhyamaka.

[65] See, for instance, Siderits (2011), whose objection—and others like it—is based on a Candrakīrtian reading of Madhyamaka. Indeed, most serious contemporary efforts to make philosophical sense of the Madhyamaka theory of conventional reality/truth have engaged it through a Candrakīrtian lens (e.g., Cowherds 2011). Regardless of whether or not Candrakīrti could field this kind of objection, Śrīgupta's Madhyamaka is more obviously equipped to respond, in large part because of Śrīgupta's repurposing of Dharmakīrti's causal/pragmatic efficacy criterion for real particulars as one of his three criteria for real

I argue that an important *virtue* of Śrīgupta's Madhyamaka is its capacity to honor scientific hierarchies as well as developments in scientific explanation without endorsing an unmitigated relativism. That's because Śrīgupta's pragmatic efficacy criterion for conventional reality supports an ontology that is not flat, but flexible.

To be sure, Mādhyamikas are—in a certain sense—ontological egalitarians, insofar as they are committed to the metaphysical impossibility of ontological independence. Strictly speaking, nothing can be said to have more ontological independence than anything else, any more than one person can be said to own more jackalopes than anyone else. So one might think that it's incoherent for a Mādhyamika—or for any anti-foundationalist for that matter—to speak of one thing's being prior to, or more fundamental than, anything else, for the rejection of foundationalism may seem to entail a flat ontology.

One, therefore, might think that it's incoherent to both reject the Fundamentality Thesis *and* uphold the Hierarchy Thesis. But the Hierarchy Thesis does not *presuppose* the Fundamentality Thesis. A hierarchical chain of metaphysical priority does not in principle require the existence of something most (or least) fundamental.[66] An indefinite (or infinite) hierarchical chain is not obviously incoherent. The fact that there is nothing absolutely fundamental no more precludes one thing's being *more* fundamental than another than the absence of a perfect Euclidean triangle in the world precludes one thing's being more

conventionals. In this, Śrīgupta is followed by Śāntarakṣita, Kamalaśīla, Haribhadra, the eighth century Bhāviveka, Atiśa, and others.

[66] Just as some kind of axiom of regularity, or axiom of foundation, is required in set theory to demand first elements in a set and rule out non-well-founded sets, similarly some constraint must be added to the metaphysical grounding relation to demand some final, ungrounded ground.

triangular than another. A metaphysical structure might thus be egalitarian in one sense, viz. insofar as everything is the same in lacking ontological independence, and nonetheless have depth, or verticality, in another sense, viz. insofar as it might instantiate asymmetry in certain of its dependence chains.

Since Mādhyamikas admit of *mutual* dependence, they can accept that a composite is dependent on its parts—in some sense—and yet the parts are also—in another sense—dependent on the composite, and thus neither is strictly prior to the other. But the admission of certain kinds of mutual dependence neither rules out the possibility of asymmetrical dependence chains nor takes it for granted. Likewise, hierarchical structures are neither ruled out *a priori* on this picture nor are they necessary. By providing contextualist qualifications to asymmetry, the Mādhyamika could admit hierarchical structures for which the direction of priority is extrinsically determined by, for instance, a given dialogical, analytic, or scientific context.[67]

Madhyamaka anti-foundationalism is thus not a picture on which there is *no* structure, but one on which the structures are richer than might be presumed. Yet this

---

[67] Other examples of contexts that might extrinsically determine priority include the analytical context of a solitary epistemic agent, and the dialogical context of more than one epistemic agent is involved in an exchange. In the latter context, the Mādhyamika may adopt the presupposed hierarchy of the interlocutor: when the opponent is an atomist, the micro will be supposed to be more fundamental, and when the opponent is a monist, the macro will be supposed to be more fundamental. This is, of course, not to say that the Mādhyamika provisionally accepts foundationalism. Rather, they can accept the direction of contextual priority in a given dependence structure in order to demonstrate that it has no final ground. Westerhoff (2016, 372) similarly argues that a kind of "opponent-relativist feature" characterizes Madhyamaka, both in the structure of its arguments and in the theory as a whole. Siderits (2003, 2011) similarly argues that Madhyamaka ought to endorse a kind of epistemological contextualism, according to which some "procedure counts as an epistemic instrument only relative to a context of inquiry, where contexts of inquiry are determined by factors such as aims of the inquirer and the methods of inquiry available to the inquirer" (2011, 178).

quasi-maximalism about structure is constrained by a pragmatic understanding of *which* structures are salient. It can accommodate hierarchical scientific structures, but at the same time, it leaves science open to pursue other kinds of non-hierarchical, non-reductionist models.

A strict hierarchic and reductionist ontology may encourage the thought that there is a single privileged way of carving up and ordering the world. But the Madhyamaka *flexible* ontology can recognize scientific insights, while also granting legitimacy to other ways of thinking about the world, e.g., to recognizing a certain *kind* of priority not just to quarks and leptons, or strings in ten-dimensional space, but also to plants, animals, people, and even—if it is useful—countries and corporations.

But conventional *truths*—as the truth-tracking claims we make about conventionally real things and structures—are not simply claims that are commonly accepted within a given society (*lokaprasiddha*); nor is causal/pragmatic efficacy underwritten by popular opinion. Śrīgupta's successor, Kamalaśīla, who endorses a version of Śrīgupta's threefold criterion of conventional reality, provides a word of caution in response to a rival Mādhyamika position that endorses a form of relativism, which sanctions common consensus as the guide to what is conventionally real.[68] Kamalaśīla observes that the

---

[68] SN (PD 3118, 1479–80); see also MAP (Ichigō 1985, 203) and MĀ (PD 3116, 1133). In his MAP (Ichigō 1985, 203), Kamalaśīla resists the definition "conventional truth" according to which it signifies commonly accepted linguistic-conceptual practices. For a translation and discussion of the relevant passage from the SN, see Tillemans (2011, 153–54), where Kamalaśīla rebuts an unnamed opponent reminiscent of Candrakīrti. For a discussion of Kamalaśīla's account of a discerning person (*prekṣāvat*), see McClintock (2010, 58–62; 2013) and Tillemans (2016, 143–44); on this term, see also Eltschinger (2007, 137–50; 2014, 195 n. 17, 219–34). This points to the difference between Mādhyamikas like Candrakīrti and those in the tradition of Śrīgupta, Śāntarakṣita, and Kamalaśīla on the status of reasoning, justification, and the sources of knowledge more generally, as well as the characterization of conventional truth (differences that, among other distinctions including the style of argumentation, are

---

54

general consensus is often mistaken and that a "judicious" or "discerning" person (*prekṣāvat*), that is, an ideal epistemic agent, differentiates between true and false conventions (*tathyasaṃvṛti* and *mithyāsaṃvṛti*), verifying the pragmatic efficacy of a given claim or theory by utilizing the epistemic instruments of *perception* and *inference*. The Mādhyamika may thus endorse the best scientific explanations of the day as conventional truths so long as their causal/pragmatic efficacy is empirically and/or rationally verifiable.[69]

Of course, the flexibility of the ontology is further constrained by the ultimate truth: no Mādhyamika can accept a theory that includes ontologically independent elements. While the conventional truth is revisable, the ultimate truth is fixed. On the other hand, reductionist ontologies, like the Abhidharma theory of conventionally real composites and ultimately real, basic "dharmas," requires a settled accounting of fundamentalia for a complete theory of their final ontology. So any revision made to accommodate scientific developments at the subatomic level will demand a revision of the Abhidharma ultimate truth. Mādhyamikas, on the other hand, need only revise the conventional truth. Surely a metaphysical picture with a fixed final ontology but revisable conventional truth is

---

implicated by the Tibetan doxographical categorizations of these figures as *Prāsaṅgika- and *Svātantrika-Mādhyamikas, respectively). As noted above, much of the secondary literature on conventional truth/reality to date, including many of the critiques of its coherence, have focused on the Candrakīrtian tradition. But the so-called *Svātantrika-Mādhyamikas have a richer set of explanatory resources at their disposal when it comes to characterizing conventional reality.

[69] This is, of course, consistent with the reclassification of testimony (*śabda*) as a subcategory of inference (*anumāna*) in the epistemological tradition of Dignāga and Dharmakīrti, a tradition followed by so-called *Svātantrika-Mādhyamikas, including Śrīgupta, Śāntarakṣita, and Kamalaśīla. This, again, should be contrasted with Mādhyamikas like Candrakīrti who rejected Dignāgian epistemology.

preferable. And given that there is, as yet, no incontrovertible scientific evidence for a bedrock of reality, so far so good.

Conclusion

I have offered a preliminary sketch of a two-stage model for understanding the metaphysical dependence structure that follows from Śrīgupta's rejection of ontological independence in his neither-one-nor-many argument, when taken together with his three criteria for conventional reality. The first level of analysis highlighted ways in which Madhyamaka metaphysical dependence relations deviate from standard accounts of the metaphysical grounding relation, ontological dependence relation, and proper parthood relation, insofar as it does not strictly honor the Hierarchy Thesis, instantiating irreflexivity and extendability but not strict asymmetry or transitivity. And given that the Madhyamaka dependence structure admits of both dependence loops as well as dependence chains that are indefinite but not actually infinite in length, this level of analysis also reveals how this structure represents an alternative model to the three standard categories of metaphysical foundationalism, infinitism, and coherentism.

The second level of analysis shows how this flexible ontology can support a contextualist form of the Hierarchy Thesis, allowing it to respect certain hierarchical structures (as well as non-hierarchical structures), but whatever structure is admitted must be earned by its pragmatic upkeep. And with its revisable theory of conventional truth, Śrīgupta's Madhyamaka can accommodate the best scientific explanations of the day, with the (rather sizable) caveat that it can never admit ontologically independent fundamentalia.

This two-stage analysis is not only a picture that Mādhyamikas like Śrīgupta would endorse, but it is my hope that it may also be instructive for at least gesturing toward some of the potential value that stands to be gained from engaging with non-standard metaphysical pictures of this kind.

2

CAN A MIND HAVE PARTS?

ŚRĪGUPTA ON MENTAL MEREOLOGY

There are plenty of skeptics about material simples, ranging from anti-atomist dualists to subjective idealists to physicalist gunk-theorists. But one is hard-pressed to find a community of skeptics about the simplicity of the mind. There's a common line of thought (often more of an intuition than an argument) that says we're entitled to infer the *metaphysical* unity of the mind from the phenomenal unity of our conscious experience.[70] This might also be accompanied by an appeal to the existence of a self that exists above and beyond the mind, which may underwrite the diachronic unity of consciousness.

Buddhist philosophers, of course, uniformly reject an enduring, unitary self, and regard momentary mental states taken collectively to constitute a mental bundle (*vijñānaskandha*) or a mental continuum (*cittasantāna*), which is only a unity by convention, or an accidental unity. That is, a mental bundle is a unity by virtue of the fact that we *treat* it as one thing, much like a bag of groceries. But what about a single, momentary mental

---

[70] Following Kant, this line of thought is often referred to as the "Achilles Argument," which Kant describes as "nothing like a mere sophistical play that a dogmatist devised in order to give his assertions a fleeting plausibility, but an inference that seems to withstand even the sharpest testing and the greatest scruples of inquiry" (1999, A351). In the end, however, Kant argues that the simplicity of the I is immediately intuited: "The proposition *I am simple* must be regarded as an immediate expression of apperception" (1999, A354–5). For a taxonomy of various versions of the Achilles Argument, see Lennon and Stainton (2008, 3–8), and for analysis of such arguments throughout the history of Western philosophy, ranging from Plato and Aristotle up to the twenty-first century, see Lennon and Stainton (2008).

state? Surely that counts as an indivisible unity. After all, unlike a mental continuum, which is plausibly conceptually divisible into temporal parts, a single moment of consciousness certainly *seems* to be indisputably simple, divisible into neither spatial nor temporal parts. In opposition to fellow Buddhists, Mādhyamikas beginning with Śrīgupta (ca. seventh–eighth century) deploy the anti-foundationalist "neither-one-nor-many argument" (*ekānekaviyogahetu*) to demonstrate that *no* mental entity—not a self or even a momentary mental state—counts as a true unity, whether synchronically or diachronically.

It's no exaggeration to say that the Madhyamaka attack on the unity of even a momentary mental state represents one of the most extreme positions on the unity of consciousness in the history of philosophy. While many have cast doubts on whether or not consciousness is necessarily and/or always unified, and others have questioned whether some number of conscious states are unified with other conscious states,[71] skepticism about *any* unified consciousness—even at a moment—is an uncommonly argued position indeed. Hume is often singled out as one of the few skeptics of any unified consciousness, though many commentators are quick to point out that he seems to have back-peddled on this point in the conclusion of his *Treatise*.[72] And in contemporary philosophy of mind, unified consciousness is largely taken for granted,[73] with debates centering instead on how best to describe or account for the unity of consciousness. Against this tide, Śrīgupta makes a case

---

[71] For skeptical stances on the unity of consciousness, see, for example, Nagel (1971), Davidson (1980), Dennett (1991, 1992), O'Brien and Opie (1998), and Rosenthal (2003).

[72] Garfield (2019a, chap. 12) represents an exception.

[73] Exceptions include Rosenthal (1986, 2002) and Garfield (2019b).

not merely for *skepticism* about unified consciousness, but instead argues that the unity of any mental state is just plain incoherent and, thus, metaphysically impossible.

But what's at stake in this question of the unity of a mental state? Unity is standardly regarded as the mark of ontological self-sufficiency, a status attributed to the likes of substances, God, and fundamental elements. In the history of Western philosophy, figures ranging from Aristotle to Leibniz were at least in agreement that being and unity are convertible; that is, whatever claims self-sufficient and substantial reality is necessarily a true and non-accidental unity.[74] Conversely, mere aggregates—like herds, heaps, and divisible hunks of matter—do not count as self-sufficient, substantially existent, or fundamental precisely because they are not true unities.

Similarly, according to Ābhidharmika Buddhists, to be a fundamental constituent of the world (*dharma*) is to exist substantially (*dravyasat*) rather than just nominally (*prajñaptisat*), and to exist substantially is to be an indivisible, partless unity; in other words, to be both a foundation and a substance is to be mereologically simple. Mādhyamikas like Śrīgupta agree with Ābhidharmikas that (i) whatever is fundamental has ontological self-sufficiency, or "ontologically independent being" (*svabhāva*), (ii) whatever is ontologically independent is necessarily a true unity, and (iii) whatever is a true unity is mereologically simple. Anything that has proper parts lays no claim to ontological independence, the thought goes, because whatever has proper parts depends for its existence on those parts, as a molecule does on its constituent atoms.

---

[74] As Leibniz puts it, "I hold this identical proposition, differentiated only by the emphasis, to be an axiom, namely, that what is not truly *one* being is not truly one *being* either. It has always been thought that 'one' and 'being' are reciprocal" (Letter to Arnauld, 30 April 1687, GP II 97/ PM 121).

But as a Mādhyamika, Śrīgupta's central commitment is, of course, that nothing in fact has ontologically independent being; everything depends for its existence on something else. This, he argues, follows from the fact that nothing is mereologically simple. Someone like Leibniz, for instance, rejects material simples only to appeal to mind-like simple substances to ground the multitude. [75] Similarly, Yogācāra Buddhist idealists like Vasubandhu reject material simples only to reduce all phenomena to the status of mere cognition (*vijñaptimātra*). Śrīgupta agrees with Yogācārins that there are no material simples, since whatever is material—whether it be a canyon or a quark—is necessarily extended, whatever is extended is divisible, and whatever is divisible has proper parts. But Śrīgupta insists that the mind is no different from matter in this regard: there *are* no simples, whether material or immaterial. But if there are no true unities, then there is nothing that has ontologically independent being, and thus, no metaphysical foundations.

In this chapter, I will reconstruct Śrīgupta's neither-one-nor-many argument, focusing on his subargument against the true unity—and thus the fundamentality—of *any* mental state. If the mind is not a true unity, then on Śrīgupta's definition, it must be divisible (at least conceptually) into proper parts. But is this application of mereological analysis to the mind justified? Can a mind really have proper parts in the same way as a molecule or a mountain? Śrīgupta argues just that, by way of an analysis of Yogācāra positions on the

---

[75] Where accepting the infinite dividedness of matter together with the simplicity of the mind drove Leibniz to argue for a form of immaterialist foundationalism, a similar pair of commitments motivated Descartes to endorse mind-body dualism, since whatever is simple cannot be the same in kind as whatever is infinitely divisible. As Descartes states, "the body is by its very nature always divisible, while the mind is utterly indivisible. For when I consider the mind, or myself in so far as I am merely a thinking thing, I am unable to distinguish any parts within myself; I understand myself to be something quite single and complete" (CSM 59, AT VII 85–86).

relation between the mind and mental content, cast as awareness (*jñāna*) and representations (*ākāra*). Later Yogācārins like Ratnākaraśānti (ca. 970–1030) and Jñānaśrīmitra (ca. 980–1050) would subsequently deploy their own iterations of the neither-one-nor-many argument, but restricting the scope of its subject, each claiming that there is some fundamental mental entity to which the argument does not apply. For Ratnākara, that foundation is awareness itself, or luminosity (*prakāśa*). For Jñānaśrī, it is a variegated representation that is non-distinct from awareness (*citrādvaitākāra*). In the course of presenting Śrīgupta's case against mental simples, I will motivate his application of mereological analysis to the mind in anticipation of subsequent responses to his influential argument by figures like Ratnākara, Jñānaśrī, and Ratnakīrti (ca. eleventh century) in late Buddhist India at Vikramaśīla, and, where relevant, I will also situate Śrīgupta's argument vis-à-vis contemporary debates in the philosophy of mind, such as the question of whether or not unified consciousness has experiential parts. As I will argue, Śrīgupta not only poses a serious challenge to his Yogācārin interlocutors, but he also carves out a lesser explored position in the contemporary space that calls into question the warrant for taking the metaphysical unity of consciousness for granted.

1. Śrīgupta's Neither-one-nor-many Argument

Throughout the history of Buddhist philosophy, the neither-one-nor-many argument has been recruited to reject the existence of a variety of (purported) pseudo-entities, from material objects to universals. In the *Ornament of the Middle Way* (*Madhyamakālaṃkāra*, MA)*,* Śāntarakṣita (ca. 725–788) applies the neither-one-nor-many

argument to an all-inclusive subject, cementing its place in Madhyamaka thought as one of the so-called four or sometimes five great arguments for the emptiness of ontologically independent being (*niḥsvabhāvatā*).[76] Śāntarakṣita's iteration of the argument is likely an expansion of that formulated in the *Commentary on the Introduction to Reality* (*Tattvāvatāravṛtti*, TAV, hereafter *Introduction to Reality*) by Śrīgupta, who is taken by the Tibetan tradition to be the teacher of Śāntarakṣita's teacher, Jñānagarbha (eighth century).

Śāntarakṣita's *Ornament of the Middle Way* bears a striking number of parallels to Śrīgupta's *Introduction to Reality,* including their nearly identical opening stanzas presenting the central inference from the reason of neither-one-nor-many. Two differences between these texts, however, bear on the present discussion. First, Śrīgupta's *Introduction to Reality* is more "internally focused," particularly when it comes to his subargument about the status of the mind: where Śāntarakṣita's iteration of the argument targets a host of Buddhist and non-Buddhist theories, as we will see, Śrīgupta directs his attention exclusively to Yogācāra accounts of the mind and mental content. Second, although Śāntarakṣita famously strikes a conciliatory tone at the end of the *Ornament of the Middle Way*, provisionally endorsing the Yogācāra doctrine that everything is mere cognition on a "conventional" level, Śrīgupta makes no such concession to Yogācārins.

The central inference of Śrīgupta's neither-one-nor-many argument runs as follows:

> [*s*] All that exists externally and internally [*p*] in reality lacks independent being, due to [*r*] lacking independent being that is either one or many, [*e*] like a reflection.[77]

---

[76] Śrīgupta's and Śāntarakṣita's neither-one-nor-many arguments are prefigured in Nāgārjuna's works; see, for example, RĀ 1.71, ŚS 32ab, and VP 33–39; see also Āryadeva's (third century) CŚ 14.19.

[77] TA 1: *phyi rol nang na gnas 'di kun // yang dag tu ni rang bzhin med // gcig dang du ma'i rang bzhin nyid // bral ba'i phyir na gzugs brnyan bzhin //* (PD 3121, 101). Cf. MA 1: *bdag dang gzhan smra'i dngos 'di dag // yang dag tu na gcig pa dang // du ma'i rang bzhin bral ba'i phyir // rang bzhin med de gzugs brnyan bzhin //* (Ichigō 1989,

63

Here, [*s*] denotes the subject of the inference (*pakṣa/dharmin*), which we might simplify as <all things>; [*p*] denotes the predicate (*sādhyadharma*), or the property to be proved: <lacking independent being>; [*r*] denotes the reason property (*sādhanadharma*): <neither one nor many>; and [*e*] denotes the example (*dṛṣṭānta*) in which the entailment relation between the reason property (neither one nor many) and predicate (not having independent being) is instantiated: a <reflection>. Formulated in the standard three-part inference of classical Buddhist logic, the argument runs as follows:

> Thesis (*pratijñā*):
>
> *All things lack independent being.*
>
> Major premise, statement of the entailment relation (*vyāpti*):
>
> *Whatever is neither one nor many does not have independent being.*
>
> Minor premise, predication of the reason property of the subject (*pakṣadharmatā*):
>
> *All things are neither one nor many.*

This argument might also be set up as an argument from dilemma as follows: *if anything has ontologically independent being, then it is either one or many*. The strategy is to demonstrate that there is nothing that can satisfy the consequent, and therefore (by modus tollens) there is nothing that can satisfy the antecedent.

What, then, are the conditions for predicating one or many, or unity/non-unity of a given subject? The operator or qualifier, "in reality" (*yang dag tu, *tattvatas*), in this formulation of the inferential argument indicates that the terms "one" and "many" in these

---

190; English translation 191). MA 1 is preserved in Sanskrit in Prajñākaramati's BCAP: *niḥsvabhāvā amī bhāvās tattvataḥ svaparoditāḥ / ekānekasvabhāvena viyogāt pratibimbavat //* (Vaidya 1960: 173).

premises should be understood as *true* unity and multiplicity. As Kamalaśīla defines it in his subcommentary on Śāntarakṣita's *Ornament of the Middle Way*, to be a true unity is to lack proper parts, viz. to be a mereological simple.[78] A proper part is something arrived at through physical or conceptual division. A true unity, then, is something that is neither physically nor conceptually divisible; it is mereologically simple and conceptually primitive. And a non-unity is anything that *does* have proper parts, the most basic of which are themselves simples. With these definitions, Śrīgupta stipulates a foundationalist structure on which the world bottoms out in mereological simples. Thus, we might rephrase the entailment relation as follows [Entailment Relation*]: *whatever is neither one simple nor many simples does not have independent being,* where "many" just signifies a determinate number greater than one. This excludes an indeterminate status, as well as that of a merely conventional unity/non-unity.

As a contradictory disjunctive pair, *<eka>* and *<aneka>* conforms to the logical, grammatical, and conceptual <F> and <not-F>, which is more precisely conveyed by the translation "unity or non-unity" as compared to the more popular translation "one or many." <Unity or non-unity> is an exhaustive, mutually exclusive pair of predicates. If *x* has independent being, then, on pain of violating the law of excluded middle, it is either unitary or non-unitary. As Śrīgupta states, "Since [unitary and non-unitary] are

---

[78] MAP *ad* k. 1: *cig pa zhes bya ba ni cha med pa nyid do //* (Ichigō 1985, 23). Kamalaśīla continues: *cig shos zhes bya ba ni du ma nyid de tha dad pa nyid ces bya ba'i tha tshig go //* (ibid.); "'Unity' refers to something's being partless. The alternative member of [this disjunctive predicate pair] is non-unity (*anekatva*), which is synonymous with 'consisting in discrete parts' (*bhedatva*)."

contradictory, existing [with independent being] in any other manner is surely untenable."[79]

However, the translation "one or many" has the virtue of more naturally reflecting another feature of this disjunctive pair. When understood as the *true* unity and *true* multiplicity of a foundationalist structure, *eka* and *aneka* are a peculiar contradictory predicate pair in that they share not only a conceptual priority relation but also a metaphysical priority relation: the *existence* of a non-unity presupposes the *existence* of some unities. A plurality presupposes singular things as its building blocks. Many and one, thus, stand in a mereological dependence relation of a composite on its proper parts. By contrast, while <non-blue>, for example, conceptually presupposes <blue>, the existence of something that is non-blue—say, a yellow school bus—certainly does not require the existence of something blue.

Śrīgupta points to this metaphysical priority relation between one and many, or a true unity and a true multiplicity, stating, "Given that [a multiplicity] consists of many unities, if one [viz. a unity] does not exist, the other [viz. a multiplicity] is also impossible."[80] Śāntarakṣita echoes this point, stating,

---

[79] TA 2cd: *'gal ba'i phyir ni rnam pa gzhan // yod par yang ni mi 'thad do //* (PD 3121, 101). To the same effect, Śāntarakṣita states in MA 62: *gcig dang du ma ma gtogs par // rnam par gzhan dang ldan pa yi // dngos po mi rung 'di gnyis ni // phan tshun spangs te gnas phyir ro //* (Ichigō 1989, 210); "Aside from unity and non-unity, an object's having some other classification is impossible, since it is established that these two are mutually exclusive." See Ichigō (1989, 211) for an alternative English translation.

[80] TAV *ad* k. 2b: *gcig mang po'i ngo bo yin pas de med na 'di yang mi srid pa . . .* (PD 3121, 102). Cf. MA 61: *dngos po gang gang rnam dpyad pa // de dang de la gcig nyid med // gang la gcig nyid yod min pa // de la du ma nyid kyang med //* (Ichigō 1989, 210); "Whatever object one analyzes, none has unity. Given that there is no unity, neither is there a non-unity." See Ichigō (1989, 211) for an alternative English translation. Here, Śāntarakṣita closely glosses Āryadeva's CŚ 14.19: *tasya tasyaikatā nāsti yo yo bhāvaḥ parīkṣyate / na santi tenāneke 'pi yenaiko 'pi na vidyate //* (Lang 1986, 132); "Whatever object one examines, none has unity.

66

Accordingly, "multiplicity" is defined as a composite of unities. If no unity exists, then neither does that [composite of unities] exist, just like if no trees exist, neither does a forest exist.[81]

Since unity is metaphysically prior to multiplicity, if unity is rejected, then multiplicity is *ipso facto* precluded. Just as a forest cannot exist without trees, a composite of simples cannot exist without simples. Since a multiplicity depends for its existence on some unities, it is not a proper candidate for ontologically independent being after all. The entailment relation is thus simplified as follows [Entailment Relation**]: *whatever is not a simple does not have independent being.* All Śrīgupta needs to do to establish the minor premise (the predication of the reason property of the all-inclusive subject) is to demonstrate that there are no true unities. The argument, thus, reduces to a refutation of simples.

Śrīgupta presents a systematic and exhaustive argument in his *Introduction to Reality*, which is represented in condensed form in the reconstruction below. P1 is the contrapositive of the entailment relation between the predicate and reason property, and the remainder of the premises work toward establishing the predication of the reason property of the all-inclusive subject (C2 and C3).

Given that there is no unity, neither is there a non-unity." See Lang (1986, 133) for an alternative English translation.

[81] MAV *ad* k. 61: *'di ltar du ma ni gcig bsags pa'i mtshan nyid do / gcig med na de yang med de / shing la sogs pa med na nags tshal la sogs pa med pa bzhin no //* (Ichigō 1985, 172).

Table 2: Neither-one-nor-many Argument with Attention to P6

P1 If there is anything that has independent being, then it is either a true unity or a true non-unity.	*Contrapositive of the entailment relation*
P2 There is a true non-unity only if there are true unities. P3 Something is a true unity if and only if it is a simple, viz. has no proper parts. P4 Whatever exists is material or mental or neither-material-nor-mental. ∴ C1 If there is a true unity, it is either a mental simple or a material simple or a neither-material-nor-mental simple. (from P3, P4) P5 There are no material simples. **P6 There are no mental simples.** P7 There are no neither-material-nor-mental simples. ∴ C2 There are no true unities. (from C1, P5, P6, P7) ∴ C3 There are no true non-unities. (from P2, C2)	*Proof of the predication of the reason property of the subject*
∴ C4 There is nothing that has independent being. (from P1, C2, C3)	*Thesis*

In what follows, I unpack Śrīgupta's subargument in support of P6, that *there are no mental simples*, which supports the proof of the predication of the reason property of the subject, i.e., that all things are indeed neither one nor many.

2. Against Immaterialist Foundationalism: A Guided Tour of Śrīgupta's Refutation of Mental Simples, P6 Unpacked

Śrīgupta's strategy in targeting mental simples differs from that applied to material simples. His case against material simples turns on the claim that matter is infinitely divisible. In other words, matter is necessarily extended, whatever is extended is divisible (into, say, a right side and a left side), whatever is divisible has parts (e.g., a right part and a left part), and whatever has parts is not simple. Since the subargument against material simples (a) targets a monadic subject, e.g., one hunk of matter, and (b) relies on the pair of

monadic properties <unitary> and <non-unitary>, we can regard the subargument against

material simples as the "monadic phase" of the argument.

With the monadic phase, Śrīgupta takes himself to have rejected material simples,

and thereby rejected substantially real matter and material foundations. So at this point in

the dialectic, he turns his critical attention to immaterialist foundationalism by targeting

Yogācāra theories of the mind and mental content. The subargument against mental

simples takes up a dyadic subject: awareness (*jñāna*) and its mental representation (*ākāra*),[82]

which are taken to jointly constitute a mental state. Here Śrīgupta utilizes two additional

property pairs to pose two dilemmas:

1. Distinct/non-distinct (*bheda/abheda*) Dilemma: are awareness and its

representation taken together one identical thing or distinct things?

2. Real/unreal (*satya/alīka*) Dilemma: are representations real in the same way that

awareness is supposed to be, or do they have some lesser, derivative ontological

status?

This pair of questions tracks a central point of contention in the *sākāra-nirākāra* (literally,

"having a representation" vs. "lacking a representation") debate among Yogācārins in late

Buddhist India.[83] To simplify, this debate centers on the question of whether or not

---

[82] *Ākāra* is a multivalent term in the history of Buddhist epistemology and philosophy of mind. See articles in Kellner and McClintock (2014) for recent scholarship on the variety of meanings of *ākāra* in different Indian Buddhist historical and philosophical contexts.

[83] In the context of realists about external objects, this same pair of terms, *sākāra-nirākāra*, signify respectively representationalist and direct realist/non-representationalist theories of perception. When defined in this way, representationalists, or Sākāravādins, would include Sautrāntikas, while direct realists, or Nirākāravādins, would include, for instance, Vaibhāṣikas and Naiyāyikas. This same set of terms is, however, also used to refer to divisions of Yogācāra theories of the status of mental content. Since Yogācārins reject the existence of real external objects, there is a sense in which all Yogācāra

cognition *necessarily* includes representations, with so-called Nirākāravādins like Ratnākaraśānti claiming that, although ordinary mental states include representations, the enlightened mental state of the dharmakāya whose cognition is necessarily veridical *cannot*. So-called Sākāravādins, on the other hand, like Jñānaśrī in his *Sākārasiddhiśāstra* and Ratnakīrti in his *Citrādvaitaprakāśavāda,* maintain that the mental states of both ordinary and ideal epistemic agents necessarily have intentional content in the form of (some kind of) representation.

The answer to the question of whether a mental state necessarily includes a representation bears on both the Distinct/non-distinct and Real/unreal Dilemmas posed by Śrīgupta. If all mental states are necessarily and intrinsically representational, as Sākāravādins would have it, then awareness and its representations are, in an important sense, non-distinct. But if that's right, then awareness and representations ought to enjoy some kind of ontological parity. By contrast, if the necessarily veridical cognition of an ideal epistemic agent (*pramāṇabhūta*) *cannot* include a representation, as Nirākāravādins claim, then awareness and a representation are not strictly identical, yet since the mental state of an ordinary epistemic agent *does* include a representation, they cannot be entirely distinct

---

accounts of *ordinary* cognition are necessarily non-representational: cognition cannot properly *represent* mind-independent external objects in the sense of genuine correspondence. Yet there is another sense in which ordinary cognition is necessarily representational for Yogācārins: direct acquaintance with (nonexistent) external objects is off the table, so the only possible objects of ordinary cognition are mental objects, viz. representations. Insofar as Yogācārins agree that the intentional object of *ordinary* cognition is a representation (*ākāra*), it would seem that they are all, in some respect, representationalists. The *sākāra-nirākāra* debate among Yogācārins instead centers on whether or not *enlightened* cognition necessarily has representational content, with Sākāravādins maintaining that it does and Nirākāravādins that it does not.

either. But if representations are never the object of veridical cognitions, then they ought not enjoy the same ontological status as awareness.

The Tibetan doxographical categories of *Satyākāravāda (*rnam pa bden par smra ba*) and *Alīkākāravāda (*rnam pa brdzun par smra ba*) (literally "theory/proponents of real representations" and "theory/proponents of unreal representations") derive from the Real/unreal Dilemma, though this pair of labels is not attested in extant Indic doxographies, where we instead find the Sākāravāda-Nirākāravāda distinction.[84] Late Indic and Tibetan Buddhist doxographers imposed these labels onto differing and internally diverse sets of thinkers in contriving subschools of Yogācāra, and one should be careful not to conflate these pairs of terms or the contentious points that their etymologies implicate. And while these doxographical labels can be problematic oversimplifications, they can nonetheless be useful conceptual touchstones in laying out the structure of Śrīgupta's argument. For ease of discussion, I will use the terms "real representationalist" and "unreal representationalist" to refer to two sets of views on the ontological status of representations, which in general overlap with *sākāra* and *nirākāra* positions, respectively.

---

[84] Tibetan doxographers commonly classified Yogācāra Sākāravādins as "proponents of real representations" (*Satyākaravādins) and Yogācāra Nirākāravādins as "proponents of unreal representations" (*Alīkāraravādins), despite the fact that these latter labels are unattested in Indian Buddhist writings. See Almogi (2010) for a helpful survey of these categorizations in late Indian Buddhist and early Tibetan doxographical writings. Indic texts presenting the classification of Yogācāra/Vijñānavāda into the two subschools of Sākāravāda and Nirākāravāda include: Advayavajra's (eleventh century) *Tattvaratnāvalī* and *Apratiṣṭhānadeśakavṛtti*, the *Paramārthabodhicittabhāvanākrama* ascribed to Aśvaghoṣa/Śūra, Candraharipāda's (eleventh century) *Ratnamālā*, Jñānavajra's (eleventh century?) *Tattvamārgadarśana*, Vajrapāṇi's (eleventh century) *Guruparamparākramopadeśa*, and the *Bka' gdams bu chos* ascribed to Atiśa (982–1054).

There are two dichotomies that can be derived here from *satya* and *alīka* (which I've translated thus far as "real" and "unreal") as applied to representations, one epistemological and the other ontological: (1) The epistemological dichotomy concerns the veridicality of the representational content of a cognition. (2) The ontological dichotomy concerns whether or not a representation *itself* is real. These two sets of dichotomies are not unrelated. In explaining the unreal representationalist position, Śāntarakṣita (MA 52/MAV *ad* 52) comments that, on this view, representations appear due to an error caused by the ripening of karmic latencies (and are thus non-veridical in an epistemological sense), but in actual fact, they do not exist (and are thus unreal in an ontological sense), likened to the illusion conjured by a magician. For the present purposes, however, I will bracket the epistemological dichotomy, since it is the ontological dichotomy that drives Śrīgupta's argument.

The real vs. unreal representation debate can be clarified by looking to how these two camps understand representations as fitting in the three natures (*trisvabhāva*) framework, an important ontological framework in Yogācāra thought comprised of: (i) the dependent nature (*paratantra-svabhāva*), (ii) the imagined nature (*parikalpita-svabhāva*), and (iii) the perfected nature (*pariniṣpanna-svabhāva*). This framework has a long and complex exegetical history, but for the present purposes, we might look to Vasubandhu's *Exposition on the Three Natures* (*Trisvabhāvanirdeśa*) for a basic overview, according to which (i) the dependent nature refers to *what* appears, which consists in some kind of causally conditioned mental activity, (ii) the imagined nature refers to the *manner* of appearance, i.e., the non-veridical superimposition of subject-object duality, and (iii) the perfected nature refers to the fact that

the dependent nature lacks the imagined nature, which is to say that appearances do not actually exist in the dualist manner in which they seem to; there is no appearing object that is distinct from a perceiving subject.[85]

Bodhibhadra (eleventh century) sums up the ontological dichotomy between real and unreal representations with reference to the three natures framework as follows:[86] According to real representationalists, representations are classified under the dependent nature and are thus existent bases for the erroneous manner of appearances, viz. the imagined nature. According to unreal representationalists, on the other hand, representations belong to the imagined nature, meaning that not only is the manner of appearance of representational content non-veridical but the representations themselves are also unreal. As with the question of whether or not consciousness is necessarily intentional, the question of the ontological status of representations has implications for the

---

[85] TN 2–4ab: *yat khyāti paratantro sau yathā khyāti sa kalpitaḥ / pratyayādhīnavṛttitvāt kalpanāmātrabhāvataḥ // tasya khyātur yathākhyānam yā sadāvidyamānatā / jñeyaḥ sa pariniṣpannaḥ svabhāvo 'nanyathātvataḥ // tatra kiṃ khyāty asatkalpaḥ tathaṃ khyāti dvayātmanā / tasya kā nāstitā tena yā tatrādvayadharmatā /* (La Vallée Poussin 1932–1933, 154); "*What* appears is the 'dependent' [nature], because of the fact it proceeds from causal conditions; *how* that [dependent nature] appears is the 'imagined' [nature], due to being mere superimposition. [TN 2] The fact that *what* appears is always bereft of [existing in] the *manner* in which it appears should be known as the 'perfected' nature, since it is unchanging. [TN 3] What appears there? The unreal construction. How does it appear? As having a dualistic nature. [TN 4ab]" For a discussion of these verses and an alternative translation, see Garfield (2002, chap. 7); on the disputed authorship of the TN, see Kapstein (2018).

[86] See Seitetsu Moriyama (1984, 10–11). Shinya Moriyama (2014, 340) summarizes the positions on this distinction of the later thinkers at Vikramaśīla: "Ratnākaraśānti held the position that mental images [viz. *ākāra*] are false (*alīka*) or nonexistent (*asat*) because they arise from wrong imagination (*abhūtaparikalpa*) and are invalidated or corrected by subsequent cognitions. Jñānaśrīmitra and Ratnakīrti held a different view, namely, that mental images are truly existent insofar as they arise in one's perception. Mental images are inseparable from the perceptual cognition itself, as they are not invalidated by any other cognition at the instant a perception occurs. . . . [T]he debate is primarily based on Dharmakīrti's theory of perception, especially his theorem of self-awareness (*svasaṃvedana*), which establishes a cognition's self-illuminating act and its inseparable relation to its mental image, together with the traditional Yogācāra theory of the three natures (*trisvabhāva*)."

relation between awareness and representations. The unreal representationalist's acceptance of an ontological disparity between unreal representations and real awareness entails an ontological distinction between representations and awareness. Indeed, if representations are unreal, then there are not actually two things to be distinct at all—there just *seem* to be. According to the real representationalist, however, the ontological parity of awareness and representations supports their non-distinction.

Śrīgupta runs a multitiered argument from dilemma utilizing the three predicate pairs real/unreal, distinct/non-distinct, and unitary/non-unitary, and relying heavily on the law of non-contradiction (LNC), according to which contradictory properties cannot be predicated of the same subject (*viruddhadharmādhyāsa*). Using the real/unreal and distinct/non-distinct disjunctive pairs, the logical space of relations between the mind and mental content is as follows:

View 1: Awareness is non-distinct from real representations. (Real Representationalist Views)

View 2: Awareness is distinct from real representations.

View 3: Awareness is non-distinct from unreal representations. (Unreal Representationalist Non-distinct Lemma)

View 4: Awareness is distinct from unreal representations. (Unreal Representationalist Distinct Lemma)

Figure 2: Range of Views on Awareness and Representations
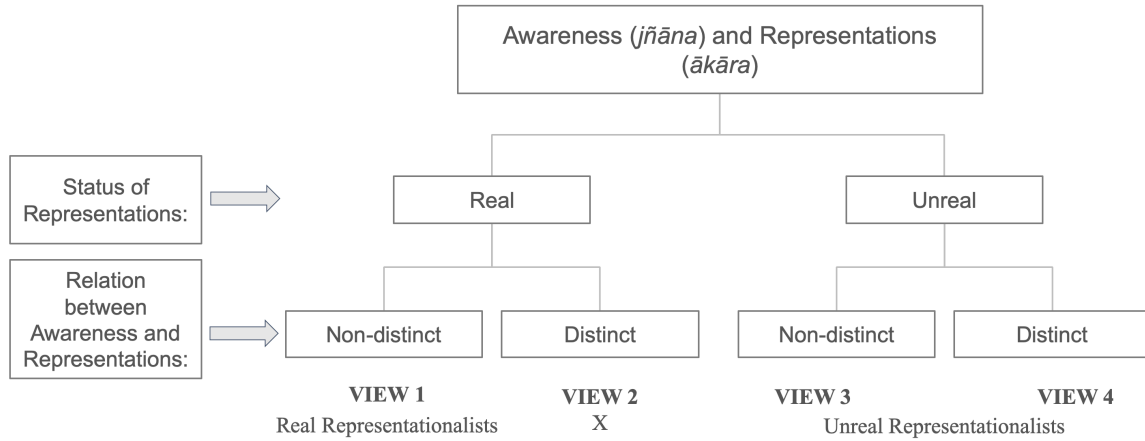
Beginning with the Real/unreal Dilemma, which asks whether or not representations are real in the same way that awareness is taken to be, Śrīgupta first addresses the position that they *are*. And as explained, the real representationalist view is linked with the view that awareness necessarily includes a representation, from which it is non-distinct. Thus, Śrīgupta begins with View 1, that *awareness is non-distinct from real representations.*

2.1 Rejection of Real Representationalist Views

View 1, *Awareness is non-distinct from real representations* (TA 4, cf. MA 46–51)

Śrīgupta's overarching strategy in targeting this first view is to go after the representation part of the mental state, arguing that it must be complex. Whoever maintains that awareness is non-distinct from, or identical with, a real representation owes some account of how they jointly constitute a truly unitary mental state. It's easy to take for granted the simplicity of the *subject* of a conscious experience, which, phenomenologically speaking, certainly doesn't *seem* divisible into parts. This motivates other intuitions like the thought that a swarm of bees could not itself be conscious, nor could a collection of scattered

neurons.[87] On the other hand, the *content* of our ordinary experience seems obviously complex. After all, in any given moment, I'm presented with a diverse array of sensory stimuli from various sensory sources. This would seem to threaten the simplicity of my conscious experience when taken as a necessary constituent of it. How, then, is a real representationalist to provide an account of a truly unitary mental state?

In contemporary philosophy of mind, we find two prevailing families of views on how unified consciousness is structured—the "experiential parts view" and the "no experiential parts view"—which may serve as helpful analogues for the present discussion.[88] On the experiential parts view, unified consciousness includes simpler experiences as its parts, whereas on the no experiential parts view, unified consciousness consists of a single, non-partite experience. Advocates of the experiential parts view will owe some explanation for how those parts are both individuated and at the same time "tied together" into a genuine unity. Proponents of the no experiential part view, on the other hand, must explain just how a conscious experience might be simple and nonetheless include a simultaneous diversity of experiential objects; e.g., the simultaneous sound of the kettle whistling, the visual presentation of the contents of my kitchen, the sensation of the cold kitchen floor under my bare feet, etc.

---

[87] See Putnam (1967) on the swarm of bees intuition and Unger (1990) on the brain separation intuition. See Barnett (2008, 334) for an argument that the "simplicity intuition," which he describes as the "naïve commitment to the principle that conscious beings must be simple," is the "source" of a host of other intuitions including these two, which have determined the debate space in contemporary philosophy of mind.

[88] Proponents of the experiential parts view include Lockwood (1989, 1994), Dainton (2005), Shoemaker (1996, 2003), Bayne and Chalmers (2003), and Bayne (2010), while proponents of the no experiential parts view include Searle (2002) and Tye (2003).

Śrīgupta sets up his argument against the real representationalists with an implicit dilemma targeting these two accounts of a unified conscious experience: if there exists a truly unified conscious experience, then it either has experiential parts or it does not. To explain the structure of this subargument from another angle: rather than asking whether *x* is one or many, as he did in the monadic phase of the argument, Śrīgupta instead considers whether each member of the dyadic subject—awareness and its representation— is one or many. In other words, are awareness and its representation *each* unitary or are they *each* manifold?

There are four logical possibilities for the numeric relation between awareness and representations: (1) one-to-many, (2) many-to-many, (3) one-to-one, and (4) many-to-one. The first three views were respectively dubbed in Tibetan doxographical literature (1) the variegated nonduality (*citrādvaita, sna tshogs gnyis med pa*) view, (2) the numerical parity of awareness and representations (*rnam shes grangs mnyam pa* or *bzung 'dzin grangs mnyam pa*) view, and (3) the half-egg (*sgo nga phyed tshal ba*) view.[89] The variegated nonduality (one-to-many) view is a version of the no experiential parts view, while the numerical parity of awareness and representations (many-to-many) view is a version of the experiential parts view. On the third, so-called "half-egg view" (or one-to-one view), a single operative awareness is paired with a simple represented aspect, like two halves of an egg, yet because awareness moves so quickly among the various simple represented aspects, it *seems as if* we are having a simultaneous experience of a complex representation, much like when we witness a

---

[89] While the "numerical parity of awareness and representations" and "half-egg" views are not attested doxographical labels in Indic writings to my knowledge, the views they signify are.

whirling firebrand, it *seems as if* we are having the visual experience of a flaming circle, when in fact we are watching a single burning ball in rapid motion. The half-egg (or one-to-one) view, like the one-to-many view, is another version of the no experiential parts view. Although Dharmakīrti, Śāntarakṣita, and others address this third view with iterations of the neither-one-nor-many argument, Śrīgupta exclusively targets the one-to-many view and the many-to-many view. However, as we will see, his argument also indirectly rules out this third view. Finally, the fourth permutation—many awarenesses to one representation—is not considered in any iteration of the argument that I am aware of, presumably because it is so unintuitive as to be deemed utterly implausible, finding no known defenders. To sum up, the logical range of real representationalist views is as follows:

1. One-to-many Lemma: Unitary awareness is non-distinct from its non-unitary real representations. (= variegated non-duality, *citrādvaita*, *sna tshogs gnyis med pa*)

2. Many-to-many Lemma: Non-unitary awareness is non-distinct from its non-unitary real representations. (= numerical parity of awareness and representations in Tibetan doxographies, *rnam shes grangs mnyam pa* or *bzung 'dzin grangs mnyam pa*)

3. One-to-one Lemma: Unitary awareness is non-distinct from its unitary real representation. (= "half-egg" view in Tibetan doxographies, *sgo nga phyed tshal ba*)[90]

---

[90] Śāntarakṣita considers and rejects this same set of three views in the context of representationalists who are realists about external objects.

4. ~~Many-to-one Lemma: Non-unitary awareness is non-distinct from its unitary real representation.~~

Figure 3: Real Representationalist View Unpacked
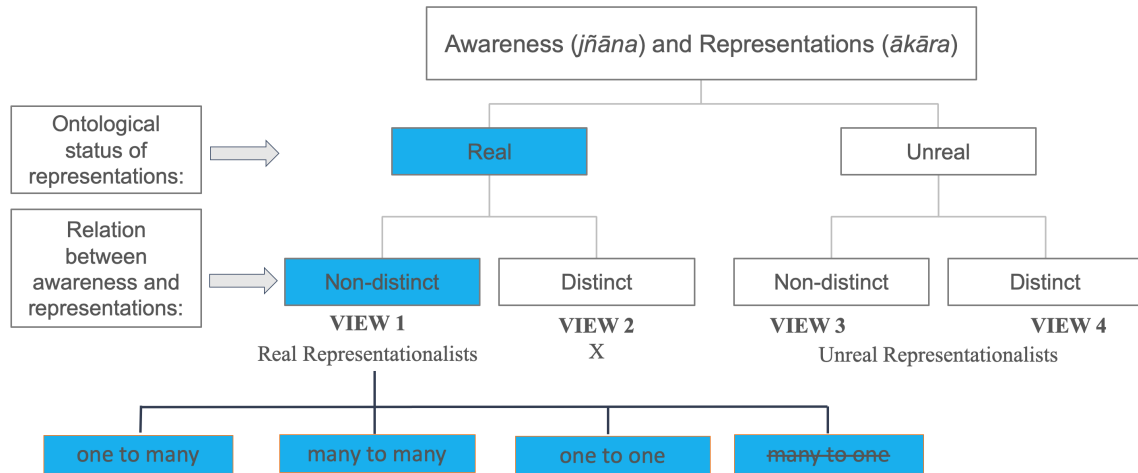
2.1.1 Argument against the One-to-many Lemma: Unitary awareness is non-distinct from non-unitary real representations (TA 4ab, cf. MA 50–51)

Awareness is not one—Argument from LNC

Śrīgupta first addresses the One-to-many Lemma, or the variegated nonduality view, with an argument to which figures like Jñānaśrīmitra and Ratnakīrti will later owe a response. This lemma corresponds to the rather intuitive view that awareness is unitary and yet real representations, which are non-distinct from awareness, are variegated. As suggested, this view might be helpfully thought of as a version of the no experiential parts view, according to which there is just one experience at a moment whose object or content is manifold. Critics of the no experiential parts view point to the internal contradiction of

something's being both simple and complex. Śrīgupta anticipates these challenges in his case against the one-to-many view, arguing as follows:

> Since representations that are non-distinct [from awareness] are variegated, the mind cannot be unitary.

> Awareness cannot be unitary because it is non-distinct from its non-unitary representations. Otherwise, on account of having contradictory properties, [awareness and its representation] would arise distinctly in two loci.[91]

Śrīgupta's strategy here is to drive a wedge between awareness and representations using the law of non-contradiction. In this argument, he takes the property, variegated (*citra, sna tshogs*), to entail non-unity. Whatever is variegated is necessarily complex, i.e., not simple. While applying mereological analysis to matter seems natural, predicating properties like divisibility and parthood of mental objects may seem unintuitive or even unwarranted. But according to Śrīgupta, the mereological dependence relation of a composite on its proper parts is "topic-neutral."[92] The composite-part relation is not limited to material things; anything that can be physically or even *conceptually* divided is partite—be it a molecule, a mongoose, or even a mind.[93] And this should not seem so strange. After all, just because

---

[91] TA 4ab and TAV *ad* k. 4ab: *rnam pa tha dad ma yin rnams // sna tshogs phyir na sems gcig min // shes pa gcig pu ma yin te // rnam pa du ma dang tha mi dad pa'i phyir ro // gzhan du na chos 'gal bar gnas pa gnyis tha dad par 'gyur te /* (PD 3121, 102–3).

[92] On issues concerning the topic-neutrality of mereology, see Johnston (2005), Varzi (2010), Donnelly (2011), and Johansson (2015).

[93] It's important to keep in mind here that conceptual divisibility is not equivalent to conceptual distinction. For instance, one might think that a single thing called by two different names (e.g., the same woman considered as a "sister" and "friend"), or a thing and its essence (e.g., matter and extension), are conceptually distinct, though ontologically identical. But neither the sister and friend nor matter and extension are conceptually divisible from one another in the mereological sense; extension is not a conceptually separable proper part of matter, nor vice versa. And neither is conceptual divisibility inclusive of the conceptual distinguishability of a formal aspect, as in, for instance, the distinguishability of a mouth from its smile. Rather, *x* is conceptually divisible in the mereological sense just in case there are conceptually isolatable proper parts *y*s that compose *x*, such that *x* is the sum of the *y*s. Someone like

we might not be capable of *physically* dividing some minute bit of matter doesn't prevent us from identifying its parts (e.g., its top, bottom, etc.). Likewise, a four-dimensionalist will find the conceptual division of a perduring object into its temporal proper parts to be perfectly reasonable, despite our inability to physically divide objects into temporal parts. And the proper parts of an abstract object like a Euclidean triangle may include its three sides and three angles.

So too, Śrīgupta would argue, the proper parts of a mental representation of a chair, for instance, may include the represented seat and represented legs.[94] Even a representation of something that may not *seem* "variegated," like a uniform patch of blue, is necessarily *phenomenally* extended, and is thus variegated insofar as it is conceptually divisible into phenomenal parts, e.g., a left side, a right side, and so on.[95]

If awareness and representations are indeed non-distinct, then a pair of contradictory properties like <unitary> and <non-unitary> should not be predicable of this one mental entity. That is, one cannot coherently maintain that awareness and

---

Descartes would, of course, maintain that the mind and thought are conceptually distinct, but not conceptually divisible, insofar as thought is the principal attribute, or essence, of the mind. Neither Śrīgupta nor his principal interlocutors would agree with this account of the relation between the mind and thought. A common account of the essence of mind in Śrīgupta's intellectual milieu would instead be reflexive awareness (*svasaṃvedana/svasaṃvitti*).

[94] See, for instance, Leech (2016) on taking seriously (rather than just metaphorically) the mereological structure of Kantian representations.

[95] This line of reasoning supposes that the representation is spatial, i.e., is phenomenally extended. However, were there some representation that belonged purely to some other modality—perhaps a sound, or a thought of an abstract object—then the argument could be run from a temporal perspective: there is no a temporally partless representation, maintaining that any moment of mind taken up for analysis necessarily has a beginning, middle, and end, each of which themselves have a beginning, middle, and end, *ad indefinitum*. See Prajñākaramati's BCAP *ad* k. 9.101 for an argument to this effect.

representations are identical, and yet awareness is one thing and a representation is many things.[96]

2.1.2 Argument against the Many-to-many Lemma: Non-unitary awareness is non-distinct from non-unitary real representations (TA 4cd, cf. MA 49, PV III.212)

Awareness is not many—Argument against mental atomism

Śrīgupta next addresses the Many-to-many Lemma, the numerical parity of awareness and representations view, arguing that it is equally incoherent for awareness to be manifold in conformity with its variegated representations. As suggested, the many-to-many view might be understood as a version of the experiential parts view, according to which unified conscious experience is constituted by as many experiences as there are objects of experience. While something seems right about saying that my experience of the sound of the kettle whistling and my simultaneous visual experience of my kitchen counter are not strictly the same experience, this view owes an account of how these simultaneous manifold experiential parts are fused or subsumed into a unified whole. Moreover, if experiences exist in a one-to-one relationship with objects of conscious experience, how are

---

[96] This argument might be formulated to run as follows:
(s) Awareness (p) is not unitary (r) because it is non-distinct from its variegated representation.

P1 If awareness is real (viz. has independent being), then awareness is either unitary or non-unitary.
P2 Real awareness is non-distinct from its real representation.
P3 Representations are variegated.
P4 Whatever is variegated is (conceptually) divisible.
P5 Whatever is divisible is not unitary.
∴ C1 Representations are not unitary. (from P3, P4, P5)
P6 Contradictory properties cannot be predicated of the same subject. (Law of Non-contradiction, LNC)
∴ C2 Awareness is not unitary. (from P2, C1, P6)

we to individuate objects to determine the precise number of experiential parts in a given experience? In other words, what counts as one object—my kitchen? My kitchen counter? The smallest visible unit of my kitchen counter? And does my tactile experience of the kettle and my visual experience of the kettle count as one or two experiences?

Śrīgupta's strategy in addressing the many-to-many view parallels that of his argument against material simples, and it might be described as an argument against mental atomism. Phenomenal simples, he argues, are impossible since any aspect of a representation that is fit to be an object of experience—say, a represented speck of blue—must be phenomenally extended; otherwise it would be imperceptible. And being phenomenally extended, it is not simple. Even our blue speck has a right side and a left side.[97] But let's grant for the moment that a phenomenal simple is possible. In that case, Śrīgupta relies on the premise that you can't get something extended from something unextended—even when that extension is merely phenomenal. Phenomenal simples as the building blocks of a variegated representation could no more constitute a phenomenally extended representation than could material simples constitute an extended hunk of matter. And given the metaphysical priority of unities to non-unities, if there are no simple representations, neither is there a true plurality of representations. If that's right, then awareness cannot exist in numerical parity to representations, since there can be no determinate number of representations with which it might correspond. As Śrīgupta puts it:

---

[97] To the contrary, Berkeley (*Principles* in *Works* vol. 2, 98) and Hume (*Treatise* 1.2.4), for instance, both argue for theory of a *minima sensibilia*, a kind of phenomenal atomism on which a perception is reducible to indivisible, unextended simples.

If, however, one contends that mind is also [non-unitary], corresponding with the number of its [manifold] representation, this is not the case:

> Since awareness that is a composite of many
> directional/perspectival parts
> is untenable, it is impossible that this is correct.

Were one to accept a simultaneous plurality of awarenesses, then representations—just like fundamental particles—could not compose a composite, as has been repeatedly established.[98]

Śrīgupta does not explicitly address the One-to-one Lemma.[99] However, his refutation of mental atomism obviates that alternative too: if a simple representation is impossible, then awareness from which it is non-distinct cannot be simple either. There is, thus, no coherent story to be told about a truly unitary mental state that either has or does not have experiential parts.

With the logical range of Yogācāra real representationalist views rejected, Śrīgupta concludes his argument against View 1, that *awareness is non-distinct from real representations*.[100]

---

[98] TA 4cd and TAV *ad* k. 4cd: *gal te 'o na sems kyang rnam pa'i grangs bzhin no zhe na / ma yin te // shes pa du ma'i phyogs bsags pa // mi rung phyir na 'thad par dka' //* [TA 4cd] *cig car du shes pa du ma khas len na / rnam pa rnams rdul phran bzhin du bsags par mi 'gyur te / ji ltar rtag tu bsgrubs pa bzhin no /* (PD 3121, 103). Cf. MA 49: *ci ste rnam pa'i grangs bzhin du // rnam par shes pa khas len na // de tshe rdul phran 'drar 'gyur ba // dpyad pa 'di las bzlog par dka' //* (Ichigō 1989, 206); "If consciousness were accepted [as non-unitary] in accord with the number of its representation, then, being similar to the case of fundamental particles, it would difficult to avoid that same analysis [here]." See Ichigō (1989, 207) for an alternative English translation. Cf. also PV 3.212.

[99] Śāntarakṣita does address the One-to-one Lemma, deducing absurd consequences of the view (MA 47–48; cf. PV 3.135, 198–99).

[100] This argument might be formulated to run as follows:
(s) Awareness (p) is not non-unitary (r) because it cannot be composed of directional parts.

P1 If awareness is real (viz. has independent being), then it is either a unity or a non-unity.
P2 Real awareness is non-distinct from its real representation.
P3 Representations are variegated. (Datum)
P4 Whatever is variegated is divisible.
P5 Whatever is (conceptually) divisible is not a unity.
∴ C1 Representations are not unitary. (follows from P3, P4, P5)

2.2 Rejection of Unreal Representationalist Views

Views 3 and 4 (TA 5, cf. MA 52–60)

Śrīgupta next moves to the other horn of the Real/unreal Dilemma, giving voice to an opponent who suggests that if representations are *unreal*, then their mereological status should not bear on that of *real* awareness.[101] In other words, the fact that some figment appearing to awareness is neither truly one nor truly many should not undermine the unity of awareness itself. In responding to the unreal representationalist view, Śrīgupta relies on the Distinct/non-distinct Dilemma, which says: if awareness is real, then it must be either distinct or non-distinct from unreal representations. These two lemmas are views 3 and 4 from our list of four possible views as I've formulated them:

P6 Contradictory properties cannot be predicated of the same subject. (LNC)

∴ C2 Awareness is not a unity. (follows from P2, C1, P6)

P7 Awareness is the same in number as its representation.

∴ C3 Awareness is a non-unity only if its representation is a non-unity. (follows from P7)

P8 A representation is a non-unity only if there are represented unities.

P9 Representations are phenomenally extended. (Datum)

P9 Whatever is phenomenally extended is (conceptually) divisible.

P10 A phenomenal representation is extended only if its basic phenomenal parts are extended.

∴ C4 There are no unitary representations.

P11 There is a phenomenally extended composite only if its parts are phenomenally extended.

P12 Whatever is phenomenally extended is (at least conceptually) divisible.

P13 Whatever is divisible is not a unity.

∴ C5 There are no non-unitary representations. (follows from P8, P9, P10, P11, P12, P13)

∴ C6 Awareness is not a non-unity. (follows from C3, C5)

∴ C7 Awareness is not real (viz. does not have independent being). (follows from P1, C2, C6)

[101] TAV *ad* k. 5: *gal te 'di rnams bden pa ma yin pa nyid khas blangs na 'di thams cad legs par smras pa ma yin nam zhe na /* (PD 3121, 103); [Objection:] "If it is accepted that these images are in fact unreal, then is it not the case that all this is well theorized?" Cf. MA 52: *ci ste ngo bo nyid du de'i // rnam pa 'di dag med pa ste // yang dag tu na rnam med pa'i // rnam par shes la nor bas snang //* (Ichigō 1989, 208); "But it [viz. awareness] is not actually endowed with these representations; they appear by virtue of an error to consciousness which is actually devoid of representations." See Ichigō (1989, 209) for an alternative English translation.

1. Non-distinct Lemma: Awareness is non-distinct from unreal representations. (View 3)

2. Distinct Lemma: Awareness is distinct from unreal representations. (View 4)

Figure 4: Unreal Representationalist Views
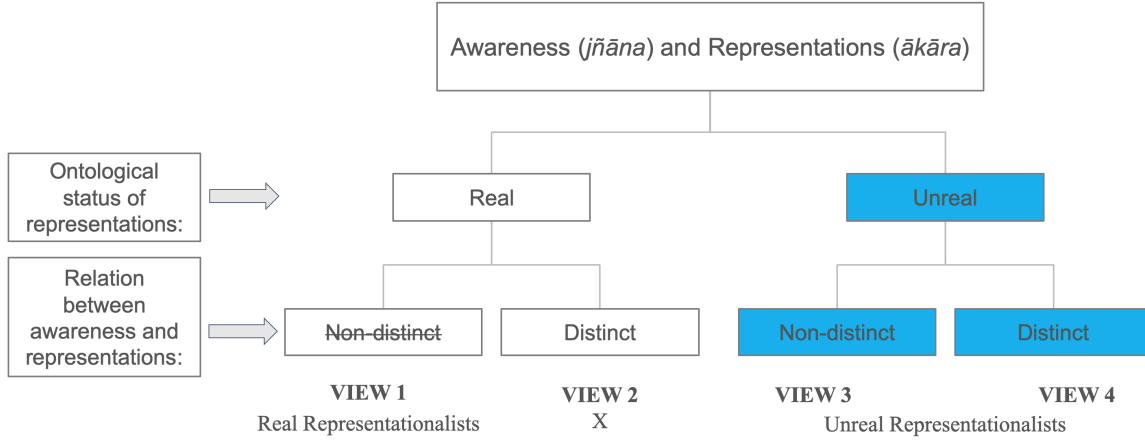
As we've seen, Śrīgupta's argument against *real* representationalists turns on a mereological analysis of awareness and its representation, which were taken to jointly constitute a given mental state. By contrast, in addressing *unreal* representationalists, Śrīgupta argues that these theorists cannot even get a coherent picture of their view up and running in order to apply any mereological analysis. As Śrīgupta sees it, the problem lies in the very fact that if one component of a mental state (as a dyadic subject) does not exist at all, then that threatens the reality of the entire composite mental state. Unreal representationalists like Ratnākaraśānti try to get around this problem by proposing that awareness and representations are distinct in one sense and non-distinct in another.

On Śrīgupta's dilemma, awareness and its representation are either distinct or non-distinct in the *strictest* sense; that is, they are either *numerically* identical—viz. one and the same particular object—or they are not numerically identical. Śrīgupta argues that neither

86

the Distinct Lemma nor the Non-distinct Lemma is available to the unreal representationalist by deducing unwanted consequences that follow from both alternatives. Moreover, according to Śrīgupta and unreal representationalists alike, awareness and representations cannot be both distinct and non-distinct *in precisely the same sense*, for this would violate the LNC. The question, then, is whether or not an unreal representationalist can differentiate between *kinds* of identity in a way that renders their story coherent. In what follows, I will consider a range of possible identity relations that an unreal representationalist might recruit in response to Śrīgupta's argument and assess their viability.

2.2.1 Argument against the Non-distinct Lemma (View 3): Awareness is non-distinct from unreal representations (TA 5ab1, cf. MA 53)

Awareness too would be unreal—Argument from LNC

To begin, Śrīgupta argues that if the Non-distinct Lemma were true, that is, if awareness were numerically identical with unreal representations, then given the LNC, awareness too would be unreal, which is, of course, an unwanted consequence for unreal representationalists. Śrīgupta argues,

> If representations were simply unreal,
> then absurd consequences would follow.[102]

---

[102] TA 5ab1: *rnam rnams mi bden nyid yin na / ha cang thal 'gyur /* (PD 3121, 103). Peking, Nar thang, and Gser bris ma editions read: *ha cang thal bar 'gyur*; Sde dge and Co ne editions read: *ha cang thal 'gyur ba*. TA 5 is not preserved unified or in consistent meter in any editions of the Tengyur. I emend the text in accordance with 'Gos lo tsā ba's *Rgyud bla ma'i 'grel bshad de kho na nyid rab tu gsal ba'i me long*, which cites the stanza as unified and in consistent meter as follows: *rnam rnams mi bden nyid yin na // ha cang thal 'gyur ma 'brel phyir // ji lta bur na nges snang 'gyur // de lta min na bden pa nyid //* (Mathes 2003, 181). 'Gos lo tsā

87

When one undergoes experiences, then not only would these [representations] be simply unreal, but it follows that cognition [viz. awareness of the representation] too would have [this same unreal] nature . . .[103]

But do unreal representationalists really endorse the *numerical* identity of awareness and representations? As an exemplar of the unreal representationalist view, Ratnākaraśānti maintains that awareness and unreal representations are non-distinct insofar as they both have the nature of luminosity,[104] viz. the "lights on" property that is the mark of conscious experience. If someone like Ratnākaraśānti were to endorse the numerical identity of awareness and its representation, it could only be contingent identity, like mud and the brick into which it has been baked. The mud and the brick can be said to share the same nature, despite the fact that the mud may persist (in a crumbled pile) once the brick has ceased. Likewise, if awareness and its representation are contingently identical, they may share the same nature of luminosity, despite the fact that awareness will persist at the state

ba attributes the stanza to the *Tattvāvatāra* (*De kho na nyid la ’jug pa*), rather than the *Tattvāvatāravṛtti*, so in ’Gos lo tsā ba’s time, the root text may still have been extant as an independent text.

[103] TAV *ad* k. 5ab1: *gal te nyams su myong na ’di dag kyang mi bden pa nyid yin te / rtogs pa’i ngo bo yang der thal bar ’gyur te /* (PD 3121, 103). Cf. MA 53: *gal te med na ji lta bur // de dag ’di ltar gsal bar tshor // de yi dngos las tha dad pa’i // shes pa de ’dra ma yin no //* (Ichigō 1989, 208); "If [representations] did not exist, how could they be clearly experienced, as is the case? Awareness is not distinct from its object [viz. representations] in that [proposed] manner." See Ichigō (1989, 209) for an alternative English translation.

This argument might be formulated as follows:
(s) Awareness (p) is not non-distinct from unreal representations (r) because it would follow that awareness is unreal.

P1 Representations are unreal.
P2 If awareness is real, then awareness is either distinct or non-distinct from representations.
P3 Contradictory properties cannot be predicated of the same subject. (LNC)
∴ C1 If awareness is non-distinct from representations, then awareness is unreal. (from P1, P2, P3)
∴ C2 If awareness is real, then awareness is distinct from its unreal representation. (from P2, C1)

[104] See, for instance, PPU, D 148a2–6 and MAU, D 227b4–7. For a helpful discussion of Ratnākaraśānti's line of reasoning on this point, see Shinya Moriyama (2011).

of enlightenment once representations have ceased. If this is indeed the picture that an unreal representationalist endorses, then Śrīgupta's subargument against this lemma goes through: so long as awareness and representations are purported to be numerically identical at some ordinary mental state t_1, then a claim about the reality of a representation at t_1 will necessarily bear on the reality of awareness at t_1.

But surely unreal representationalists have a weaker identity claim in mind. Ratnākaraśānti, for instance, argues that, despite their identical nature, awareness and representations are distinct insofar as the existence of representations is successfully refuted by the neither-one-nor-many argument, while that of awareness is not. Whoever claims that awareness and representations are distinct in one sense and non-distinct in another sense cannot coherently endorse their strict numerical identity, which demands sameness of all properties.

Perhaps, then, the unreal representationalist should say that representations and awareness are not strictly numerically identical, but instead (partially) qualitatively identical, i.e., the same type but not the same token. For instance, an existence claim about a candle flame in the kitchen need not bear on the existence of a bonfire on the beach, despite the fact that both fires share the same nature. Still, even if they share the same nature of luminosity, awareness and representations don't look like the same *kind* of thing. Perhaps instead the unreal representationalist ought to appeal to identical material constitution (like the co-located statue and clay), claiming simply that awareness and representations are composed of the same "stuff," viz. luminosity. Unreal representationalists might then distinguish awareness and representations based on their relative identity, claiming that

they are two distinct kinds of things that happen to be "co-located" and constituted by the same stuff, each possessing different identity and persistence conditions. Still, this picture looks strange in the case of awareness and its representation, which seem less like overlapping objects and more like relata of some kind. The unreal representationalists may instead intend a kind of adverbialism, that the representation is a *way* of being aware, or that awareness and its unreal representation are an object and its intrinsic, accidental property, akin to a mouth and its smile. One might say that a smile, which is asymmetrically ontologically dependent on the mouth, is modally distinct from the mouth, though not substantially distinct from it, sharing a single locus.

To sum up, if an unreal representationalist implausibly claims that awareness and representations are strictly numerically identical, then they are vulnerable to Śrīgupta's argument from the LNC. But if they instead endorse some alternative to numerical identity, like any of those just discussed, this move would place them squarely in the Distinct Lemma camp, to which Śrīgupta next turns.

2.2.2 Argument against the Distinct Lemma (View 4): Awareness is distinct from unreal representations (TA 5b2cd, cf. MA 57–58, 60cd)

Śrīgupta puts forward a two-stage argument against the Distinct Lemma which turns on two premises: [Relation Requirement] representations must stand in some kind of relation with awareness, and [Real Relatum Requirement] only real things can stand in relations. The two stages of the argument run as follows: (i) Unreal representations do not meet the Real Relata Requirement, so they cannot meet the Relation Requirement,

90

rendering them unable to account for our experience. (ii) Or else, if the unreal representationalist insists that their account *does* meet the Relation Requirement, then it follows that representations are real, and their position is contradicted.

Phase 1: Representations could not be related to awareness and thus could not appear with spatiotemporal determination—Reductio ad absurdum

If unreal representations are not strictly numerically identical with awareness, Śrīgupta reasons, representations must nonetheless stand in *some* kind of relation with awareness in order to account for our experience of them (Relation Requirement). Perhaps they stand in a causal relation or alternatively in some kind of non-numerical identity relation as just discussed. Only real things, however, can stand in relations with other real things, whether that be a causal relation or (*any* kind of) identity relation (Real Relata Requirement). A dragon can't cause a real forest fire, nor can it be partially qualitatively identical with some real winged animal. [105] A dragon can't claim identical material constitution with any real hunk of matter, nor can it stand in a subject-property relation with a real counterpart. But if unreal representations are mere figments that stand in no

---

[105] One might worry that a hallucination of a dragon, for instance, can have very real effects (engendering fear, motivating us to act, etc.), despite the fact that it misrepresents reality to us, and thus unreal things *can* stand in a causal relation. But the proposed relatum would presumably be the hallucination *qua* some real mental event, as opposed to the content represented in the hallucination, which does not correspond to any real referent. Likewise, the subject of this argument is the representation *itself*—not the represented content. So just as a nonexistent hallucination could not cause any fear or motivate any action, the thought goes, neither could any nonexistent representation stand in any relation with awareness. To borrow Descartes' formal reality vs. objective reality distinction, Śrīgupta takes the unreal representation view to mean that representations don't even have formal reality as ideas, and so any discussion of objective reality is baseless.

relation to awareness, that leaves no way to account for the spatiotemporal determinacy

and consistency that is the default of our ordinary experience. Śrīgupta argues as follows:

> Due to being unrelated [to awareness],
> how could [representations] determinately appear?[106]

> Furthermore, since unreal [representations] could not be related to [real] awareness,
> appearances [viz. representations] could not have [spatiotemporal] determinacy.[107]

---

[106] TA 5b2c: *ma 'brel phyir / ji lta bur na nges snang 'gyur /*. All editions of the Tengyur read: *ma 'brel <u>ba'i</u> phyir / ji lta bur na snang 'gyur* (PD 3121, 103). I emend the text with the omission of *ba'i* and the addition of *nges* in accordance with the stanza as preserved in 'Gos lo tsā ba's *Rgyud bla ma'i 'grel bshad* (Mathes 2003, 181). Here, I understand *nges snang* (**niyatapratibhāsa* or **niyatākāra*) to refer to appearing representational content that is spatiotemporally determinate, fixed, or delimited. 'Gos lo tsā ba glosses the meaning of this stanza as follows: *yul snang 'di dang rtog pa ma 'brel na ni rtog pa'i snang ba ma yin par ha cang thal bar 'gyur zhing / rtog pa de ji srid yod pa de srid du yul snang de yang nges par yod pa 'di ji ltar 'thad ces zungs pa yin te /* (Mathes 2003, 181); "If cognition were not connected with its appearing object, then it would absurdly follow that [it] could not be the appearing [object] of that cognition, and accordingly, it is logical that it is only insofar as the cognition exists that its appearing object also exists determinately."

[107] TAV *ad* k. 5b2c: *gzhan yang brdzun pa rnams dang shes par ma 'brel ba'i phyir snang ba nges pa dang ldan par mi 'gyur ro //* (PD 3121, 103).

This argument might be formulated to run as follows:
(s) Awareness (p) is not distinct from unreal representations (r₁) because it would absurdly follow that awareness and representations could not be related, and (r₂) because unreal representations that are unrelated to awareness could not appear to awareness with spatiotemporal determinacy.

P1 Representations are unreal.
P2 If awareness is real, then awareness is either distinct or non-distinct from representations.
P3 Contradictory properties cannot be predicated of the same subject. (LNC)
∴ C1 If awareness is non-distinct from representations, then awareness is unreal. (from P1, P2, P3)
∴ C2 If awareness is real, then awareness is distinct from its unreal representations. (from P2, C1)
P4 Awareness is distinct from representations only if awareness stands in some relation to representations.
P5 Whatever is unreal does not stand in any relation.
∴ C3 Awareness is not distinct from its unreal representations. (from P4, P5)
∴ C4 Awareness is not real. (from C2, C3)
Moreover:
P6 Representations appear with spatiotemporal determinacy. (Datum)
P7 Representations appear with spatiotemporal determinacy only if representations stand if some relation to awareness.
∴ C5 Awareness is not distinct from its unreal representations. (from P4, P5, P6, P7)
∴ C6 Awareness is not real. (from C2, C5)

Phase 2: Or else representations would be real—Proof by contradiction

If, on the other hand, the defender of the Distinct Lemma insists that representations *do* meet the Relation Requirement, then given the Real Relata Requirement, they will be forced to contradict their original claim that representations are unreal. Śrīgupta continues,

> Otherwise, [representations] would in fact be real.[108]

> Were one to accept [representations] as related [to awareness] in virtue of their appearing determinately, then [representations] would in fact be real, since otherwise it would be impossible [for them] to stand in either an identity relation (*tādātmya*) or a causal relation (*tadutpatti*) [with awareness].[109]

---

[108] TA 5d: *de lta min na bden pa nyid //* (PD 3121, 103). On TA 5d and TAV AŚ 2, cf. MA 58cd, 60cd: *rgyu dang ldan na gang zhig gis // gzhan gyi dbang las bzlog par 'gyur //. . . de yi mthu yis byung na ni // de yang gzhan gyi dbang nyid do //* (Ichigō 1989, 210); "If [a representation] *does* have a cause, then how could its being dependent (*paratantra*) [and thus not unreal] be avoided? [58cd] . . . If [a representation] arises through the force of [error], then in that case too it is surely dependent (*paratantra*). [60cd]" See Ichigō (1989, 211) for an alternative English translation.

[109] TAV *ad* k. 5d: *nges par snang ba'i dbang gis 'brel par khas len na ni bden pa kho nar 'gyur te / gzhan du na de'i bdag nyid dang de las byung ba mi srid pa'i phyir ro //* (PD 3121, 103). Cf. MA 57: *gang phyir de yod nges tshor ba // shes dang 'brel ba ci zhig yod // bdag med de yi bdag nyid dang // de las byung ba ma yin no //* (Ichigō 1989, 210); "How could there be any relation between a determinately perceived [unreal representation] and awareness? An unreal [representation] could neither stand in an identity relation nor a causal relation with [awareness]."

This argument might be formulated to run as follows:
(s) Awareness (p) is not distinct from unreal representations (r) because it would absurdly follow that (unreal) representations would be real.

P1 Representations are unreal.
P2 If awareness is real, then awareness is either distinct or non-distinct from representations.
P3 Contradictory properties cannot be predicated of the same subject. (LNC)
∴ C1 If awareness is non-distinct from representations, then awareness is unreal. (from P1, P2, P3)
∴ C2 If awareness is real, then awareness is distinct from its unreal representations. (from P2, C1)
P4 Awareness is distinct from representations only if awareness stands in some relation to representations.
P5 Whatever is unreal does not stand in any relation.
∴ C3* Awareness is distinct from representations only if its representations are real. (from P4, P5)
∴ C4* Awareness is not distinct from representations. (from P1, C3)
∴ C5* Awareness is not real. (from C2, C4)

Notice that the unwanted consequence in the second phase of the argument against the Distinct Lemma—that representations would be real—is in fact View 2 from our list of four possible Yogācāra views on the mind and mental content:

View 1:   ~~Awareness is non-distinct from real representations. (Real~~
           ~~Representationalist Views)~~

View 2:   Awareness is distinct from real representations.

View 3:   ~~Awareness is non-distinct from unreal representations. (Unreal~~
           ~~Representationalist Non-distinct Lemma)~~

View 4:   ~~Awareness is distinct from unreal representations. (Unreal~~
           ~~Representationalist Distinct Lemma)~~

Up to this point in the argument, the first, third, and fourth views have been considered and rejected in that order. View 2, that representations are real and distinct from awareness, is the only logical possibility that remains.

Figure 5: View 2 as the Final Alternative

2.3 Case against Nondual Awareness

2.3.1 Argument against View 2: Awareness is distinct from real representations

Awareness could not be nondual—Reductio ad absurdum

While the implication that representations would be real if they were distinct from awareness is already an unacceptable consequence for the unreal representationalist, Śrīgupta makes the further point that accepting representations as real and distinct from awareness is at odds with the thesis common to Yogācārins of all stripes: that the subjective representation (*grahakākāra*) and the objective/intentional representation (*grāhyākāra*) are, in actual fact, nondual (*advaya*). View 2 is, thus, no real option for any Yogācārin, given their commitment to the ultimate absence of a real distinction between subject and object, viz. nondual awareness. After all, awareness could not be nondual with representations from which it is really distinct. Śrīgupta thus shuts down this final option, arguing,

> Given that [according to you] it is inadmissible to say that [unreal] matter is related [to awareness], since that [unreal representation] would likewise [be unrelated to awareness], [you must] accept representations as real; [but] in that case, awareness that is devoid of a subject and object [distinction] would be untenable.[110]

---

[110] TAV *ad* k. 5d: *gzugs ni 'brel zhes mi 'thad na / de lta yin na rnam pa bden par khas blang ba'i phyir shes pa gzung ba dang / 'dzin pa dang bral mi 'thad do //* (PD 3121, 103).

This argument might be formulated to run as follows:

(s) Awareness (p) is not nondual (r) because it would follow that representations are real.

P1 Representations are unreal.
P2 If awareness is real, then awareness is either distinct or non-distinct from representations.
P3 Contradictory properties cannot be predicated of the same subject. (LNC)
∴ C1 If awareness is non-distinct from representations, then awareness is unreal. (from P1, P2, P3)
∴ C2 If awareness is real, then awareness is distinct from its unreal representations. (from P2, C1)
P4 Awareness is distinct from representations only if awareness stands in some relation to representations.
P5 Whatever is unreal does not stand in any relation.
∴ C3* Awareness is distinct from representations only if its representations are real. (from P4, P5)
P6 If awareness is distinct from representations, then awareness is not nondual with representations.
P7 Awareness is nondual with representations. (Yogācāra commitment)
∴ C4* Awareness is not distinct from representations. (from P1, C3)
∴ C5* Awareness is not real. (from C2, C4)

With this, Śrīgupta takes himself to have shown that all four logically possible ways that a mental state could exist as a true unity are, in fact, incoherent and thus metaphysically impossible.

> View 1:   ~~Awareness is non-distinct from real representations. (Real Representationalist Views)~~
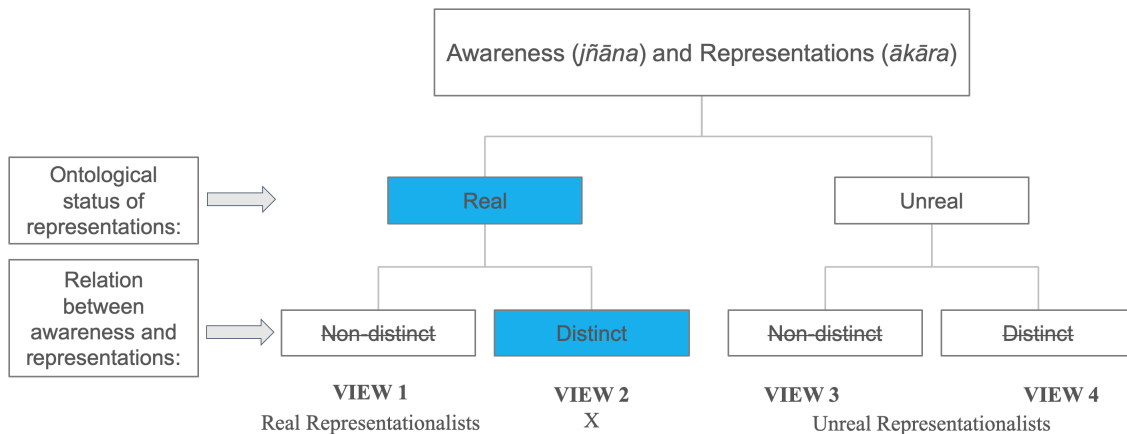>
> View 2:   ~~Awareness is distinct from real representations.~~
>
> View 3:   ~~Awareness is non-distinct from unreal representations. (Unreal Representationalist Non-distinct Lemma)~~
>
> View 4:   ~~Awareness is distinct from unreal representations. (Unreal Representationalist Distinct Lemma)~~

Figure 6: All Four Possibilities Are Rejected
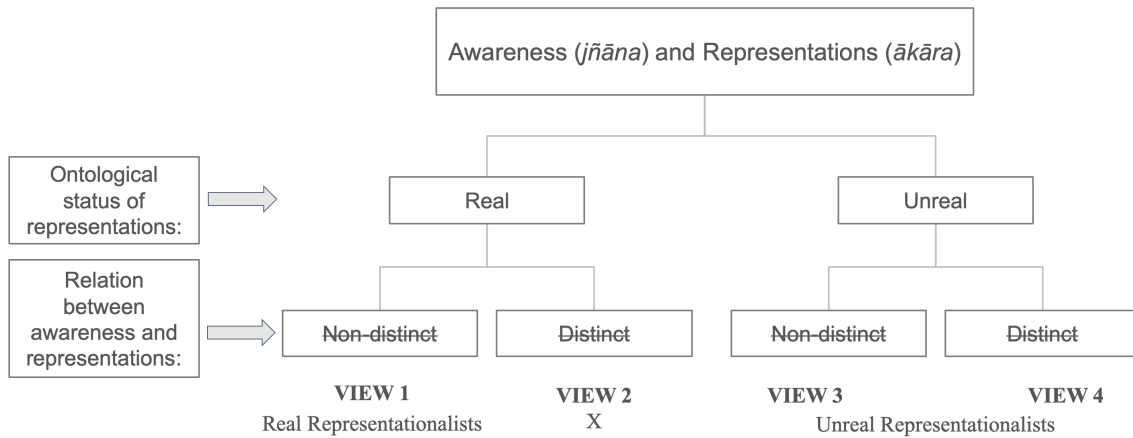
At this point, Śrīgupta supposes that his immaterialist foundationalist opponent may object that this whole exercise of conceptually distinguishing between awareness and its content as subjective and objective features of a mental state and then analyzing their relation is entirely misguided. We might imagine a Yogācārin interlocutor retorting to the above line

96

of reasoning, "Of course your inquiry into the nature of dualistic awareness showed it to be incoherent; the subject-object dualism of awareness and its representations is merely an error. In actual fact, the mind is just one simple entity: nondual awareness."[111] In response to this line of thought, Śrīgupta takes up the subject of nondual awareness as the last resort for an immaterialist foundationalist who claims that the mind or any mental state exists as a true unity.

2.3.2 Nondual Awareness Is Incoherent (TA 6, TAV AŚ 3–4)

Śrīgupta argues that the very concept of nondual awareness is incoherent, being internally contradictory. In brief, if nondual awareness were truly *nondual*, then it could not meet the definition of awareness, and if it were truly *aware*, then it could not meet the definition of nonduality. This line of reasoning presupposes a definition of awareness that requires it to have an intentional object. To be *aware* is necessarily and by definition to be aware *of* something; this is the "intentionality demand" on awareness.[112] Moreover,

---

[111] As Kamalaśīla puts the objection in his MAP *ad* k. 60: *'o na yongs su ma dag pa'i gnas skabs na shes pa na tshogs su snang ba brdzun pa kho na yin du chug kyang yongs su dag pa'i gnas skabs na ni rang bzhin gcig pa gnyis su med pa'i ngo bo kho nar 'gyur te /* (Ichigō 1985, 159); =AAA (Wogihara, 1932–1935, 633.24–6): *tarhi apariśuddhāvasthāyāṁ citrāvabhāsam alīkam eva jñānam, pariśuddhāvasthāyāṁ bhrāntivigamād advayarūpam evaikasvabhāvaṁ bhaviṣyatīti*; "Well, although in the impure state, awareness simply consists in unreal variegated appearance, in the completely pure state, there is simply the unitary nature [of awareness] that has a nondual character. Thus, your reason [that awareness is neither-one-nor-many] is unestablished due to being doubtful (*saṃdigdhāsiddha*)."

[112] Vasubandhu defines *vijñāna*, which Śrīgupta uses interchangeably with *jñāna*, as follows in his PSk: *vijñānaṃ katamat / ālambanaṃ vijñaptiḥ //* 112 //; "What is consciousness? The cognition of a phenomenal object." In his PSkV, Sthiramati explains that here, "'Phenomenal object' refers to [any] object of the mind or of a mental activity, including any of the six kinds from matter to mental objects. The 'cognition' of that [phenomenal object] refers to apprehending, being aware of, and understanding"; *ālambanaṃ cittacaittaviṣayaḥ / sa punaḥ ṣaḍ-prakāraḥ / rūpaṃ yāvad dharmāḥ / tasya vijñaptir grahaṇam avabodhaḥ pratipattir ity arthaḥ /; de yang rnam par rig pa ni 'dzin pa dang / rtog pa dang khong du chud pa zhes bya ba'i tha tshig go /* (Kramer 2013, 89). Similarly, Vasubandhu defines the *vijñānaskandha* as follows in AKB *ad* 1.16a: *vijñānam*

97

Śrīgupta takes nonduality to entail simplicity, where a simple cannot be conceptually divisible into subjective and objective aspects; this is the "simplicity demand" on nonduality.

Phase 1: Nondual awareness could not be aware

Śrīgupta begins by arguing that if nondual awareness were truly nondual, then it could not qualify as *awareness*. He asks rhetorically, "But if [awareness] were free from duality, then how could it be *aware*?"[113] That is, if a mental state were simple and thus necessarily devoid of subjective and objective conceptual parts, then how could it meet the intentionality demand on awareness? Awareness would have nothing to be aware *of*. As Śrīgupta sees it, awareness in the absence of an object of awareness is nonsensical, much like knowing in the absence of an object of knowledge is nonsensical. Suppose I claim, "I know." And you then ask me, "You know what?" And I reply, "nothing." You'd surely think I've lost the plot. Knowing nothing cannot rightly be called "knowing" at all; the same goes for awareness. Yet if a mental state necessarily includes both subjective awareness and an intentional object, then being conceptually divisible into these two

---

prativijñaptiḥ / [1.16a] *viṣayaṃ viṣayaṃ prati vijñaptir upalabdhir vijñānaskandha ity ucyate /* (Pradhan 1967, 11.6–7); "'Consciousness is individual cognition.' [AKK 1.16c] It is said here that the consciousness aggregate is the understanding that consists in the cognition of individual objects."

[113] TAV *ad* k. 6ab1: *'on te gnyis las nges par grol ba yin na / de ni ji ltar na shes pa yin /* (PD 3121, 104). Cf. MA 59: *de med na ni shes de yang // rnam pa med pa nyid kyis 'gyur // shel sgong dag pa 'dra ba yi* // shes pa rab tu tshor ba med //* (Ichigō 1989, 210). *MAV reads *yi*, while MA reads *yin*; I follow the MAV: "If that [representation] were unreal, then surely awareness would lack representations. Awareness that is like a clear crystal ball has no perception [of anything]." See Ichigō (1989, 211) for an alternative English translation. Cf. also MA 55: *rnam pa 'di la shes pa'i don // dngos su 'thad pa ma yin te // shes pa'i bdag dang bral ba'i phyir // rnam mkha'i me tog la sogs bzhin //* (Ichigō 1989, 208); "The meaning of awareness would not actually be apt with respect to this [unreal] representation, since awareness would lack its own nature [as something that has an intentional object], just as [awareness of] a sky flower, etc., [is meaningless]." See Ichigō (1989, 211) for an alternative English translation.

aspects, it could not count as "nondual" after all. Thus, Śrīgupta argues, the intentionality demand on awareness is incompatible with the simplicity demand on nonduality.[114]

Phase 2: If nondual awareness were reflexively aware, it could not be nondual

One response to phase 1 is to take issue with the definition of awareness as something that necessarily has an intentional object (P1). Another route is to maintain that nondual awareness still meets the intentionality demand because, owing to its intrinsic property of reflexivity, it takes *itself* as its object. Śrīgupta charges that recourse to the reflexivity of awareness merely shifts the problem. If awareness takes itself as an object, then, having smuggled in objective and subjective features, such awareness would once again fail the simplicity demand. The text reads,

> If one says, "Well, it is due to reflexive awareness [that nondual awareness is rightly said to be 'aware']."

[Response:] That is not so, since [in that case, awareness] could not be nondual.

Upon analysis, [this] is incorrect.

The analysis of [nondual reflexive awareness] as the true nature of the cognitive object and agent could not be the true state of affairs.[115]

---

[114] This argument might be formulated to run as follows:
(s) Awareness (p) is not nondual (r) because it is aware.

P1 Awareness is something that cognizes an object, viz. has an objective feature. (Definition)
P2 Whatever is nondual does not have subjective or objective features. (Definition)
∴ C Whatever is nondual is *not* an awareness. (from P1, P2)

[115] TA 6b2cd and TAV *ad* k. 6b2cd: *'on te* [conj.: *//*] *rang* [D, C: om. *rang*] *rig pa'i phyir ro zhe na /* [TA 6b2c] *de lta ma yin te / gnyis su med pa mi rung ba'i phyir ro // gal te brtags na yang dag min /* [TA 6d] *rig bya dang rig pa po'i dngos por brtag pa ni / de kho na nyid ma yin par 'gyur ro /* (PD 3121, 104).

This argument might be formulated to run as follows:
(s) Awareness (p) is not nondual (r) because it is reflexively aware.

Śrīgupta then summarizes his argument against the coherence, and thus the metaphysical possibility, of nondual awareness as a true unity as follows:

> Due to lacking a cognitive object, [nondual awareness] could not cognize anything else.
> Due to being nondual, [awareness] could not [cognize] itself.
> Upon examination, [nonduality] cannot be the nature [of awareness].
> Tell me, what other option is there?[116]

Phase 3: Claiming that nondual awareness is inexplicable is a sophistry

Śrīgupta gives voice to a final objection from the defender of nondual awareness who accepts real nondual awareness, yet refuses to defend it philosophically on the grounds that it is inexplicable, or inexpressible. This stance, Śrīgupta alleges, is simply an unpersuasive cop-out:

> If one [idly] stretches out one's legs, saying,
> "That state [of nondual awareness] is inexplicable,"
> that will [only] satisfy gullible individuals
> who form beliefs based on what is commonly accepted.[117]

P1 Awareness cognizes itself.
P2 Whatever is nondual does not have subjective or objective features. (Definition)
P3 Whatever cognizes something has an objective feature.
∴ C Whatever is an awareness is *not* nondual. (from P1, P2, P3)

[116] TAV AŚ 3: *rig bya med phyir gzhan mi rig // gnyis su med phyir bdag nyid min // brtags na yang dag nyid mi 'gyur // rnam pa gzhan gang yin pa smros //* (PD 3121, 104). TAV AŚ 3abc is cited in *Vipaśyanotpādanopāya*, a work of unknown authorship, with an alternate, preferable translation: *shes bya med phyir gzhan rig min // gnyis su med phyir bdag rig min // gal te brtags na yang dag min //* (PD 3611, 1462). The translation of TAV AŚ 3b here makes clear that we are to understand *bdag/bdag nyid* as the object of *rig*; I have thus accounted for *rig* in brackets in my translation.

[117] TAV AŚ 4: *gal te gnas skabs de bstan par // bya min zhe na rkang brkyang ste // grags pas dad pa bskyed pa yi // dad ldan rnams la mdzes pa yin //* (PD 3121, 104).

With the rejection of the reality of nondual awareness, Śrīgupta takes himself to have defeated all possible Yogācāra accounts of the true unity of any mental state.

2.4 Conclusion of the Subargument against Mental Simples (TA 7ab)

Śrīgupta concludes his refutation of the unity of mind by once again appealing to some version of the LNC to argue that whatever applies to the mind (*citta*) also applies to the mental activities (*caitta*):

> Since the mind does not exist like that [i.e., as truly unitary or non-unitary], mental activities [that are truly unitary or non-unitary] are also denied.

> Thus, due to sharing the same fate, given that there is no [such] mind, it is indeed easy to understand that there also are no [such] mental activities.[118]

Given that the variety of mental activities are merely modes of the mind, they "share the same fate" (**ekayogakṣama*), meaning that if the mind falls short of some ontological status, then mental activities are not entitled to that status either. If the mind is not a true unity, neither are mental activities. Therefore, there are no mental unities. And given the metaphysical priority of unities to non-unity, neither are there any mental non-unities.

Conclusion

As complicated as this argument is, it's only the beginning of the story, both historically and philosophically. Śrīgupta's iteration of this subargument set the stage for centuries of intra-Buddhist debates concerning the nature and ontological status of the

---

[118] TA 7ab and TAV *ad* k. 7ab: *sems ni de ltar med pa'i phyir // sems las byung ba rnams kyang bsal //* [TA 7ab] *de ltar na grub pa dang bde ba gcig pa'i phyir / sems med na sems las byung ba rnams kyang med par khong du chud par sla ba nyid do //* (PD 3121, 104).

mind and mental content. Not only did subsequent Indian Mādhyamika thinkers—most notably Śāntarakṣita, Kamalaśīla, and Jitāri (late tenth century)[119]—pick up and elaborate on this line of reasoning, but some of the most influential Tibetan Buddhist philosophers throughout the subsequent millennium authored texts devoted to this argument, including Phya pa Chos kyi seng ge (1109–1169),[120] Tsong kha pa Blo bzang grags pa (1357–1419),[121] and 'Jam mgon 'Ju Mi pham mgya mtsho (1846–1912).[122] Of equal historical significance is the range of Yogācāra authors who felt compelled to respond to this argument in competing ways, catalyzing them to refine their theories of the ontological status of representations and the mind, develop subtler accounts of the relation between the mind and mental content, and clarify the criteria for existence itself. As I've gestured to throughout, Jñānaśrīmitra and Ratnākaraśānti represent two of the central figures advancing these debates at Vikramaśīla in the eleventh century.[123]

---

[119] See Jitāri's SVB *ad* k. 7

[120] See Phya pa's *Dbu ma rgyan gyi 'grel pa*, *Dbu ma snang ba'i 'grel pa*, and *Dbu ma shar gsum gyi ston thun*. See Hugon (2015) for an analysis of Phya pa's iteration of the neither-one-nor-many argument.

[121] Tsong kha pa's *Dbu ma rgyan gyi brjed byang*, as recorded by his student Rgyal mtshab Dar ma rin chen (in *Gsung 'bum Tsong kha pa* [*zhol*], vol. 16, 543–64, Zhol par khang) and *Dbu ma rgyan gyi zin bris* (in *Gsung 'bum Tsong kha pa* [*zhol*], vol. 15, 783–814, Zhol par khang). See Blumenthal (2004) for an English translation of the *Dbu ma rgyan gyi brjed byang*. See also Tillemans (1982, 1983, 1984).

[122] Mi pham's *Dbu ma rgyan gyi rnam bshad 'jam dbyangs bla ma dgyes pa'i zhal lung* (in *Gsung 'bum Mi pham rgya mtsho*, vol. 12, 11–426, Rdzong gsar dgon pa'i par khang); see Doctor (2004) for an English translation of this text.

[123] It has not been possible to do justice to their views here; these authors have been referenced simply as conceptual touchstones to help motivate Śrīgupta's argument. See Tomlinson (2019) for an in-depth study of the debate between Jñānaśrīmitra and Ratnākaraśānti on the status of *ākāras*. It is also important to note that Śrīgupta picks up threads of versions of this subargument going back at least to Dharmakīrti and developed by Dharmakīrti's many commentators for the next half millennium. Prajñākaragupta (ca. mid eighth to early ninth century) is one commentator in particular whose treatment of nondual awareness in relation to this argument merits more careful consideration vis-à-vis Śrīgupta's treatment

This subargument is also just the beginning of the story when it comes to Śrīgupta's account of how the mind *does* exist. After all, this argument does not set out to prove that the mind doesn't exist at all, but rather to show that the mind lacks ontological independence since it is neither truly one nor truly many. If the mind isn't a true unity, what *is* it like? And if there are no ontologically independent foundations (whether material or immaterial) and metaphysical foundationalism is false, then what kind of structure *do* the mind and world conform to?

The positive account begins to take shape from Śrīgupta's influential threefold criterion for conventional reality (*saṃvṛtisatya*),[124] according to which whatever there is—whether material or immaterial—exists "conventionally," and whatever exists conventionally (i) is satisfactory only when not analyzed (*avicāraramaṇīya* or *avicāramanohara*), (ii) is interdependently originated (*pratītyasamutpanna*), and (iii) has the capacity for causal, or pragmatic, efficacy (*arthakriyāśakti* or *arthakriyāsāmarthya*).[125] In other words, (i) any object—

---

of nondual awareness, which may even prove helpful for more decisively determining Śrīgupta's relative chronology; see PVA *ad* kk. 3.197–207. For a discussion of this section of the PVA, see Inami (2011).

[124] Subsequent endorsements of this threefold criterion include, for instance, Jñānagarbha's SDV 8, 12, 21; Śāntarakṣita's MA 64; Kamalaśīla's MAP *ad* 64; Haribhadra's AAA (Wogihara 1932–1935, 594.18–25); the ca. eighth century Bhāviveka's MAS 9–11 and MRP I.4; and Atiśa's SDA 3.

[125] TA 11: *ma brtags gcig pu nyams dga' ste // de 'dra las byung de bzhin no // dngos po de dag de lta bu'i // don bya de dang de byed do //* (PD 3121, 105); "[1] Satisfactory only when not analyzed, [2] [things] arise from [causes] similar to themselves. [3] Those things enact their respective forms of causal efficacy." The TAV continues: *de lta bas na phyi rol dang nang na snang ba'i dngos po brtag pa'i spungs mi bzod pa rang dang mthun pa'i rgyus bskyed pa 'di dag ni gang las tha snyad 'dir 'gyur ba don bya ba ma brtags na nyams dga' ba nyid de dang der nye bar byed do //* (PD 3121, 105–6); "Thus, regarding these things that appear both externally and internally, which cannot withstand the pressure of analysis and which are produced from causes similar to themselves, based on which conventions (*vyavahāra) then come into being—if one has not examined their causal efficacy, one will approach satisfaction here and there." As Eckel (2008, 25) points out, Śrīgupta's TAV appears to be the earliest extant text in which we find this threefold characterization of conventional reality, with the first criterion as listed above possibly adapted from Candrakīrti (e.g., MAv 6.35), the second inherited from Nāgārjuna, and the third a repurposing of Dharmakīrti's criterion for

including the mind—satisfies our ordinary notions of independent existence and unity only when not subjected to analysis. That's because, upon inspection, (ii) each object turns out to be dependent for its very existence on some collection of parts, with each of those parts dependent upon its own parts, *ad indefinitum*. The picture, then, is one of aggregates—or conventional unities and beings—all the way down. Nevertheless (iii) anything that earns its keep as conventionally real must have causal efficacy, able to carry out its function in accordance with our pragmatic expectations: the water in my glass is conventionally real; the water of a mirage is not.

This account of aggregates all the way down is not an unfamiliar position when it comes to the material world, but what about the mind? If this picture applies to the mind too, and if the unity and being of any aggregate (such as a flock or an army) are conventions mentally designated in dependence on some parts (such as some individual sheep or soldiers), what could it mean to say that the unity and being of the *mind too* is mentally designated in dependence on its own parts? In what sense could the mind itself be mind-dependent without falling into a vicious regress or vicious circularity? Take, for instance, one moment of mind m_1 at t_1. Is the unity and being of m_1 at t_1 self-designated or is it designated by a subsequent moment of mind, m_2 at t_2? The self-designation alternative looks to be viciously circular. On the other hand, retrospective designation would seem to result in a vicious regress, compounded by the problem that the present moment of mind could

---

ultimately real particulars (*svalakṣaṇa*) in PV 3.3. On these three criteria, see also Eckel (1987, 137–38 n. 104).

never claim conventional unity and being *in the present* but must "wait in the wings" for the subsequent moment of mind to come along.

But lurking in this line of questioning is the presupposition of the kind of determinacy that belongs to the very foundationalist framework that Śrīgupta aims to reject. After all, on his view, we can't speak of *one* moment of mind prior to its designated unity and being *at all*. And just as there's no problem with acknowledging that there were some individual sheep prior to their designated unity and being *qua* flock, likewise, there's no problem with acknowledging that there was some mental stuff prior to its designation as <one moment of mind>. The claim is simply that, whatever we attend to earns the conventional unity and being that we experience it as having in the very moment and by virtue of its designation as such.

As I argue elsewhere, when Śrīgupta's neither-one-nor-many argument is taken together with these three criteria for conventional reality, the resulting metaphysical dependence structure is best characterized as a kind of metaphysical indefinitism.[126] On this picture, the mind and the world conform to a structure that admits of dependence chains that are indefinite (though not actually infinite) in length and dependence loops of greater than zero length. But much interpretive work remains to be done when it comes to filling out the details of this picture as it relates to the status and structure of the mind itself, and, unfortunately, Śrīgupta's terse text avails us little in this regard.

---

[126] See Chapter 1, "No Unity, No Problem: Madhyamaka Metaphysical Indefinitism."

There's one final respect in which Śrīgupta's argument against mental simples is just the beginning of the story: it presses on a host of philosophical questions with which all theorists of the mind have had to wrestle, raising many more questions than it answers. For instance, what kind of relation *should* we think that the mind and mental content share? If they are identical, what kind of identity relation do they enjoy, and if they are distinct, what kind of distinction should we draw between them?[127] And is it right to think of the mind and mental content as the kinds of things that could stand in any relation *at all*? What kind of ontological status does mental content enjoy relative to the mind? And what of the mind itself?[128]

But perhaps the most important question that Śrīgupta raises is: why *do* we so commonly take the unity of consciousness for granted, and are we justified in doing so? Specifically, does the mind's intentional structure undermine its simplicity, and if not, why not? The intuition that consciousness is simple is closely related to the foundationalist intuition, which supposes that everything—at the end of the day—is founded in some kind of fundamental entity, or entities, that constitute the bedrock of reality. When it comes to that intuition too, much more ink has been spilled describing *how* foundationalism might be true than arguing *that* it is true. Śrīgupta's neither-one-nor-many argument cautions that

---

[127] This question might be cast as inquiring into the relationship between the formal reality and the objective reality of thought for Descartes, as a prime example in the Western philosophical tradition. Indeed, it is on this very question that the well-known Arnauld-Malebranche debate on Cartesian ideas centered. For an overview of this debate, see Moreau (2000).

[128] One may look to Hume's *Treatise* I.iv.5–6 for a prime example in the Western philosophical tradition of an inquiry into this final line of questioning.

one ought not mistake either of these intuitions for brute facts that require no justification.

But exploring these many questions is a project for another day.

PART I BIBLIOGRAPHY

ABBREVIATIONS

AAA   *Abhisamayālaṃkārāloka* (Haribhadra). Wogihara (1932–1935).

AG   *Philosophical Essays.* Leibniz (1989).

AKB   *Abhidharmakośabhāṣya* (Vasubandhu). Pradhan (1967).

AKK   *Abhidharmakośa* (Vasubandhu). Pradhan (1967).

AS   *Abhidharmasamuccaya* (Asaṅga). Pradhan (1950).

AŚ   *antaraśloka* (transitional stanza)

AT   *Oeuvres de Descartes* (Descartes). Adam and Tannery (1964–1974).

BCAP   *Bodhicaryāvatārapañjikā* (Prajñākaramati). Vaidya (1960).

C   Bstan 'gyur Co ne

CŚ   *Catuḥśatakaśāstra* (Āryadeva). Lang (1986).

CSM   *The Philosophical Writings of Descartes*, vol. 2, translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch. (1984).

D   *sDe dge Tibetan Tripiṭaka – bsTan 'gyur Preserved at the Faculty of Letters, University of Tokyo*, edited by Z. Yamaguchi, et al. Tokyo: Tokyo University Press, 1977–1988.

G   Bstan 'gyur Gser bris ma. Tianjin: Tuanjin guji chubanshe, 1988.

GP   *Die Philosophische Schriften von Gottfried Wilhelm Leibniz*. Leibniz (1960). (Reference is to volume and page.)

JNĀ   *Jñānaśrīmitranibandhāvali* (Jñānaśrīmitra). Thakur (1959).

MA   *Madhyamakālaṃkāra* (Śāntarakṣita). Ichigō (1989).

MAP   *Madhyamakālaṃkārapañjikā* (Kamalaśīla). Ichigō (1985).

MAS   *Madhyamakārthasaṃgraha* (the eighth century Bhāviveka). PD 3084, vol. 58, 851–53.

MAU	*Madhyamakālaṃkāropadeśa* (Ratnākaraśānti). PD 3314, 604–25.
MAv	*Madhyamakāvatāra* (Candrakīrti). Chapter 6 in Li (2015).
MAV	*Madhyamakālaṃkāravṛtti* (Śāntarakṣita). Ichigō (1985).
MĀ	*Madhyamakāloka* (Kamalaśīla). PD 3116, vol. 62, 1114–1403; partial edition in Keira (2004).
MMK	*Mūlamadhyamakakārikā* (Nāgārjuna). Ye (2011).
MRP	*Madhyamakaratnapradīpa* (the eighth century Bhāviveka). PD 3081, vol. 57, 1487–1567.
MU	*Madhyamakopadeśa* (Atiśa). PD 3148, vol. 64, 283–86.
N	Bstan 'gyur Snar thang
P	*The Tibetan Tripiṭaka. Peking Edition − Kept in the Library of the Otani University, Kyoto − Reprinted Under the Supervision of the Otani University, Kyoto*, 168 vols., edited by D. T. Suzuki. Tokyo/Kyoto: Tibetan Tripitaka Research Institute, 1955–1961.
PD	Bstan 'gyur Dpe bsdur ma. Beijing: Krung go'i bod rig pa'i dpe skrun khang, 1994–2008. (Reference is to text number and page number.)
PM	*Leibniz: Philosophical Writings*. Leibniz (1973).
PP	*Prasannapadā* (Candrakīrti). Chapter 1 in MacDonald (2015).
PPU	*Prajñāparamitopadeśa* (Ratnākaraśānti). PD 3308, vol. 78, 357–434.
PSk	*Pañcaskandhaka* (Vasubandhu). Li and Steinkellner (2008).
PSkV	*Pañcaskandhakavibhāṣā* (Sthiramati). Kramer (2013).
PV	*Pramāṇavārttika* (Dharmakīrti). Miyasaka (1971/1972).
PVA	*Pramāṇavārttikālaṃkāra* (Prajñākaragupta). Sāṅkṛtyāyana (1953).
PVin	*Pramāṇaviniścaya* (Dharmakīrti). Chapters 1 and 2 in Steinkellner (2007), chapter 3 in Hugon and Tomabechi (2011).
RĀ	*Ratnāvalī* (Nāgārjuna). Hahn (1982).
ŚS	*Śūnyatāsaptati* (Nāgārjuna). Lindtner (1982, 34–69).

SDA	*Satyadvayāvatāra* (Atiśa). Lindtner (1981).
SDV	*Satyadvayavibhaṅga* (Jñānagarbha). Eckel (1987).
SN	*Sarvadharmaniḥsvabhāvasiddhi* (Kamalaśīla). PD 3118, vol. 62, 1479–1527.
SSŚ	*Sākārasiddhiśāstra* (Jñānaśrīmitra). Thakur (1959).
SVB	*Sugatamatavibhaṅgabhāṣya* (Jitāri). PD 3129, vol. 63, 887–1034.
TA	*Tattvāvatāra* (Śrīgupta). See TAV.
TAV	*Tattvāvatāravrtti* (Śrīgupta). PD 3121, vol. 63, 101–12; C 3892, *dbu ma, ha* 39b1–43a5; D 3892, *dbu ma, ha* 39b4–43b5; G 3295, *dbu ma, ha* 56a–62a; N 4064, *dbu ma, ha* 41a5–45b2; P 5292, *dbu ma, ha* 44b2–49a5.
TN	*Trisvabhāvanirdeśa* (Vasubandhu). La Vallée Poussin (1932–33).
Treatise	*A Treatise of Human Nature.* Hume (2000). (Reference is to book, part, and section)
TS	*Tattvasaṃgraha* (Śāntarakṣita). Shastri (1997), Bahirarthaparīkṣā Chapter in Saccone (2018, 155–221); PD 3497, vol. 107, 3–352.
VP	*Vaidālyaprakaraṇa* (Nāgārjuna). Tola and Dragonetti (1995).
Vś	*Viṃśikā* (Vasubandhu). Silk (2016).
VV	*Vigrahavyāvartinī* (Nāgārjuna). Bhattacharya (1978).
Works	*The Works of George Berkeley, Bishop of Cloyne.* Berkeley (1948–1957).
YŚ	*Yuktiṣaṣṭikākārikā* (Nāgārjuna). Lindtner (1982).

REFERENCES

Aczel, Peter. 1998. *Non-well-founded Sets.* Stanford: Stanford Center for the Study of Language and Information.

Adam, Charles and Paul Tannery, eds. 1964–1974. *Oeuvres de Descartes.* 2nd ed. Paris: Vrin.

Aitken, Allison. Forthcoming. *Introduction to Reality: A Tibetan Critical Edition, Annotated English Translation, and Philosophical Introduction to Śrīgupta's Tattvāvatāravrtti.* Cambridge, MA: Harvard Oriental Series.

Aitken, Allison, and Jeffrey K. McDonough. 2019. "Somethings and Nothings: Śrīgupta and Leibniz on Being and Unity." *Philosophy East and West*: doi: 10.1353/pew.0.0189.

Akahane, Ritsu. 2003. "離 一 多 性 を 証 因 と す る 無 自 性 論 証 」 と *avicāraikaramaṇīya* を め ぐ る 問題." ("*On the Proof of *Niḥsvabhāvatā* by Logical Reason, the Neither One nor Many Argument, and *Avicāraikaramaṇīya.*") *Indogaku bukkyōgaku kenkyū* (*Journal of Indian and Buddhist Studies*) 51: 124–28.

Almogi, Orna. 2009. *Rong-zom-pa's Discourses on Buddhology: A Study of Various Conceptions of Buddhahood in Indian Sources with Special Reference to the Controversy Surrounding the Existence of Gnosis (jñāna: ye shes) as Presented by the Eleventh-century Tibetan Scholar Rong-zom Chos-kyi-bzang-po.* Studia Philologica Buddhica. Monograph series 24. Tokyo: The International Institute for Buddhist Studies of the International College for Postgraduate Buddhist Studies.

———. 2010. "Māyopamādvayavāda versus Sarvadharmāpratiṣṭhānavāda: A Late Indian Subclassification of Madhyamaka and Its Reception in Tibet." *Journal of the International College of Postgraduate Buddhist Studies* 14: 135–212.

Arnold, Dan. 2008. "Buddhist Idealism, Epistemic and Otherwise: Thoughts on the Alternating Perspectives of Dharmakīrti." *Sophia* 47: 3–28.

———. 2010. "Nāgārjuna's 'Middle Way': A Non-eliminative Understanding of Selflessness." *Revue Internationale De Philosophie 253*, no. 3: 367–95.

Audi, Paul. 2012. "A Clarification and Defense of the Notion of Grounding." In *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Schnieder, 101–21. Cambridge: Cambridge University Press.

Barnes, Elizabeth. 2018. "Symmetrical Dependence." In *Reality and Its Structure: Essays in Fundamentality*, edited by Ricki Bliss and Graham Priest, 50–69. Oxford: Oxford University Press.

Barnett, David. 2008. "The Simplicity Intuition and Its Hidden Influence on Philosophy of Mind." *Noûs* 42, no. 2: 308–35.

Bayne, Tim. 2010. *The Unity of Consciousness*. Oxford: Oxford University Press.

Bayne, Tim, and David J. Chalmers. 2003. "What Is the Unity of Consciousness?" In *The Unity of Consciousness: Binding, Integration and Dissociation*, edited by Axel Cleeremans, 23–25. Oxford: Oxford University Press.

Berkeley, George. 1948–1957. *The Works of George Berkeley, Bishop of Cloyne*. Edited by A. A. Luce and T. E. Jessop. 9 vols. London: Thomas Nelson and Sons.

Berker, Selim. 2018. "The Unity of Grounding." *Mind* 127, no. 507: 729–77.

Bhattacharya, Kamaleswar, trans. 1978. *The Dialectical Method of Nāgārjuna: Vigrahavyāvartanī*. Edited by Elgin H. Johnston and Arnold Kunst. 2nd ed. Delhi: Motilal Banarsidass.

Bliss, Ricki. 2013. "Viciousness and the Structure of Reality." *Philosophical Studies* 166, no. 2: 399–418.

———. 2014. "Viciousness and Circles of Ground." *Metaphilosophy* 45, no. 2: 245–56.

———. 2015. "Metaphysical Grounding: Understanding the Structure of Reality." *Philosophical Review* 124, no. 3: 410–15.

Bliss, Ricki, and Graham Priest. 2018a. "The Geography of Fundamentality: An Overview." In *Reality and Its Structure: Essays in Fundamentality*, edited by Ricki Bliss and Graham Priest, 1–34. Oxford: Oxford University Press.

———. 2018b. "Metaphysical Dependence, East and West." In *Buddhist Philosophy: A Comparative Approach,* edited by Steven M. Emmanuel, 63–86. Hoboken, NJ: John Wiley & Sons.

Bliss, Ricki, and Kelly Trogdon. 2016. "Metaphysical Grounding." *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2016/entries/grounding/.

Blumenthal, James. 2004. *The Ornament of the Middle Way: A Study of the Madhyamaka Thought of Śāntarakṣita: Including Translations of Śāntarakṣita's Madhyamakālamkāra (The Ornament of the Middle Way) and Gyel-tsab's Dbu ma rgyan gyi brjed byang (Remembering "The Ornament of the Middle Way")*. Ithaca, NY: Snow Lion Publications.

Bohn, Einar Duenger. 2009. "Must There Be a Top Level?" *Philosophical Quarterly* 59, no. 235: 193–201.

———. 2018. "Indefinitely Descending Ground." In *Reality and Its Structure: Essays in Fundamentality*, edited by Ricki Bliss and Graham Priest, 167–81. Oxford: Oxford University Press.

Brook, Andrew, and Paul Raymont. 2017. "The Unity of Consciousness." *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), edited by Edward N. Zalta https://plato.stanford.edu/archives/sum2017/entries/consciousness-unity/.

Bu ston Rin chen grub. 1989. *Chos 'byung gsung rab rin po che'i gter mdzod*. Dehradun: Drikung Partrun Khang.

Buescher, Hartmut. 2010. "Vasubandhu's *Pañcaskandhaka*." *Indo-Iranian Journal* 53, no. 4: 331–58.

Cabezón, José Ignacio. 1994. *Buddhism and Language: A Study of Indo-Tibetan Scholasticism*. Albany: State University of New York Press.

Cameron, Ross P. 2014. "Parts Generate the Whole, But They Are Not Identical to It." In *Composition as Identity*, edited by Aaron J. Cotnoir and Donald L. M. Baxter, 90–107. Oxford: Oxford University Press.

Correia, Fabrice, and Benjamin Schnieder. 2012. "Grounding: An Opinionated Introduction." In *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Schnieder, 1–36. Cambridge: Cambridge University Press.

Cowherds, eds. 2011. *Moonshadows: Conventional Truth in Buddhist Philosophy*. New York: Oxford University Press.

Dainton, Barry. 2005. *Stream of Consciousness: Unity and Continuity in Conscious Experience*. 2nd ed. London: Routledge.

Davidson, Donald. 1980. "Paradoxes of Irrationality." In *Philosophical Essays on Freud*, edited by Richard Wollheim and James Hopkins, 289–305. Cambridge: Cambridge University Press.

Dennett, Daniel C. 1991. *Consciousness Explained*. New York: Little, Brown.

———. 1992. "The Self as Centre of Narrative Gravity." In *Self and Consciousness: Multiple Perspectives*, edited by Frank S. Kessel, Pamela M. Cole, and Dale L. Johnson, 103–15. Hillsdale, NJ: Erlbaum Publishers.

Descartes, René. 1984. *Meditations on First Philosophy*. In *The Philosophical Writings of Descartes*, translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch, vol. 2. New York: Cambridge University Press.

Doctor, Thomas, 2004. *Speech of Delight: Mipham's Commentary on Śāntarakṣita's Ornament of the Middle Way*. Ithaca, NY: Snow Lion Publications.

Donnelly, Maureen. 2011. "Using Mereological Principles to Support Metaphysics." *Philosophical Quarterly* 61: 225–46.

Dreyfus, Georges, and Sara McClintock. 2003. *The Svātantrika-Prāsaṅgika Distinction: What Difference Does a Difference Make?* Boston: Wisdom Publications.

Duckworth, Douglas S., Malcolm David Eckel, Jay L. Garfield, John Powers, Yeshes Thabkhas, and Sonam Thakchöe. 2016. *Dignāga's Investigation of the Percept: A Philosophical Legacy in India and Tibet*. New York: Oxford University Press.

Eckel, Malcolm David. 1987. *Jñānagarbha's Commentary on the Distinction Between the Two Truths: An Eighth Century Handbook of Madhyamaka Philosophy*. Albany: State University of New York Press.

———. 2008. *Bhāviveka and His Buddhist Opponents: Chapters 4 and 5 of the Verses on the Heart of the Middle Way* (*Madhyamakahṛdayakārikāḥ*) *with the Commentary Entitled The Flame of Reason* (*Tarkajvālā*). Harvard Oriental Series 70. Cambridge, MA: Harvard University Department of South Asian Studies.

Ejima, Yasunori. 1980. *Chūgan shisō no tenkai—Bhāvavikeka Kenkyū.* Tokyo: Shunjōsha.

Eltschinger, Vincent. 2007. "On 7th and 8th Century Buddhist Accounts of Human Action, Practical Rationality and Soteriology." In *Pramāṇakīrtiḥ: Papers Dedicated to Ernst Steinkellner on the Occasion of His 70th Birthday, Part 1*, edited by Birgit Kellner, Horst Lasic, M. T. Much, and Helmut Tauscher, 135–62. Wien: Arbeitskreis für Tibetische und Buddhistische Studien Universität Wien.

———. 2014. *Buddhist Epistemology as Apologetics—Studies on the History, Self-understanding and Dogmatic Foundations of Late Indian Buddhist Philosophy*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

Ferraro, Guiseppe. 2017. "Realistic-Antimetaphysical Reading vs. Any Nihilistic Interpretation of Madhyamaka." *Journal of Indian Philosophy* 45, no. 1: 73–98.

Fine, Kit. 2010. "Some Puzzles of Ground." *Notre Dame Journal of Formal Logic* 51: 97–118.

———. 2012. "A Guide to Ground." In *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Schnieder, 37–80. Cambridge: Cambridge University Press.

French, Steven. 2014. *The Structure of the World: Metaphysics and Representation*. New York: Oxford University Press.

Garfield, Jay L. 1995. *The Fundamental Wisdom of the Middle Way: Nāgārjuna's Mūlamadhyamakakārikā.* New York: Oxford University Press.

———. 2002. *Empty Words: Buddhist Philosophy and Cross-Cultural Interpretation*. New York: Oxford University Press.

———. 2019a. *The Concealed Influence of Custom: Hume's* Treatise *from the Inside Out.* New York: Oxford University Press.

———. 2019b. "Givenness and Primal Confusion." In *Wilfrid Sellars and Buddhist Philosophy: Freedom from Foundations*, edited by Jay L. Garfield, 113–29. New York: Routledge.

Garfield, Jay L., and Graham Priest. 2003. "Nāgārjuna and the Limits of Thought." *Philosophy East and West* 53, no. 1: 1–21.

Garfield, Jay L., and Ngawang Samten. 2006. *Ocean of Reasoning: A Great Commentary on Nāgārjuna's Mūlamadhyamakakārikā*. Oxford, New York: Oxford University Press.

Goodman, Charles. 2016. "From Madhyamaka to Consequentialism: A Road Map." In *Moonpaths: Ethics and Emptiness*, edited by the Cowherds, 141–58. Oxford: Oxford University Press.

Gold, Jonathan C. 2014. *Paving the Great Way: Vasubandhu's Unifying Buddhist Philosophy*. New York: Columbia University Press.

'Gos lo tsā ba Gzhon nu dpal. 1971. *Deb gter sngon po*. Śata-Piṭaka Series vol. 212, edited by Lokesh Chandra. New Delhi: The International Academy of Indian Culture.

Hahn, Michael, ed. 1982. *Nāgārjuna's Ratnāvalī*. Vol. 1. Bonn: Indica et Tibetica.

Hayes, Richard. 1994. "Nāgārjuna's Appeal." *Journal of Indian Philosophy* 22: 299–378.

Hugon, Pascale. 2015. "Proving Emptiness: The Epistemological Background for the 'Neither One Nor Many' Argument and the Nature of Its Probandum in Phya Pa Chos Kyi Seng Ge's Works." *Journal of Buddhist Philosophy* 1, no. 1: 58–94.

Hugon, Pascale, and Toru Tomabechi. 2011. *Dharmakīrti's Pramāṇaviniścaya: Chapter 3*. Sanskrit Texts from the Tibetan Autonomous Region 8. Beijing, Vienna: China Tibetology Publishing House and Austrian Academy of Sciences Press.

Hume, David. *A Treatise of Human Nature*. Edited by David Fate Norton and Mary J. Norton. New York: Oxford University Press, 2000.

Huntington Jr., C. W. 1983. "A 'Nonreferential' View of Language and Conceptual Thought in the Work of Tsoṅ-kha-pa." *Philosophy East and West* 33, no. 4: 325–39.

Ichigō, Masamichi, ed. 1985. *Madhyamakālaṃkāra of Śāntarakṣita with His Own Commentary or Vṛtti and with the Subcommentary or Pañjikā of Kamalaśīla*. Kyoto: Buneido.

———. 1989. "Śāntarakṣita's *Madhyamakālaṃkāra*." In *Studies in the Literature of the Great Vehicle: Three Mahāyāna Buddhist Texts*, edited by Luis O. Goméz and Jonathan A. Silk, 141–240. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.

Inada, Kenneth K. 1970. *Nāgārjuna: A Translation of His* Mūlamadhyamakakārikā *with an Introductory Essay*. Tokyo: Hokoseido Press.

Inami, Masahiro. 2011. "Nondual Cognition." In *Religion and Logic in Buddhist Philosophical Analysis*, *Proceedings of the Fourth International Dharmakīrti Conference*, *Vienna August 23–27, 2005*, edited by Helmut Krasser, Horst Lasic, Eli Franco, and Birgit Kellner, 177–96. Vienna: Verlag der Österrichischen Akademie der Wissenschaften.

'Jam mgon 'Ju Mi pham rgya mtsho. [undated xylograph]. *Dbu ma rgyan gyi rnam bshad 'jam dbyangs bla ma dgyes pa'i zhal lung*. In *Gsung 'bum Mi pham rgya mtsho*, vol. 12, 11–426. Rdzong gsar dgon pa'i par khang.

Jenkins, Carrie S. 2011. "Is Metaphysical Dependence Irreflexive?" *The Monist* 94, no. 2: 267–76.

Johansson, Ingvar. 2015. "Applied Mereology." *Metascience*, 24: 239–45.

Johnston, Mark. 2005. "Constitution." In *The Oxford Handbook of Contemporary Philosophy*, edited by Frank Jackson and Michael Smith, 636–75. Oxford: Oxford University Press.

Kajiyama, Yuichi. 1989. "Controversy between the Sākāra- and Nirākāra-vādins of the Yogācāra School-Some Materials." In *Y. Kajiyama, Studies in Buddhist Philosophy (Selected Papers)*, edited by Katsumi Mimaki et al., 389–400. Kyoto: Rinsen Book Co. Originally published 1965.

Kalupahana, David J. 1986. *Nāgārjuna: The Philosophy of the Middle Way*. Albany: State University of New York Press.

Kant, Immanuel. 1999. *Critique of Pure Reason*. Edited and translated by Paul Guyer and Allen W. Wood. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press.

———. 2002. *Theoretical Philosophy after 1781*. Edited by Henry Allison and Peter Heath; translated by Gary Hatfield, Michael Friedman, Henry Allison, and Peter Heath. Cambridge: Cambridge University Press.

Kapstein, Matthew. 2001. *Reason's Traces*. Somerville, MA: Wisdom Publications.

———. 2018. "Who Wrote the *Trisvabhāvanirdeśa*? Reflections on an Enigmatic Text and Its Place in the History of Buddhist Philosophy." *Journal of Indian Philosophy* 46: 1–30.

Keira, Ryusei. 2004. *Mādhyamika and Epistemology: A Study of Kamalaśīla's Method for Proving the Voidness of All Dharmas: Introduction, Annotated Translations and Tibetan Texts of Selected Sections of the Second Chapter of the Madhyamakāloka*. Wiener Studien zur Tibetologie und Buddhismuskunde; Heft 59. Wien: Arbeitskreis für Tibetische und Buddhistische Studien Universität Wien.

Kellner, Birgit. 2010. "Self-Awareness (svasaṃvedana) in Dignāga's 'Pramāṇasamuccaya' and '-vṛtti': A Close Reading." *Journal of Indian Philosophy* 38, no. 3: 203–31.

———. 2017a. "Proofs of Idealism in Buddhist Epistemology: Dharmakīrti's Refutation of External Objects." In *Indian Epistemology and Metaphysics*, edited by Jörg Tuske, 103–28. London: Bloomsbury.

———. 2017b. "Proving Idealism: Dharmakīrti." In *The Oxford Handbook of Indian Philosophy*, edited by Jonardon Ganeri, 307–26. New York: Oxford University Press.

Kellner, Birgit, and Sara L. McClintock, eds. 2014. ākāra *in Buddhist Philosophical and Soteriological Analysis*. Special issue, *Journal of Indian Philosophy* 42, nos. 2–3.

Kellner, Birgit, and John Taber. 2014. "Studies in Yogācāra-Vijñānavāda Idealism I: The Interpretation of Vasubandhu's *Viṃśikā*." *Asiatische Studien/Etudes Asiatiques* 68, no. 3: 709–56.

Kobayashi, Mamoru. 1992. "*Tattvāvatāra* of Śrīgupta: Japanese Annotated Translation I." *Ronshū* 19: 37–56.

———. 1994. "*Tattvāvatāra* of Śrīgupta: Japanese Annotated Translation II." *Mikkyō Bunka* 85: 99–80.

Kramer, Jowita. 2013. *Sthiramati's Pañcaskandhakavibhāṣā, Part 1: Critical Edition*. Sanskrit Texts from the Tibetan Autonomous Region, vol. 16. Beijing, Vienna: China Tibetology Research Center and Austrian Academy of Sciences.

Krämer, Stephan. 2013. "A Simpler Puzzle of Ground." *Thought* 2: 85–89.

Kyuma, Taiken. 2010. "Jñānaśrīmitra on the Definition of Existence." In *From Vasubandhu to Caitanya: Studies in Indian Philosophy and Its Textual History*, edited by Johannes Bronkhorst and Karin Preisendanz, 121–36. Delhi: Motilal Banarsidass.

La Vallée Poussin, Louis de, ed. 1903–1913. *Mūlamadhyamakakārikās de Nāgārjuna avec la Prasannapadā*. Bibliotheca Buddhica iv, St. Pétersbourg, Bibliotheca Buddhica IV. (Addenda by de Jong, 25–59 and 217–52.)

———, ed. 1907–1912. *Madhyamakāvatāra par Candrakīrti: Traduction Tibétaine*. Bibliotheca Buddhica 9. St. Petersburg: L'Académie Impériale des Sciences.

———, ed. 1932–1933. "Le Petit Traite de Vasubandhu-Nagarjuna sur les Trois Natures." *Melanges Chinois et Bouddhiques* 2: 147–61.

Ladyman, James, and Don Ross, with David Spurrett and John Collier. 2007. *Every Thing Must Go: Metaphysics Naturalized*. New York: Oxford University Press.

Lang, Karen. 1986. *Āryadeva's* Catuḥśataka*: On the Bodhisattva's Cultivation of Merit and Knowledge*. Indiske Studier VII. Copenhagen: Akademisk Forlag.

Leech, Jessica. 2016. "X—The Mereology of Representation." *Proceedings of the Aristotelian Society* 116, no. 2: 205–28.

Leibniz, G. W. 1960. *Die Philosophische Schriften von Gottfried Wilhelm Leibniz*. Edited by C. I. Gerhardt. Berlin: Weidmann; reprinted Hildesheim: Olms.

———. 1973. *Leibniz: Philosophical Writings*. Translated by Mary Morris and edited by G. H. R. Parkinson. London: J. M. Dent & Sons.

———. 1989. *Philosophical Essays*. Edited and translated by Roger Ariew and Daniel Garber. Indianapolis: Hackett.

Lennon, Thomas, and Robert Stainton, eds. 2008. *The Achilles of Rationalist Psychology*. Dordrecht: Springer.

Li Xuezhu. 2015. "*Madhyamakāvatāra-kārikā* Chapter 6." *Journal of Indian Philosophy* 43, no. 1: 1–30.

Li Xuezhu and Ernst Steinkellner, eds. 2008. *Vasubandhu's Pañcaskandhaka*. Sanskrit Texts from the Tibetan Autonomous Region, vol. 4. Beijing, Vienna: China Tibetology Research Center and Austrian Academy of Sciences.

Lindtner, Christian. 1981. "Atiśa's *Introduction to the Two Truths*, and Its Sources." *Journal of Indian Philosophy* 9: 161–214.

———. 1982. *Nagarjuniana: Studies in the Writings and Philosophy of Nāgārjuna*. Indiske Studier IV. Copenhagen: Akademisk Forlag.

Lockwood, Michael. 1989. *Mind, Brain and the Quantum*. Oxford: Blackwell Publishers.

———. 1994. "Issues of Unity and Objectivity." In *Objectivity, Simulation, and the Unity of Consciousness*, edited by Christopher Peacocke, 89–95. Oxford: Oxford University Press.

MacDonald, Anne. 1988. *Blo gsal grub mtha'*. Master's thesis, University of British Columbia.

———, ed. and trans. 2015. *In Clear Words: The Prasannapadā, Chapter One*. 2 vols. Verlag der Österreichischen Akademie der Wissenschaften.

MacKenzie, Matthew. 2008. "Ontological Deflationism in Madhyamaka." *Contemporary Buddhism* 9, no. 2: 197–207.

Markosian, Ned. 1998. "Simples." *Australasian Journal of Philosophy* 76, no. 2: 213–28.

———. 2004a. "Simples, Stuff, and Simple People." *The Monist* 87, no. 3: 405–28.

———. 2004b. "SoC It to Me? Reply to McDaniel on MaxCon Simples." *Australasian Journal of Philosophy* 82, no. 2: 332–40.

Mathes, Klaus-Dieter, ed. 2003. *'Gos Lo tsā ba gZhon nu dpal's Commentary on the Ratnagotravibhāga-vyākhyā*. Nepal Research Centre Publications, no. 24. Stuttgart: Franz Steiner Verlag.

Matilal, B. K. 1970. "Reference and Existence in Nyāya and Buddhist Logic." *Journal of Indian Philosophy* 1: 83–110.

Matsumoto, Shiro. 1978. "Jñāngarbha の 二 諦 説" (Theory of Two Truths of Jñānagarbha). *Bukkyō-Gaku* 5: 109–37.

McClintock, Sara. 2010. *Omniscience and the Rhetoric of Reason: Śāntarakṣita and Kamalaśīla on Rationality, Argumentation, and Religious Authority*. Boston: Wisdom Publications.

———. 2013. "Kamalaśīla and Śāntarakṣita on Scripture and Reason—The Limits and Extent of 'Practical Rationality' in the Tattvasaṃgraha and Pañjikā." In *Scriptural Authority, Reason, and Action: Proceedings of a Panel at the 14th World Sanskrit Conference, Kyoto, September 1st–5th 2009*, edited by Vincent Eltschinger and Helmut Krasser, 209–38. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

McDaniel, Kris. 2003. "Against MaxCon Simples." *Australasian Journal of Philosophy* 81: 265–75.

———. 2007. "Extended Simples." *Philosophical Studies* 133, no. 1: 131–41.

Miyasaka, Yūsho. 1971/1972. "Pramāṇavārttika-Kārikā (Sanskrit and Tibetan)." *Acta Indologica* 2: 1–206.

Moreau, Denis. 2000. "The Malebranche-Arnauld Debate." In *The Cambridge Companion to Malebranche*, edited by Steven Nadler, 112–38. Cambridge: Cambridge University Press.

Morganti, Matteo. 2014. "Metaphysical Infinitism and the Regress of Being." *Metaphilosophy* 45, no. 2: 232–44.

———. 2015. "Dependence, Justification and Explanation: Must Reality Be Well-Founded?" *Erkenntnis* 80, no. 3: 555–72.

———. 2018. "From Ontic Structural Realism to Metaphysical Coherentism." *European Journal for Philosophy of Science* 9, no. 1: 1–20.

Moriyama, Seitetsu. 1984. "Kamalaśīla's and Haribhadra's Refutation of the Satyākāra and Alikākāra-vādins of the Yogacara School." *Journal of Indian and Buddhist Studies* (*Indogaku Bukkyogaku Kenkyu*) 33, no. 1: 393–89.

Moriyama, Shinya. 2011. "On Ratnākaraśānti's theory of cognition with false mental images (*alīkākāravijñānavāda*) and the neither-one-nor-many argument." The XVIth Congress of the International Association of Buddhist Studies. New Taipei City: Dharma Drum Buddhist College.

———. 2014. "Ratnākaraśānti's Theory of Cognition with False Mental Images (**alīkākāravāda*) and the Neither-One-Nor-Many Argument." *Journal of Indian Philosophy* 42, nos. 2–3: 339–51.

Murti, T. R. V. 1960. *The Central Philosophy of Buddhism, a Study of the Mādhyamika System.* London: George Allen and Unwin.

Nagel, Thomas. 1971. "Brain Bisection and the Unity of Consciousness." *Synthese* 22, no. 3: 396–413.

O'Brien, Gerard, and Jon Opie. 1998. "The Disunity of Consciousness." *Australasian Journal of Philosophy* 76, no. 3: 378–95.

Oetke, Claus. 1992. "Doctrine and Argument in Vijñānavāda-Buddhism." *Wiener Zeitschrift Für Die Kunde Südasiens / Vienna Journal of South Asian Studies* 36: 217–25.

Phya pa Chos kyi seng ge. 2006a. *Dbu ma rgyan gyi 'grel pa.* In *Bka' gdams gsung 'bum*, vol. 6, 433–518. Chengdu: Si khron mi rigs dpe skrun khang.

———. 2006b. *Dbu ma snang ba'i 'grel pa.* In *Bka' gdams gsung 'bum phyogs bsgrigs thengs dang po*, vol. 6, 266–428. Chengdu: Si khron mi rigs dpe skrun khang.

Pradhan, Pralhad, ed. 1950. *Abhidharma Samuccaya of Asaṅga*. Santiniketan: Visva-Bharati.

———. 1967. *Abhidharmakosabhasya of Vasubandhu*. 2nd ed. Patna: K. P. Jayaswal Research Center. Reprinted 1975.

Priest, Graham. 2009. "The Structure of Emptiness." *Philosophy East and West* 59, no. 4: 467–80.

———. 2014. *One: Being an Investigation into the Unity of Reality and of Its Parts, Including the Singular Object Which Is Nothingness*. Oxford: Oxford University Press.

———. 2018. "Buddhist Dependence." In *Reality and Its Structure: Essays in Fundamentality*, edited by Ricki Bliss and Graham Priest, 1–34. Oxford: Oxford University Press.

Putnam, Hilary. 1967. "The Nature of Mental States." In *Readings in Philosophy of Psychology 1980*, edited by Ned Block, 223–31. Cambridge: Harvard University Press.

Raven, Michael. 2003. "Unity of Consciousness and the Self." *Proceedings of the Aristotelian Society* 103: 325–52.

———. 2015. "Ground." *Philosophy Compass* 10, no. 5: 322–33.

Roerich, George N. 1949/1953. *The Blue Annals*. Calcutta: Royal Asiatic Society of Bengal.

Rosen, Gideon. 2010. "Metaphysical Dependence: Grounding and Reduction." In *Modality: Metaphysics, Logic, and Epistemology*, edited by Bob Hale and Aviv Hoffmann, 109–35. Oxford: Oxford University Press.

Rosenthal, David M. 1986. "Two Concepts of Consciousness." *Philosophical Studies* 49, no. 3: 329–59.

———. 2002. "Persons, Minds, and Consciousness." In *The Philosophy of Marjorie Grene*, edited by Randall E. Auxier and Lewis Edwin Hahn (Library of Living Philosophers, Vol. 29), La Salle, Illinois: Open Court.

———. 2003. "Unity of Consciousness and the Self." *Proceedings of the Aristotelian Society* 103: 325–352.

———. 2009. "On What Grounds What." In *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David Chalmers, David Manley, and Ryan Wasserman, 347–83. Oxford: Oxford University Press.

———. 2016. "Grounding on the Image of Causation." *Philosophical Studies* 173: 49–100.

Saccone, Margherita Serena. 2018. *On the Nature of Things: A Buddhist Debate on Cognitions and Their Objects, with a Manuscript Description by Paolo Giunta*. Vienna: Arbeitskreis für Tibetische und Buddhistische Studien Universität Wien.

Salvini, Mattia. 2011. "*Upādāyaprajñaptiḥ* and the Meaning of Absolutives: Grammar and Syntax in the Interpretation of Madhyamaka." *Journal of Indian Philosophy* 39: 229–44.

Sāṅkṛtyāyana, Rāhula, ed. 1953. *Pramāṇavārtika-bhāshyam or Vārtikālaṅkāraḥ of Prajñākaragupta: Being a Commentary on Dharmakīrti's Pramāṇavārtikam.* Tibetan Sanskrit Works Series 1. Patna: K. P. Jayaswal Research Institute.

Schaffer, Jonathan. 2003. "Is There a Fundamental Level?" *Noûs* 37: 498–517.

_____. 2009. "On What Grounds What." In *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David John Chalmers, David Manley, and Ryan Wasserman, 347–83. Oxford University Press.

_____. 2012. "Grounding, Transitivity, and Contrastivity." In *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Schnieder, 122–38. Cambridge: Cambridge University Press.

Searle, John R. 2002. "Consciousness." In *Consciousness and Language*, 36–60. Cambridge: Cambridge University Press. Originally published in 2000, *Annual Review of Neuroscience* 23: 557–78.

Seyfort Ruegg, David. 1981. *The Literature of the Madhyamaka School of Philosophy in India.* Wiesbaden: Otto Harrassowitz Verlag.

Shastri, Swami Dwarikadas. 1981. *Tattvasaṃgrahapañjikā. Tattvasaṃgraha of Śāntarakṣita, with the Commentary "Pañjikā" of Śrī Kamalaśīla.* Series I, 2 vols. Varanasi: Bauddha Bhāratī.

Shirasaki, Kenjo. 1984. "The Sugatamatavibhaṅgabhāṣya of Jitāri (I)." *Kōbe Joshi Daigaku Kiyō: Bungaku Bunen* 17, no. 1: 77–107.

Shoemaker, Sydney S. 1996. *The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press.

———. 2003. "Consciousness and Co-Consciousness." In *The Unity of Consciousness: Binding, Integration and Dissociation*, edited by Axel Cleeremans, 59–71. Oxford: Oxford University Press.

Siderits, Mark. 1988. "Nāgārjuna as Anti-Realist." *Journal of Indian Philosophy* 16.4: 311–25.

———. 1989. "Thinking on Empty: Madhyamaka Anti-Realism and Canons of Rationality." In *Rationality in Question: On Eastern and Western Views of Rationality*, edited by Shlomo Biderman and Ben-Ami Scharfstein, 231–49. Leiden: Brill.

———. 2003. *Personal Identity and Buddhist Philosophy: Empty Persons.* Burlington, VT: Ashgate.

———. 2011. "Is Everything Connected to Everything Else? What the Gopīs Know." In *Moonshadows: Conventional Truth in Buddhist Philosophy*, edited by the Cowherds. New York: Oxford University Press.

———. 2016. *Studies in Buddhist Philosophy*. Edited by Jan Westerhoff and Christopher V. Jones. Oxford, New York: Oxford University Press.

Siderits, Mark, and Shōryū Katsura. 2013. *Nāgārjuna's Middle Way: Mūlamadhyamakakārikā.* Somerville, MA: Wisdom Publications.

Silk, Jonathan. 2016. *Materials towards the Study of Vasubandhu's Viṃśikā (I): Sanskrit and Tibetan Critical Editions of the Verses and Autocommentary, an English Translation and Annotations.* Harvard Oriental Series 81. Cambridge, MA: Department of South Asian Studies, Harvard University.

Spackman, John. 2014. "Between Nihilism and Anti-Essentialism: A Conceptualist Interpretation of Nāgārjuna." *Philosophy East and West* 64, no. 1: 151–73.

Sprung, Mervyn. 1977. "Non-cognitive Language in Mādhyamika Buddhism." In *Buddhist Thought and Asian Civilization: Essays in Honor of Herbert V. Guenther on his Sixtieth Birthday*, edited by Leslie S. Kawamura and K. Scott, 241–53. Emeryville, CA: Dharma Publishing.

Steinkellner, Ernst, ed. 2007. *Dharmakīrti's Pramāṇaviniścaya, Chapters 1 and 2.* Beijing, Vienna: China Tibetology Publishing House and Austrian Academy of Sciences Press.

Streng, Frederick. 1967. *Emptiness: A Study in Religious Meaning.* Nashville: Abdingdon Press.

Tahko, Tuomas E. 2014. "Boring Infinite Descent." *Metaphilosophy* 45, no. 2: 257–69.

Tahko, Tuomas E., and Jonathan E. Lowe. 2016. "Ontological Dependence." *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2016/entries/dependence-ontological/.

Tauscher, Helmut, ed. 1999. *Phya pa Chos kyi seṅ ge, Dbu ma śar gsum gyi stoṅ thun.* Wiener Studien zur Tibetologie und Buddhismuskunde; Heft 43. Wien: Arbeitskreis für Tibetische und Buddhistische Studien, Universitat Wien.

Thakur, Anantalal, ed. 1959. *Jñānaśrīmitranibandhāvali (Buddhist Philosophical Works of Jñānaśrīmitra).* Tibetan Sanskrit Works Series V. Patna: Kashi Prasad Jayaswal Research Institute.

Thompson, Naomi. 2018. "Metaphysical Interdependence, Epistemic Coherentism, and Holistic Explanation." In *Reality and Its Structure: Essays in Fundamentality*, edited by Ricki Bliss and Graham Priest, 107–25. Oxford: Oxford University Press.

Tillemans, Tom. T. F. 1982. "The 'Neither One Nor Many' Argument for *śūnyatā* and its Tibetan Interpretations: Background Information and Source Materials." *Études de Lettres* (University of Lausanne) 3: 103–28.

———. 1983. "The 'Neither One Nor Many' Argument for *śūnyatā* and Its Tibetan Interpretations." In *Contributions on Tibetan and Buddhist Religion and Philosophy*, edited by Ernst Steinkellner and Helmut Tauscher, 305–20. Vienna: Arbeitskreis für Tibetische und Buddhistische Studien.

———. 1984. "Two Tibetan Texts on the 'Neither One Nor Many' Argument for *Śūnyatā*." *Journal of Indian Philosophy* 12: 357–88.

———. 2011. "How Far Can a Mādhyamika Reform Conventional Truth? Dismal Relativism, Fictionalism, Easy-Easy Truth, and the Alternatives." In *Moonshadows: Conventional Truth in Buddhist Philosophy*, edited by the Cowherds. New York: Oxford University Press.

———. 2016. *How Do Mādhyamikas Think? And Other Essays on the Buddhist Philosophy of the Middle.* Somerville: Wisdom Publications.

Tillemans, Tom, and Donald S. Lopez. 1998. "What Can One Reasonably Say About Nonexistence? A Tibetan Work on the Problem of Āśrayāsiddha." *Journal of Indian Philosophy* 26, no. 2: 99–129.

Tola, Fernando, and Carmen Dragonetti. 1995. *Nāgārjuna's Refutation of Logic (Nyāya): Vaidalyaprakaraṇa / Źib mo rnam par ḥthag pa źes bya baḥi rab tu byed pa*. Delhi: Motilal Banarsidass.

Tomlinson, David. 2019. "Buddhahood and Philosophy of Mind: Ratnākaraśānti, Jñānaśrīmitra, and the Debate over Mental Content (*ākāra*)." PhD diss., University of Chicago.

Trogdon, Kelly. 2013. "An Introduction to Grounding." In *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-Dependence*, edited by Miguel Hoeltje, Benjamin Schnieder, Alex Steinberg, 97–122. Munich: Philosophia Verlag.

———. 2018. "Inheritance Arguments for Fundamentality." In *Reality and Its Structure: Essays in Fundamentality*, edited by Ricki Bliss and Graham Priest, 182–98. Oxford: Oxford University Press.

Tsong kha pa Blo bzang grags pa. [undated xylograph]. *Dbu ma rgyan gyi brjed byang*. In *Gsung 'bum Tsong kha pa (zhol)*, vol. 16, 543–64. Zhol par khang.

———. [undated xylograph]. *Dbu ma rgyan gyi zin bris*. In *Gsung 'bum Tsong kha pa (zhol)*, vol. 15, 783–814. Zhol par khang.

Tye, Michael. 2003. *Consciousness and Persons: Unity and Identity*. Cambridge, MA: MIT Press.

Unger, Peter. 1990. *Identity, Consciousness, and Value*. New York: Oxford University Press.

Vaidya, P. L., ed. 1960. *Bodhicaryāvatāra of Sāntideva with the commentary Pañjikā of Prajñākaramati*. Buddhist Sanskrit Texts 12. Darbhanga: Mithila Institute.

Varzi, Achille. 2010. "On the Boundary between Material and Formal Ontology." In *Interdisciplinary Ontology, Vol. 3: Proceedings of the Third Interdisciplinary Ontology Meeting*, edited by Barry Smith, Riichiro Mizoguchi, and Sumio Nakagawa, 3–8. Tokyo: Keio University.

———. 2019. "Mereology." *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/spr2019/entries/mereology/.

Walser, Joseph. 2005. *Nāgārjuna in Context: Mahāyāna Buddhism and Early Indian Culture*. New York: Columbia University Press.

Westerhoff, Jan. 2009. *Nāgārjuna's Madhyamaka: A Philosophical Introduction.* Oxford: Oxford University Press.

———. 2011. "The Merely Conventional Existence of the World." In Cowherds, *Moonshadows: Conventional Truth in Buddhist Philosophy.* New York: Oxford University Press, 189–212.

———. 2016. "On the Nihilist Interpretation of Madhyamaka." *Journal of Indian Philosophy* 44, no. 2: 337–76.

———. 2017a. "Madhyamaka and Modern Western Philosophy: A Report." *Buddhist Studies Review* 33, nos. 1–2: 281–302.

———. 2017b. "Nāgārjuna on Emptiness: A Comprehensive Critique of Foundationalism." In *The Oxford Handbook of Indian Philosophy*, edited by Jonardon Ganeri, 93–119. New York: Oxford University Press.

Wilson, Jessica. 2014. "No Work for a Theory of Grounding." *Inquiry* 57: 535–79.

Wogihara, Unrai, ed. 1932–1935. *Abhisamayālaṃkārāloka Prajñāpāramitāvyākhyā* (*Commentary on Aṣṭasāhasrikā-Prajñāpāramitā*) *by Haribhadra Together with the Text Commented On*. Tokyo: The Toyo Bunko.

Wood, Thomas E. 1994. *Nāgārjunian Disputations: A Philosophical Journey Through an Indian Looking-Glass*. Monographs of the Society for Asian and Comparative Philosophy, no. 11. Honolulu: University of Hawaii Press.

Ye Shaoyong, ed. 2011. *Zhong lun song: Fan Zang Han he jiao, dao du, yi zhu*. Shanghai: Zhongxi shuju.

PART II

LOCKE ON RELATIONS AND PERSONS

The self is a relation that relates itself to itself.

~ Kierkegaard
The Sickness unto Death

3

LOCKE'S RELATIONAL ACCOUNT OF PERSONS

Locke's account of persons was groundbreaking for decoupling personal identity from substantial identity. But what—metaphysically speaking—*is* a person on Locke's view? Locke sets out a threefold taxonomy of complex ideas, from which interpreters commonly derive a threefold ontology of substances, modes, and relations. If he is to have a metaphysical account of what a person is, then, it will have to fall in one of these three categories. Given these options, it has been commonly supposed that a person is most plausibly a substance or a mode.[129] In what follows, I'll argue for the neglected alternative: a Lockean person is a relation.[130]

There is, in fact, some striking *prima facie* evidence for this reading: Locke presents his account of persons in a chapter on relations.[131] Indeed, the subsequent chapter is

---

[129] Proponents of substance-interpretations include Alston and Bennett (1988), Chappell (1990), Winkler (1991), Gordon-Roth (2015), and Rickless (2015), while proponents of mode-interpretations include Law (1769), Mattern (1980), Uzgalis (1990), Thiel (2011), and LoLordo (2012).

[130] While this project was under way, Simendic (2015) published an article making a case that Locke's person is a relation. Although there is some overlap, my interpretation differs significantly from his, and the points of agreement and disagreement are noted throughout. Simendic's account does not attend to the unique structure of Locke's ideas of relations or their peculiar metaphysical status, both of which are addressed in this paper. While Simendic's relation-interpretation is a start in the right direction, this paper develops a more detailed account of what it means for a person to be a relation, both as an idea and from an ontological perspective. Moreover, Simendic does not distinguish between the Lockean self and the Lockean person, regarding them as first- and third-person equivalents, and this distinction is central to the interpretation I put forward.

[131] John Locke, *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch (Oxford: Oxford University Press, 1975), II.xxvii, "Of Identity and Diversity." References to the *Essay* are by book, chapter, and section.

entitled "Of other relations." It sounds odd to us today to say that a person is a relation. But Locke does not mean the same thing by "relation" that we do. Once equipped with a proper understanding of Lockean relations, we will see why a person is best categorized not as a substance (like a body or a soul) or as a mode (like a number or an activity), but as a relation (like a father or a friend). Yes, for Locke, fathers themselves are *relations*, not relational *things*.[132] So too are friends, foes, professors, possessors, and, I argue, persons.

Locke couches his discussion of ontology in terms of how we form *ideas* of substances, modes, and relations. I therefore begin with an analysis of the unique structure of Lockean *ideas* of relations and explain how the idea of a person conforms to that structure. I also shed light on the historically overlooked distinction between the Lockean self and the Lockean person. Next, I will tease out two metaphysical features of Lockean relations and show how they are instantiated in the case of persons. Finally, I will reveal how these two metaphysical features support Locke's forensic intuitions about personhood while also defusing a number of classic objections to Locke's account of personal identity. Along the way, I will also explain how a relation-interpretation of Lockean persons may lay claim to the principal advantages of both the substance- and mode-interpretations.

1. The Idea of Person as an Idea of a Relation

Before exploring what the idea of a person as an idea of a relation might look like, it is first necessary to sketch out the unique structure of Lockean ideas of relations. Locke's

---

[132] While fatherhood is also a kind of abstracted relation, it is subsidiary to the relation father.

account of ideas of relations does not align with contemporary intuitions. The first example he offers of an idea of a relation in his introductory chapter on the topic is the idea of husband. What is striking about so many of Locke's stock examples of ideas of relations—like <husband>, <general>, <king>, <subject>, <possessor>, etc.—is that they seem not to be ideas of *relations* at all, but of *things* that stand in relations to other things. Locke claims, however, that an idea of a relation is any idea that is not simply an idea "of Things, as they are in themselves," but that "intimate[s] some other" thing or idea (II.xxv.1). He offers the example of Cajus, which, when considered as a <man>, is an idea of a substance, but when considered as a <husband>, intimates some other individual—namely, Cajus's wife—and thus counts as an idea of a relation. He explains that an idea of a relation is one that "is not confined to that precise Object: It can carry an *Idea*, as it were, beyond it self, or, at least, look beyond it, to see how it stands in conformity to any other" (II.xxv.1).

Locke observes that ideas of relations are a far more pervasive part of our mental lives than may be initially supposed. He takes names of relations that have correlative terms, such as "bigger" and "less," "father" and "son," or "cause" and "effect," to signify "obvious" examples of ideas of relations, but he insists that there are a great many ideas of relations that are often improperly understood because they lack a correlative term, for example, "concubine" (II.xxv.2).[133] Locke observes that language often misleads us when it comes to relations, for many absolute terms "conceal a tacit, less observable relation" (e.g.,

---

[133] As Locke states elsewhere, "relation is commonly over-look'd, v.g. A *Patron* and *Client,* are easily allowed to be Relations: but a Constable or Dictator, are not so readily, at first hearing, considered as such. Because there is no peculiar Name for those who are under the Command of a Dictator, or Constable, expressing a Relation to either of them" (II.xxviii.3).

"imperfect," "great," etc.), suggesting that more of our ideas than may be readily apparent have direction, pointing secondarily to something beyond the thing that they primarily denominate (II.xxv.3).

Locke further defines an idea of a relation as a complex idea "which consists in the consideration and comparing one *Idea* with another" (II.xxv.7).[134] He states that ideas of relations presuppose "two Ideas, or Things, either in themselves really separate, or considered as distinct, and then a ground or occasion for their comparison" (II.xxv.6). Examples of grounds for comparison include "Extent, Degrees, Time, Place, or any other Circumstances" (II.xi.4). Ideas of relations, then, involve considering one idea or thing, *x,* vis-à-vis some other idea(s) or thing(s), *y,* in respect of some ground for comparison, G (xGy).[135] This class of ideas includes familiar ideas of relations, like <bigger>, as well as what might seem to the non-Lockean as relational things, such as <mother>. Locke classifies ideas like <bigger> as "proportional relations" (II.xxviii.1), while ideas like <mother> belong to the category of "natural relations" (II.xxviii.2).[136] Still, all ideas of relations—regardless of the subclass to which they belong—conform to the same structure.

---

[134] Locke frequently uses the language of (ideas of) relations "consisting in consideration/comparing." However, since ideas are not the same as mental acts for Locke, we might understand that he means here "ideas of relations *result from* considering and comparing one idea with another," as has been suggested by Rickless (2018, 77). I suggest that relations *themselves* consist in or involve in part the mental act of comparing/consideration (more on that below).

[135] Notice that the ground for comparison, G, stands in for the relation, R, as in standard formalizations. This is because, on Locke's view, the idea of the relation R is signified *collectively* by <xGy>, i.e., the consideration of x vis-à-vis y in respect of some ground for comparison, G; the relation is *not* something bridging x and y. For instance, the relation <father> consists in the consideration of some man x vis-à-vis some child y in respect of generation G. This should be contrasted with "standard" accounts according to which some father x stands in the relation of fatherhood R in respect of some child, y, xRy.

[136] In addition to ideas of proportional and natural relations, Locke also lists ideas of instituted, or voluntary, relations such as <citizen> (II.xxviii.3), and ideas of moral relations such as <good> and

The formation of the idea of the relation <bigger>, for instance, involves considering (the idea of) some object x vis-à-vis (the idea of) some object y in respect of <size> (G). Similarly, the formation of the idea of the relation <mother> involves considering (the idea of) a woman x vis-à-vis (the idea of) a child y in respect of <generation> (G).[137] Notice that in all of these examples, the idea of the relation denominates more than one relatum, but one relatum is primary ($\boldsymbol{x}$Gy, where the primary relatum is in the first place and bolded). In the case of <bigger>, the object regarded as greater in size compared with some other object or some comparison class is the primary relatum. In the case of <mother>, the woman considered as the parent in respect of some child is the primary relatum. Lockean ideas of relations, then, are ways of considering some idea/thing (the primary relatum) vis-à-vis some other idea/thing or set of ideas/things (the secondary relatum/relata) in respect of some ground for comparison. The relata may be simple ideas or (complex ideas of) substances, modes, or further relations.

Now, one might think: maybe it's the ground for comparison that is actually the relation. But that can't be right. For the examples that Locke gives of grounds for comparison generally look to be modes, for instance, size, extent, degree, time, place, etc. And in the case of the relations under which we consider men, the ground for comparison is often a mode of activity, e.g., generation (in the case of a parent), teaching (in the case of a professor), wielding power (in the case of a king), etc. To state the obvious: if grounds of

---

<evil>, which represent the agreement or disagreement of our voluntary actions with a rule (II.xxviii.4), as additional classes of relations apart from those whose ground for comparison is time, place, and causality.

[137] See II.xxviii.18 for Locke's analysis of the idea of the relation <father> in such terms.

comparison are (generally) modes, and modes are distinct from relations, then grounds of comparison must also be distinct from relations.

There are two ideas of relations that will be important for understanding persons as relations: identity and agency. As for the first, Locke regards the ideas of identity and diversity to be two of the most fundamental kinds of ideas of relations, also characterizing them respectively as the agreement and disagreement of ideas.[138] The formation of the simplest cases of synchronic identity involve considering one idea vis-à-vis itself in respect of sameness, e.g., considering the idea of a circle vis-à-vis the idea of a circle in respect of sameness. While the *synchronic* agreement (S) between <circle> and <circle> (*a*S*a*) is an immediately apparent object of intuitive knowledge, Locke recognizes that cases of *diachronic* identity are less apparent.[139] He says,

> Another occasion, the mind often takes of comparing, is the very Being of things, when, considering any thing as existing at any determin'd time and place, we compare it with it self existing at another time, and thereon form the *Ideas* of [diachronic] *Identity* and *Diversity*. (II.xxvii.1)[140]

As Locke explains, perceiving an idea of the relation of diachronic identity requires us to compare something with itself at a different time and place. As we will see, the idea of

---

[138] Propositions as the agreement and disagreement of ideas is actually the proper object of all varieties of knowledge for Locke, suggesting that in some respect, all objects of knowledge are relations. Ideas of relations are, in fact, a ubiquitous part of our mental lives, for, as Locke remarks, all of our ideas "when attentively considered" include some kind of relation (II.xxi.3).

[139] There are, of course, more complex cases of synchronic identity relations, as in the case of mathematical equivalences that require demonstration to determine. Since these cases do not bear directly on the present discussion, I will leave them aside.

[140] In presenting his fourfold account of knowledge, Locke clarifies that although he treats them as separate categories of knowledge, or ways in which the mind may perceive the agreement or disagreement of ideas, "Identity and Co-existence are truly nothing but Relations" (IV.i.7).

<person> involves an idea of a relation of diachronic sameness of self. Note that it is not two persons that are being related in this diachronic identity relation—the relata are *selves*. It will be important to distinguish between selves and persons going forward. Not incidentally, the fact that <person> involves a diachronic identity relation explains why Locke presents his account of persons in the chapter "Of Identity and Diversity."

When we understand the diachronic identity of anything—whether it be <same atom>, <same man>, or <same self>—the mental act of comparing takes up the ideas of $<x_1$ at $t_1>$ and $<x_2$ at $t_2>$ in respect of some ground for comparison, namely, an identity condition. We thereby arrive at the relational idea <same *x*.> The difference between ideas such as <man> and <atom> and the idea <person>, I argue, is that <person> is *only ever* an idea of a relation, while <man> or <atom>, when considered synchronically, are ideas of substances. <Same man> is an idea of a relation (of identity), while <man> is an idea of a single substance.[141] Yet, I suggest, both <same self> and <person> are ideas of relations, and indeed the idea of <person> is constituted in part by the idea of <same self>. This is because Locke uses the term "person" in a technical, forensic sense. Whether literally a legal matter or a question of moral accountability, forensics is necessarily pastward looking, so <person> is always an idea of a relation of diachronic identity, uniting a present self with a past self, who is the agent of some action.

---

[141] Locke classifies the idea of man as an idea of a single substance, and defines this "ordinary *Idea* of *a Man*" as "a combination of the *Ideas* of a certain sort of Figure, with the powers of Motion, Thought, and Reasoning, joined to Substance" (II.xii.6). He speaks about the relational idea of "same man" differently, describing its identity condition as "a participation in the same continued Life, by constantly fleeting Particles of Matter, in succession vitally united to the same organized Body" (II.xxvii.6).

An <agent>, much like <professor> or <possessor>, is another relational idea under which we can consider a man, and this denomination is also important for understanding <person>. Although Locke does not explicitly classify <agent> in his taxonomy of ideas, he uses the word "agent" interchangeably with "cause," and he explicitly classifies <cause> as an idea of a relation (II.xxvi.2). Moreover, as noted, Locke takes sets of correlative ideas (including <cause> and <effect>) to be obvious examples of ideas of relations (II.xxv.2), and <agent> and <patient> is one such pair of mutually implicative correlatives.[142] The relation <agent> is, in fact, jointly constituted by a number of relations, much like the relation <grandfather>.[143] An agent is something that engages in an action that affects some patient by virtue of wielding some power. So if the primary relatum of an <agent> is some man *x*, then <agent> is a way of considering some man (1) vis-à-vis some patient in respect of activity, (2) vis-à-vis some action in respect of generation,[144] (3) vis-à-vis some power in respect of possession.[145] It is the second of these agential relations that will be most relevant for understanding persons. Of course, Locke often speaks of a man as an "agent," but this no more makes agents substances than

---

[142] <Agent> may also implicate other ideas in addition to <patient>, such as <action>, <object>, <instrument>, or <power>.

[143] The man considered under the relation <grandfather> necessarily stands in relations not only vis-à-vis a grandchild but also vis-à-vis his own child who is the parent of the grandchild in addition to the woman who bore his child, etc.

[144] When Locke equates "agent" with "cause," he cites "generation" as the ground of comparison.

[145] This is much like the relation <possessor>, which can be a way of considering a man vis-à-vis some other thing in respect of ownership. This threefold unpacking of <agent> is not intended to be exhaustive. As noted above, agent may also involve a relation with some instrument or with some indirect object that is distinct from the patient, etc.

denominating a man by the relation "father" makes fathers substances.[146] And just as <paternity> is an abstraction of the relation <father>, signifying the state or the fact of being a father, similarly, <agency> is an abstraction of the relation <agent>, signifying the state or the fact of being an agent. In order to carry out its forensic role, the relation <person> involves relations of identity as well as relations of agency.

The fact that the Lockean idea of person is a forensic notion is critical for understanding how <person> conforms to the structure of Locke's ideas of relations. <Person> carries out the role of linking a present subject of judgment with a past agent for the purposes of tracking a morally or legally responsible party. Locke states that "person" is:

> A Forensick Term appropriating Actions and their Merit; and so belongs only to intelligent Agents capable of a Law, and Happiness and Misery. This personality extends it self beyond present Existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it self past Actions, just upon the same ground, and for the same reason that it does the present. (II.xxvii.26)

It is as persons that we appropriate actions and their merits in both the temporal domain of legal accountability and the ultimate domain of resurrection and final moral judgment.

Despite the fact that in ordinary language we tend to use "man" or "human being" and "person" interchangeably, it is well known that Locke distinguishes man from person. Person is not coextensive with man because they have different diachronic identity conditions: sameness of consciousness for person and participation in same continued life

---

[146] Similarly, Locke states that "powers belong only to agents, and are attributes only of substances" (II.xxi.16), but this is just to say that only substances can be the primary relatum of the relation <agent> considered vis-à-vis power in respect of possession.

for man. Locke draws this distinction between person and man based on the intuition that it is unjust to assign reward or punishment for a deed that one is unaware of having committed (II.xxvii.26). Consider a presently existent individual *S* in respect of some past deed *x*. Past deed *x* belongs to *S* the *man* iff *x* was undertaken by a man who is the *same* as *S*, that is, a man who participated in the same continued life as *S*. Yet at present, *S* is only able to consciously appropriate a limited subset of the actions he has committed throughout his lifetime.[147] And while a long-forgotten action undertaken in his infancy is rightly said to belong to *S* the *man*, it does *not* belong to *S* the *person,* who is presently incapable of consciously appropriating it. While it seems natural for us to say that a man is responsible for his deeds, it is important for Locke that we assess persons, not men, for moral responsibility.

Sameness of man is not only insufficient to link a prior agent with the appropriate present recipient of judgment—it is not even necessary. Locke famously offers a variety of thought experiments to illustrate this point, such as that of the prince and the cobbler, in which the soul together with the consciousness of a prince comes to inhabit the body of a cobbler (II.xxvii.15). Locke observes that while the cobbler is the same man as before, he is

---

[147] See, for instance, Locke's account of the sleeping and waking Socrates as two distinct persons (II.xxvii.19). Although he acknowledges that we might pick out a man and a person by the same name, like "Socrates," this does not commit him to a relative identity theory. See Chappell (1989) for a compelling case against such a reading of Locke. Socrates the man and Socrates the person pick out different referents, both synchronically and diachronically. Likewise, the aggregate of atoms with which Socrates the man is co-located has different synchronic and diachronic identity conditions from the man. There is not one *thing* that is both (i) a mass of atoms and (ii) a rational, thinking being joined to a human body. Likewise, person as a relation and man as a substance have different synchronic and diachronic identity conditions, so there is not one thing that is strictly both a person and a man in a moment or over time. Cf. Stuart (2013b) for a more recent argument in favor of the relative identity reading.

now certainly a different person. And although the cobbler and prince are different *men*, the past prince and the present cobbler sharing the same consciousness constitute one *person*. Thus, just as sameness of man does not necessarily track sameness of person, difference of man does not provide sufficient grounds for differentiating a past agent from a present recipient of judgment.

Person is no more coextensive with any material or immaterial substance than it is with man. Locke considers it conceivable (and thus possible) that consciousness may not always be annexed to the same substance, potentially changing bodies and even souls. Since neither sameness of man nor sameness of any substance can guarantee that reward or punishment is justly meted out, only persons can unite the present and past selves partaking in the same consciousness, regardless of the substance or substances in which that consciousness inheres.

Some interpreters read Locke's person more broadly than the forensic role calls for, suggesting that the forensic role is just one aspect of a person, who can also be, say, a friend or an artist.[148] But this appears to be influenced by our ordinary way of conceiving the term independent of the technical sense that Locke has in mind. Locke is clear that these other roles are ways of considering a *man*—not a person. He says,

> there is *no one thing*, whether simple *Idea*, Substance, Mode, or Relation, or Name of either of them, *which is not capable of almost an infinite number of* Considerations, in reference to other things: and therefore this makes no small part of Men's Thoughts and Words. *v.g.* One single **Man** may at once be concerned in, and sustain all these following *Relations*, and many more, *viz.* Father, Brother, Son, Grandfather,

---

[148] For instance, in arguing for a substance-interpretation of Lockean persons, Gordon-Roth states, "Although Locke identifies persons as moral agents when he calls 'person' a forensic term, I do not think that persons are merely moral agents. Persons are also artists, friends, etc. Thus, being a moral agent is just one aspect of being a person (albeit a very important one)" (2015, 109 n. 36).

> Grandson, Father-in-Law, Son-in-Law, Husband, **Friend**, Enemy, Subject, General, Judge, Patron, Client, Professor, European, English-man, Islander, Servant, Master, Possessor, Captain, Superior, Inferior, Bigger, Less, Older, Younger, Contemporary, Like, Unlike, etc. (II.xxv.7; bold added)

Here, Locke provides a lengthy list of examples of relations that consist in considering one *man* as the primary relatum vis-à-vis some secondary relatum in respect of some ground. Much like <friend>, <patron>, and <professor>, <person> is just another relation under which we might consider a *man*. But unlike these other relations, <person> involves a relation of identity. And just as there is no puzzle in understanding why <professor> and the woman who is the substance denominated as the primary relatum enjoy different diachronic identity conditions (for she remains the same substance before, after, and throughout any gaps in holding a professorship), the same holds true for <person> and the man denominated as the primary relatum.[149]

By seeing how the idea of a person conforms to the structure of Locke's ideas of relations, we can see how a person might have distinct diachronic identity conditions from a collective substance, such as a mass of atoms, and a single substance, such as a body, a soul, or even a man, and still be co-located with them. First, if the idea of a person is an idea of a relation, what are a person's relata and what is its ground for comparison? And how is the idea of person *qua* idea of a relation (hereafter <person$_R$>) to be differentiated from the ideas of personal identity or sameness of self as ideas of a relation? Locke states that "person" stands for:

---

[149] While "man" is the technical term Locke uses to refer to the human being *qua* organism, I use "woman," "man," and "human being" interchangeably in this sense.

a thinking, intelligent Being, that has reason and reflection, and can consider **it self as it self,** the **same** thinking thing **in different times and places**; which it does only by that consciousness, which is inseparable from thinking . . . (II.xxvii.9; bold added)

A person, then, is foremost something that "can consider *it self as it self* . . . in different times and places." Thus, <person$_R$> involves a *diachronic* identity relation of a present self vis-à-vis some past self in respect of sameness of consciousness. In order to carry out its forensic role, however, <person$_R$> must also involve a relation of agency. As Locke tells us, "person" is "a Forensick Term **appropriating Actions** and their Merit; and so belongs only to *intelligent **Agents***" who, by virtue of consciousness, "**becomes concerned and accountable, owns and imputes to it self past Actions**" (II.xxvii.26; bold added). A person is, then, necessarily a diachronic agent, that is, an appropriator of a past action. <Person$_R$> not only involves [Relation 1] a Diachronic Identity Relation in which the present self identifies with some past self, but also [Relation 2] a Diachronic Agential Relation whereby the present self "owns" and "imputes to itself" the action(s) of that past self.

We can represent the structure of <person$_R$> as involving these two constitutive relations, [R1] a Diachronic Identity Relation and [R2] a Diachronic Agential Relation, with the following diagram:

Figure 7: The Idea of a Person as an Idea of a Relation
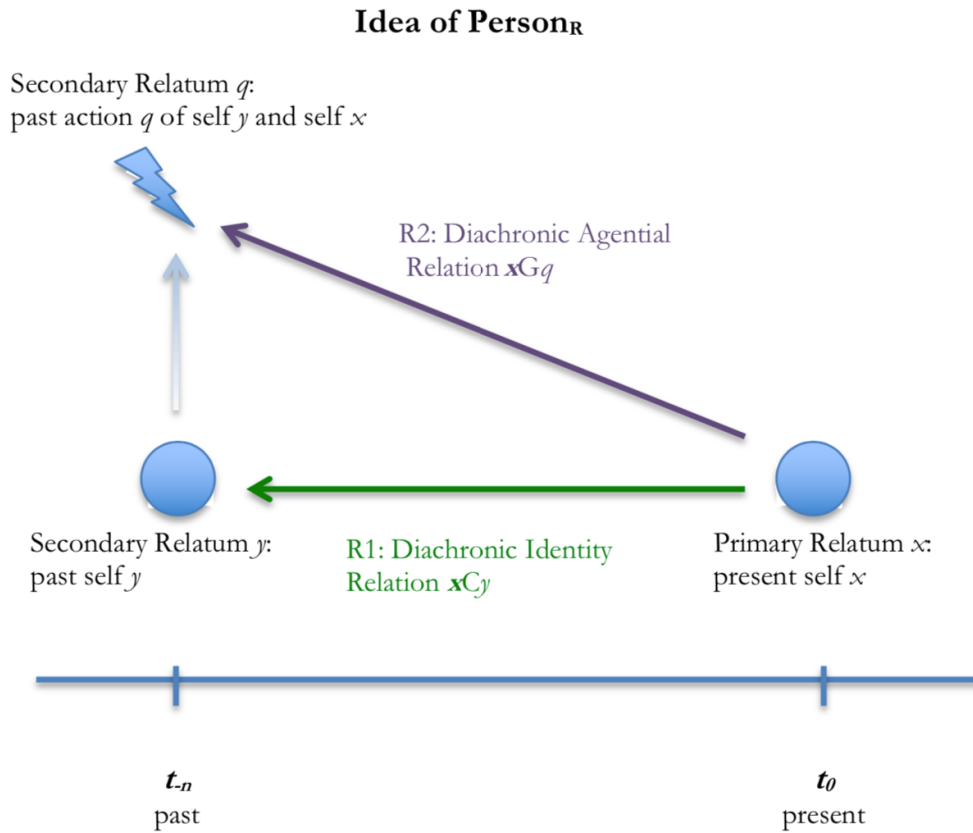
R1, Diachronic Identity Relation: *<person_R> involves the consideration of the present self* x *vis-à-vis (at least one) past self* y *in respect of sameness of consciousness,* C. The Diachronic Identity Relation is signified by xCy, where C stands for the ground for comparison, viz. sameness of consciousness, *x* stands for the present self as its primary relatum, and *y* stands for a past self as its secondary relatum. Because forensics is necessarily pastward looking, <person_R> always involves an idea of a Diachronic Identity Relation, uniting the present self with at least one past self.

R2, Diachronic Agential Relation: *<person_R> involves the consideration of the present self* x *vis-à-vis (at least one) past action* q *of (at least one) past self* y *in respect of generation,* G. The

Diachronic Agential Relation is signified by $\boldsymbol{x}Gq$, where G stands for the ground for comparison, viz. generation, x stands for the present self, and q stands for a past action (of self y). The relation of central concern to forensics is that of agency born by the *present* self in respect of some *past* action.[150] R1, the Diachronic Identity Relation of <same self>, in a sense, underwrites and makes possible R2, the Diachronic Agential Relation, the present self's conscious appropriation of some past action, which is necessary for the Lockean person to underwrite moral accountability.

2. The Idea of Person$_R$ and the Idea of Self$_R$

As we have seen, <person$_R$> is a way of considering a present self in respect of a past self and its action. The fact that the past self and present self are relata of the <person$_R$> indicates a critical and historically overlooked distinction between Locke's treatment of "self" and "person."[151] To be sure, Locke frequently uses the terms "self" and "person" interchangeably, and they are commonly interpreted as first- and third-person

---

[150] Cf. Simendic's relation-interpretation, wherein he maintains that person *qua* relation is an idea "comprised of two principal elements—substance and personal identity" (2015, 80). Simendic regards personal identity as "the relation between multiple diachronic iterations of the same consciousness" (ibid.). He takes a person to be a complex idea involving some kind of relation between the *substance*, which is taken to be a person by substance-interpretation proponents, and the *mode* of the continuation of consciousness, taken to be a person by mode-interpretation proponents. Simendic's interpretation is, in effect, a hybrid of the substance- and mode-interpretations.

[151] Lähteenmäki also distinguishes the Lockean self from the Lockean person, arguing that "a person is the sensible self taken together with those thoughts and actions the sensible self attributes to itself as its own by being conscious of them" (2018, 165). Although Lähteenmäki articulates the distinction between self and person differently from how it is presented here, he does not commit to a position on the ontological status of persons or elaborate on how precisely the person results from taking the sensible self together with the consciously appropriated actions, and so his account is not necessarily incompatible with the relation-interpretation.

equivalents.[152] They are not, however, strictly equivalent. For Locke, the person carries out the forensic role in his system of linking a present subject of judgment with a past agent of some action. It is the "person"—not the "self"—that is a diachronic forensic entity tracking moral accountability (II.xxvii.26), and it is the self—not the person—that is the synchronic object of knowledge of the cogito (IV.iii.21). A <person> is *necessarily* diachronic, whereas a <self> *need* not exist for more than a moment (though it *can* and most often *does*).

<Self> falls into the category of obvious Lockean ideas of relations inasmuch as it is a member of a correlative pair, namely, <self> and <other>.[153] This is true of both the generic self and the personal self. The generic self may pick out the synchronic identity of anything, whether talking about "the circle itself," "the cat himself," or "the woman herself." Each of these "selves" signifies a synchronic identity relation of *x*'s being the self-same *x*. Moreover, <*x*-self> also implicates the relation of being distinct from what is other than *x*.

The idea of a *personal* self, which we most commonly associate with the term "self," has more than one unique feature not shared with the generic self. First, a personal self is necessarily conscious, and reflexively so. My being a personal self requires that I consciously identify myself *as* myself in a synchronic identity relation. Second, a personal self also involves an agential relation. That's because a personal self necessarily knows itself to be the agent of its own thoughts and actions.

---

[152] Simendic, for instance, takes "self" and "person" to be first- and third-person equivalents (2015, 81). Simendic does not accept any synchronic relations as constituents of persons, since he does not appear to acknowledge the existence of a synchronic self.

[153] See II.xxv.2.

It is the personal self, both present and past, that are among the relata of the ⟨person$_R$⟩. (In the analysis that follows, I will use "self" to refer strictly to the personal self.) Locke explains the self as follows:

> When we see, hear, smell, taste, feel, meditate, or will any thing, we know that we do so. Thus it is always as to our **present** Sensations and Perceptions: And by this every one is **to himself**, that which he calls *self*. (II.xxvii.9; bold added)

Locke emphasizes that I am my *self* to my *self* in the present moment by virtue of the fact that I engage in thinking and know that I do so; that is, I am conscious of myself *as* a thinker. Thinking is the action of an "intellectual Agent," which occurs by exercising some power (II.xxii.11). As Locke has it, the soul does not always think, and thus thinking is the action and not the essence of the soul (II.xix.4), but the self, on the other hand, necessarily thinks (II.xxvii.17).[154] And since thinking is invariably accompanied by consciousness (II.i.11), whenever we think, "we know that we do so." Accordingly, when it comes to the kind of mental action most central to moral accountability—volition—whenever we "will any thing, we know that we do so." That is, whenever I will to do anything, I am conscious that I am the agent of that willing (and likewise the agent of whatever action follows from that willing). Thus, we might take "self" to be synonymous with "self-conscious agent."

The picture of the self as a self-conscious agent is somewhat complicated by the fact that it wields not only active but also passive power. On Locke's view, thinking occurs by virtue of the self-conscious agent wielding either an active power like will, as in the case of

---

[154] Thus, while a soul and a thinking thing (whatever kind of substance it may be) might persist uninterruptedly, a self (and thus, the potential relata for a person$_R$) exists only and always whenever thinking occurs.

volition, or a passive power like understanding, as in the case of perception (II.vi.2).[155]

Locke notes that, although in both cases mental operations are "called and counted *Actions*,

yet, if nearly considered, will not be found to be always perfectly so" (II.xxi.72). Mental

operations like perception that are a function of passive powers may, Locke suggests, more

properly be classified as passions. And although we do indeed regard the perceiver as an

intellectual agent by virtue of the fact that it wields a passive power, there is some "external

Agent" that wields the active power in such cases (II.xxi.72). Yet even when the self *seems*

to wield only a passive power, as a subject or patient of passions, sensations of pleasure and

pain, etc., the self nonetheless possesses active power as a capacity and is in that respect

properly called and counted as an "agent," just like someone need not always be teaching

to be properly called and counted as a "teacher."

The peculiar mental action of consciousness further complicates the picture of the

self *qua* self-conscious agent. Consciousness, whereby we know that we perceive (or will or

see, etc.), presumably results from the passive power of understanding. Yet both the agent

and patient of consciousness can be nowhere *but* internal to the subject. There looks to be,

then, an important sense in which consciousness is *active* insofar as being conscious of our

thoughts and actions, our pleasures and pains, *as* our own is something that we do *to*

ourselves, something that has its source *in* the thinking thing, just like memory, imagination,

and reflection. We might say, then, that the self is always some kind of agent, in addition

---

[155] Locke (confusingly) sometimes identifies thinking with perception (e.g., II.vi.2), but more often uses "thinking" as the umbrella term for all modes/activities of the mind (e.g., II.xix.2). He identifies two kinds of actions, which he takes to be exhaustive: motion and thinking, with motion being the action of bodies and thinking the action of minds, or souls. Thinking, thus, denominates all mental activities, including perception, volition, remembering, reflecting, etc. (II.xxi.72; II.xix.4).

to being a patient. Indeed, it is imperative that the present self as the primary relatum of the person$_R$ be in some sense both an agent *and* a patient. This is because the present self must appropriate responsibility for some past action by identifying as its agent, and yet also be a suitable subject, or patient of judgment and reward or punishment.[156]

In sum, the self is not *simply* a thinking thing, i.e., it is not simply a substance considered in and of itself. Rather, it is a thinking thing minimally considered under two relations, with a thinking thing as the primary relatum of both:

1. Synchronic Identity Relation: <self$_R$> involves the consideration of a thinking thing x vis-à-vis itself in respect of sameness, S (xSx).

2. Synchronic Agential Relation: <self$_R$> involves the consideration of a thinking thing x vis-à-vis some present thought or action z in respect of generation, G (xGz).[157]

The self as a twofold relation is represented in the following diagram:

---

[156] It is this feature of the personal self as the subject, particularly of pleasure and pain, that enables concernment for the future self as well. Locke emphasizes the centrality of the personal self's role as a subject of sensations when he says, "if we take wholly away all Consciousness of our Actions and Sensations, especially of Pleasure and Pain, and the concernment that accompanies it, it will be hard to know wherein to place personal Identity" (II.i.11). See Lähteenmäki (2018) for an analysis of the Lockean self as a sensible subject.

[157] We might understand the agency relation to obtain vis-à-vis a thought or volitional action, as well as vis-à-vis itself as the patient by virtue of the reflexive activity of consciousness.

Figure 8: The Idea of a Self as an Idea of a Relation

Idea of Self$_R$

Secondary Relatum z: present thought/action z
of thinking thing x

Synchronic Agential
Relation $\boldsymbol{x}G\boldsymbol{z}$

Synchronic Identity
Relation $\boldsymbol{x}S\boldsymbol{x}$

Primary Relatum x:
present thinking thing x

$\boldsymbol{t_0}$
present
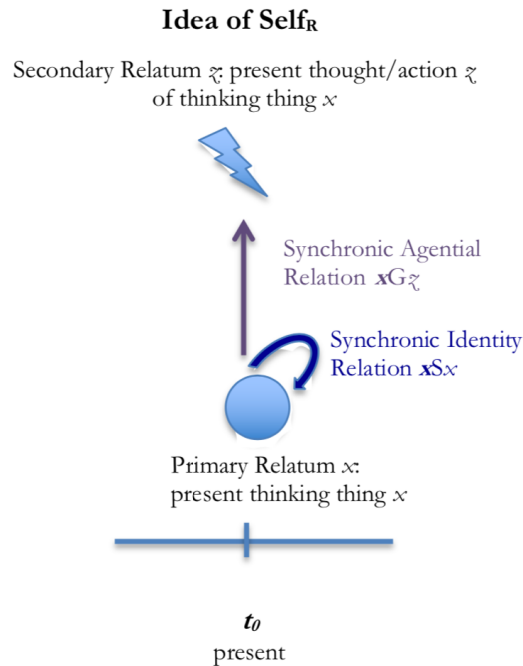
Locke reiterates that the self involves both a Synchronic Identity Relation and a Synchronic Agential Relation when he states:

> For it is by the consciousness it has of its **present** Thoughts and Actions that it is *self* **to it** *self* **now**, and so will be the **same** *self* as far as the same consciousness can extend to Actions past or to come. (II.xxvii.10; bold added)[158]

In other words, (1) we are our *self* to our *self* in the present moment, which is made possible by the reflexivity of consciousness that necessarily accompanies thinking, viz. the Synchronic Identity Relation. And (2) we are necessarily conscious of ourselves as the agent of our occurrent thoughts and (intentional) physical actions, viz. the Synchronic Agential

---

[158] Similarly, "*Self* is that conscious thinking thing, (whatever Substance, made up or whether Spiritual, or Material, Simple or Compounded, it matters not) which is sensible, or conscious of Pleasure and Pain, capable of Happiness or Misery, and so is concern'd for it *self*, as far as that consciousness extends" (II.xxvii.17).

147

Relation. Here, Locke suggests that the intuitive knowledge of the Synchronic Identity Relation that is partly constitutive of the self is made possible by knowledge of the Synchronic Agential Relation, "For it is by consciousness it has of its present Thoughts and Actions, that it is *self* to it *self*." Yet it is not that consciousness of the Synchronic Agential Relation is *prior* to the Synchronic Identity Relation. Rather, since consciousness is (i) necessarily reflexive and (ii) invariably co-occurs with all acts of thinking,[159] Locke takes intuitive knowledge of the Synchronic Identity Relation and the Synchronic Agential Relation to naturally and necessarily co-occur.

In virtue of the consciousness that attends my *present* thoughts and actions, I am my *self* to my *self* now, and I can also be the *same self* across time. While the present self may indeed identify with a past or future self and so be concerned for it and appropriate its actions, it *need* not. One can be a self for only a moment—*not so for a person*. In keeping with its forensic, pastward-looking role, person *necessarily* "extends itself beyond present existence to what is past" (II.xxvii.26); it is that "same thinking thing in different times and places" (II.xxvii.9). In other words, <personR> signifies, in part, the relation of the diachronic sameness of self.[160] And the fact that <personR> denominates relations among its relata

---

[159] Locke insists, "It [is] impossible for any one to perceive, without perceiving, that he does perceive" (II.xxvii.9). He defines consciousness as simply "the perception of what passes in a Man's own mind" (I.i.19), and later elaborates on the inextricable link between consciousness and thinking/perception, stating: "consciousness always accompanies thinking"; it is "inseparable from thinking, and it seems to me essential to it" (II.xxvii.9). Locke even defines "thinking" in terms of consciousness: "thinking consists in being conscious that one thinks" (II.i.19). For a careful analysis of what it means for Lockean consciousness to perceive, see Lähteenmäki (2011), where he details how consciousness perceives simultaneously not only the subject and object *qua* ideas, but also the *act* of thinking.

[160] Although I believe that drawing this distinction between the self and the person renders Locke's account most coherent, if you remain unpersuaded, you may still endorse a version of the relation-

---

(i.e., a past self and present self) is not unusual. $<\text{Person}_R>$ *qua* $<\text{same self}_R>$ signifies the diachronic identity relation *of a relation,* much as $<$same sister$>$ does.[161]

The Synchronic Identity Relations of the past and present selves that are the relata of the $<\text{person}_R>$ provide the epistemic building blocks that make possible knowledge of the Diachronic Identity Relation that is partly constitutive of the $<\text{person}_R>$. And knowledge of the Synchronic Agential Relation of a past self, when taken together with the Diachronic Identity Relation, makes possible knowledge of the second constitutive relation of the $<\text{person}_R>$, the Diachronic Agential Relation.

If $<\text{person}_R>$ always signifies, at least in part, [R1] the Diachronic Identity Relation, then that seems to entail that the expression "same person" is actually redundant, and I think that in a sense it is. I propose that although Locke frequently uses the expression "same person," what he means, strictly speaking, is "same (personal) self." Understanding persons as relations also seems to entail the collapse of the ideas of person and personal identity, though considered more carefully, "personal identity" refers to the Diachronic Identity Relation that obtains between instantiations of the personal self, and which is constitutive of the person.

---

interpretation. Where I take the self to be among the relata of the person, one could substitute agent, thinking thing, or man.

[161] Given that self too is an idea of a twofold relation ($<\text{self}_R>$)—a Synchronic Identity Relation and a Synchronic Agential Relation—we can gain a more fine-grained understanding of $<\text{person}_R>$ by unpacking the $<\text{present self}_R>$ and $<\text{past self}_R>$ that are its relata. As argued, $<\text{person}_R>$ is foremost [R1] the Diachronic Identity Relation of the present self with a past self, by means of which the present self may then appropriate the action of the past self in the form of [R2] the Diachronic Agential Relation. These two relations jointly satisfy the forensic role of person. Yet these two diachronic relations are made possible only by a set of four conceptually prior relations that collectively constitute the $<\text{present self}_R>$ and $<\text{past self}_R>$: [R3] the Present Synchronic Identity Relation, [R4] the Present Agential Relation, [R5] the Past Synchronic Identity Relation, and [R6] the Past Agential Relation.

This is intended to be the minimally complex structure of <person_R>. Of course, the instances of past self-conscious agency with which a given present self identifies are likely numerous. Moreover, the present self may also look forward with concernment for the future self, which may constitute further secondary relata.[162] Although <person_R> can be dissected into multiple relations, the overarching idea of person_R is all of the constitutive relations taken collectively, the relation of being oneself to oneself, the same thinking and acting thing as that which existed in a different time and place.

The fact that the <person_R> itself is analyzable into a number of constitutive relations does not make it a strange or unusually complex relation. Many other Lockean relations can also be analyzed into multiple relational constituents. For instance, the idea of the relation <grandfather> signifies a man S who stands in the relation of begetter to a child P (Relation 1), where P stands in relation of begetter to some child Q (Relation 2), and Q stands in relation of grandchild to S (Relation 3) (II.xxv.7). Similarly, the idea of the relation <king> may be analyzed into an immense number of constitutive relations, equal in number to the population of subjects in respect of whom the man *qua* king rules. Moreover, the fact that <person_R> denominates actions among its relata and links its relata

---

[162] Locke also gestures toward the future in his discussion of personal identity, counting as the same self whoever's pain and happiness we are concerned with as our own. The difference between the past and future relata of the person is that the past relata must directly or indirectly denominate actual past substances, while any future relata do not denominate existent substances, only *ideas*. However, given Locke's forensic account of persons, judgments of sameness of self are always made looking toward the past to determine moral accountability. Still, so long as some present individual is concerned with his future self, that *idea* of the future self might constitute one of the relata of the relation <person>. That relational component of the <person_R> is not used to underwrite moral accountability, but only anticipates a *potential* future responsible subject of the present agent's actions. Locke speaks elsewhere of relations that point toward some future function in II.xxvi.6.

by means of conscious appropriation is also not unique. Take, for example, the idea of the relation <murderer>: it denominates not only the man who is the perpetrator and the man who is the victim but also the activity of killing and the mental activity of intention.

Notice that substance has not featured in this account of person at all. This is important given Locke's insistence that sameness of substance is not necessary for sameness of self. Substances are not even among the *immediate* relata of <person$_R$>. When the relata themselves are further unpacked in <self$_R$>, we arrive at the thinking thing that is the primary relatum of the present self and past self. Though these thinking things consist in substances of some kind, Locke is agnostic about their constitution: He says, "*Self* is that conscious thinking thing, (whatever Substance, made up or whether Spiritual, or Material, Simple or Compounded, it matters not)" (II.xxvii.17). Each of the relata of <person$_R$> directly or indirectly denominates some substance, for consciousness can occur only as a mode of a (thinking) substance, and actions—whether mental or physical—are also modes of substances. In this way, <person$_R$> denominates present and past "thinking, intelligent Being[s]."[163] But, in keeping with Locke's non-substantialist account of personhood, the

---

[163] In arguing for a substance-interpretation of persons, Rickless (2015) gives voice to the only objection to a relation-interpretation of which I am aware, pointing to II.xxvii.9, where Locke states that "person" stands for "a thinking, intelligent Being, that . . . can consider it self as it self, the same thinking thing in different times and places." Rickless emphasizes that the Lockean "person" stands for a positive being and suggests that in this passage, Locke "may, however, be telling us that persons are not *relations*" (Rickless 2015, 111). Yet the reason that Locke classifies substances and modes, but not relations, as positive beings is that an idea of a substance or a mode is "the complex idea of one thing," while ideas of relations arise from the consideration of "two *Ideas*, or Things, either in themselves really separate, or considered as distinct" in respect of some ground for comparison (II.xxv.6). According to Locke's definition, the person does *not* signify just *one* positive being (substance-interpretations) or an affection of a being (mode-interpretations). Rather, it signifies one being that considers itself the same as some other thinking thing(s) "in different times and places." What Locke does *not* want to say is precisely that the person is necessarily tied to the same substance or the same "one being" over time.

substance denominated by the present relatum may be otherwise *entirely unrelated* to the substance or substances denominated by the past relata. Like the relation <king>, <person_R> may link multiple, distinct substances. <Person_R> unites present and past self-conscious agents, regardless of the substance or substances in which each inheres/inhered, for the purpose of tracking moral responsibility.

3. The Metaphysics of Persons_R

If the *idea* of a person is an *idea* of a relation, then a person itself is a relation. But what precisely this means is not obvious, given that Locke's ontological account of relations is much disputed. Ontological interpretations of Lockean relations (often linked with modes)[164] range wildly, including everything from reductionism to conceptualism and from anti-realism to realism.[165] Some of the ambiguity of Locke's account derives from his habit of conflating talk of *ideas* of things with the things themselves. And given that a principal component of Locke's project is developing a taxonomy of ideas, teasing apart his account of ideas from his ontological theory is not a straightforward enterprise. Rendering Locke's account of relations coherent requires determining when Locke is addressing ideas of relations and when he is speaking of relations in the world (if indeed there are any). By examining Locke's account of relations, two metaphysical features emerge that endow the relation-interpretation of persons with a number of virtues, including the resources to claim

---

[164] Stuart (2013a) represents an exception, arguing that modes exist in the world while relations do not.

[165] For recent surveys of the range of ontological interpretations of Lockean relations, see Rickless (2018) and Ott (2009, 159–69).

the principal advantages of both the substance- and mode-interpretations. Persons as relations are (1) mind-dependent and yet also (2) well-founded by virtue of answering to the existence of the mind-independent substances in which their relata terminate.

Much of the interpretive debate concerning the metaphysics of Lockean relations centers on Locke's claim that relations are "not contained in the real existence of Things, but something extraneous, and superinduced" (II.xxv.8; see also II.xxv.4).[166] One might think that Locke is suggesting here that relations are mere creatures of the mind that do not correspond to objects in the world in any way. But that can't be right, since in the very next breath, he tells us that relations "belong" to substances (II.xxv.8). Locke insists that ideas of relations are not "Copies of Things" in nature, being mere "*Idea*[*s*] of my own making," yet he also claims that ideas of relations "denominate" actual "relations . . . **as they come to exist**" (III.x.33; bold added). Relations, thus, do pick out something in the world, even if ideas of relations are not copies of mind-independent things—a kind of reverse direction of fit from ideas of substances.

That relations are somehow both mind-dependent and yet also owe their existence to the mind-independent world of particulars is confirmed by Locke's definition of the *nature* of relations:

> *The nature* therefore *of Relation* consists in the referring, or comparing two things, one to another; from which comparison, one or both comes to be denominated. And if either of those things be removed, or cease to be, the **relation ceases**, and the denomination consequent to it, though the other receive in itself no alteration at all.

---

[166] This passage occupies the center of much of the interpretive debate concerning the ontology of Lockean relations. For varied interpretations of this passage, see, for instance, Bennett (1971, 253–54) for a reductionist interpretation; Langton (2000) for the interpretation that relations are in some sense irreducible, mind-independent entities; Green (1885, 35), Gibson (1917, 193–95), and Stuart (2013a) for anti-realist takes; and Odegard (1969) and Rickless (2018) for realist interpretations.

> *V.g. Cajus,* whom I consider to day as a Father, ceases to be so to morrow, only by the death of his Son, without any alteration made in himself. (II.xxv.5)

Here, Locke tells us that relations "consist in" something mental,[167] viz. the mental act of referring or comparing. Nevertheless, relations "denominate" extramental things in the world that conform to them. Moreover, Locke states that a relation *itself*—not the *idea* of a relation—"ceases" upon the cessation of any of its relata and, consequently, the relata are no longer denominated by it.[168] It is not the case that the *idea* of father must cease upon the death of Cajus's son. Rather, Cajus himself ceases to *be* a father in the absence of any children. Similarly, if all the subjects of a kingdom perished, that does not mean that the *idea* of the king must cease. Rather, the man himself ceases to *be* a king in the absence of any subjects to rule.[169] If I consider that subject-less man under the relation king, then since there is no such state of affairs conformable to that relation, (i) my judgment is mistaken, (ii) my relational idea <king> is "false" in the Lockean sense of true/false ideas,[170] and (iii) that particular relation does not exist in the world. Thus, relations *themselves* consist (at least in part) in the mental act of comparison, while also *existentially* depending on the mind-independent particular substances in which their relata terminate.

---

[167] Locke at times states that a relation is *a way of* comparing, referring, or considering *x* in respect of *y* (II.xxv.7; III.x.33) and at other times that a relation *is* or *consists in* comparing, referring, or considering *x* in respect of *y* (II.xxv.10; II.xxv.5; II.xii.7).

[168] It would be difficult to make sense of Locke's discussion of the cessation of a relation if he intended this passage to be a discussion of mere ideas. Cf. Rickless (2018, 12), who argues that in passages such as this, Locke really means to speak of *ideas* of relations, not relations in and of themselves.

[169] Locke elsewhere emphasizes this point when discussing natural relations such as father, stating that relations are "as lasting as the Subjects to which they belong" (II.xxviii.2).

[170] "When-ever the Mind refers any of its *Ideas* to any thing extraneous to them, they are then *capable to be called true or false*" (II.xxxii.4).

We are now in a position to put a finer point on the two metaphysical features of Lockean relations: (1) relations are mind-dependent insofar as they involve the mental act of comparison, but (2) the mind-independent particular substances provide (i) truthmakers for our judgments about relations, (ii) truthmakers for our ideas of relations in the Lockean sense of true/false ideas, and (iii) a metaphysical foundation for relations themselves.[171] Neither a strict conceptualist/anti-realist account that locates relations solely in the mind nor a realist account that grants relations full-fledged mind-independence sufficiently captures the complex status of Lockean relations, which have, as it were, one foot in the mind-dependent domain, insofar as they consist in the mental act of comparison, and the other foot in the mind-independent domain, insofar as they depend for their very existence on the existence of the particular substances in which their relata terminate.

By applying these two metaphysical features of Lockean relations to the case of a person$_R$, we can derive two claims: (1) A person$_R$ is mind-dependent insofar as it is an idea arrived at through a mental act of considering the present self (i) vis-à-vis some past self/selves in respect of identity and (ii) vis-à-vis some past action(s) in respect of agency. But (2) the mind-independent particular substances in which the relata of a person$_R$ terminate provide a metaphysical foundation for that person$_R$ and serve as truthmakers for our idea of and judgments about that particular person$_R$.

---

[171] Ott's (2009, 159–69) foundational conceptualism interpretation of Lockean relations nicely captures these two dimensions of the ontology of Lockean relations. Ott sums up his foundational conceptualist account of Lockean relations as follows: "[W]hile relations are fully mind-dependent and have no real being, it remains the case that the mind-independent world provides a foundation (and a justification) for us to form the ideas of relations that we do" (2009, 167).

These two metaphysical features of person_R jointly yield significant payoffs for the relation-interpretation, both epistemological and metaphysical. A thorough treatment of all of these interpretive advantages is beyond the scope of this chapter, so for the present purposes, I will briefly survey several of the most significant virtues of this account. On the epistemological front, Locke famously endorses a strong form of privileged access when it comes to first-personal judgments concerning one's own person. Locke claims that my own person is something I can "be sure of" (II.xxvii.23), and he is confident that on Judgment Day, "no one shall be made to answer for what he knows nothing of; but shall receive his Doom, his Conscience accusing or excusing him" (II.xxvii.22). Locke ought not, however, endorse infallibility when it comes to first-personal judgments about our own person. Many interpretations of Lockean personal identity, including the popular "Memory Theory," commit him to infallibility as well as circularity. On a relation-interpretation, however, in keeping with Locke's forensic intuitions, the mind-dependence of the person_R affords epistemic privilege in first-person judgments of our own personhood, supported by the fact that the immediate relata of the person_R—our past and present selves and actions—are accessed by intuitive knowledge, by virtue of consciousness and memory. Yet the fact that actually existent persons_R must answer to substances as their foundations and truthmakers provides a non-circular basis for excluding infallibility. In this way, the relation-interpretation precludes false memories and delusional imaginings from constituting persons.

Understanding persons as relations also delivers metaphysical payoffs for Locke's account of persons. The metaphysically thin picture of the person as a mind-dependent

relation not only honors the non-substantialist spirit of Locke's account of personal identity but also sidesteps Reid's failure of transitivity objection. This is because the identity relations that are constitutive of the person$_R$ *are* transitive relations and are not threatened by the intransitivity of memory. Importantly, the primary relatum of person$_R$ is always the *present* self, so person$_R$ is invariably relativized to the present moment. Whatever past self-conscious agents *y, r, s* with which the *present* self *x* identifies and whose actions the present self *x* consciously appropriates constitute person$_R$ $\boldsymbol{x}$G*yrs*, and the transitivity of identity holds among that set of selves, past and present. In addition, persons as relations can, in virtue of their relata, be founded in a number of *distinct* substances at once, without tying persons to the *same* substance across time. Recognizing persons as relations is, thus, naturally compatible with Locke's account of the person as surviving change in substance, whether material or even immaterial.

Moreover, these two metaphysical features of Lockean relations enable the relation-interpretation to claim the principal advantages of the substance- and mode-interpretations. First, the fact that persons$_R$ are in some sense mind-dependent enables the relation-interpretation to claim the principal advantage of the mode-interpretations. Second, the fact that particular, existent persons$_R$ must nonetheless be well-founded equips the relation-interpretation with the resources to claim the principal advantage of the substance-interpretations.[172]

---

[172] Simendic (2015, 94–95) also points to the textual evidence that supports a relation-interpretation's claim to the respective principal advantages of the substance- and mode-interpretations, but he does not discuss the metaphysical features of Lockean relations or how/why these advantages follow from these metaphysical features.

First, because persons play the forensic role in Locke's system of tracking the morally/legally accountable agent, the idea of a person *ought* to be the kind of idea that we can plug into his demonstrative moral science. In fact, it ought to be one of the principle moral ideas. Proponents of mode-interpretations of persons have emphasized that since the real essence of modes can be perfectly known and ideas of modes are always adequate, modes are uniquely suited to Locke's demonstrative science of ethics.[173] If Locke's forensic <person> is to be a part of moral demonstrations, mode-interpreters reason, then persons *must* be modes and not substances. The relation-interpretation, however, lays equal claim to this advantage: by virtue of the mind-dependence of relations, their real essence can also be perfectly known, and ideas of relations too are always adequate.

If persons were substances, they could not be proper objects of a demonstrative moral science. To the contrary, they would be objects of natural philosophy, which Locke famously insists falls short of demonstration, reaching, at best, probable opinion. This is because our ideas of substances are never adequate, meaning that they do not perfectly represent their archetypes since we can never be assured that our nominal essences correspond to the real essences of substances owing to the fact that their microconstitution lies beyond the scope of human knowledge.[174] By contrast, in the domain of morality, as in mathematics, we may attain demonstrative knowledge precisely because moral ideas are

---

[173] See LoLordo (2010, 662–64; 2012, 79–82); in fact, LoLordo uses the example of a relation, <father>, to illustrate this point (2012, 81).

[174] Locke emphasizes that complex ideas of substances are never adequate (IV.iv.12), because they represent mind-independent archetypes that consist in an "abstruse hidden Constitution" (III.xi.17); see also II.xxxi.13. Of nominal vs. real essences of substances, see III.vi.2.

adequate and the real essences of their referents are knowable. Since relations are in some respect mind-dependent, being "something extraneous, and superinduced" (II.xxv.8) and consisting "in the referring, or comparing two things" (III.x.33), the idea of person as an idea of a relation is always an adequate idea, perfectly representing its archetype (II.xxxi.3 and II.xxxi.14).[175] Likewise, Locke claims that moral words must stand for things whose real essence can be perfectly known.[176] By contrast with substances, the real essence of a relation like the person$_R$ *just is* its nominal essence, and so it cannot but be perfectly known.[177]

Now, one might worry that, given the second metaphysical feature of persons as relations, viz. their answering to the existence of particular substances, <person$_R$> is *not* always adequate. Yet the adequacy of <person$_R$> concerns only the archetype, or the

---

[175] On the adequacy of moral ideas, Locke states, "our *moral Ideas*, as well as mathematical, being *Archetypes* themselves, and so adequate, and complete *Ideas*" (IV.iv.7); see IV.iv.7–8. Moreover, on the clarity and distinctness of ideas of relations as compared with those of substances, see II.xxv.8 and II.xxviii.19.

[176] Locke states, "Morality is capable of Demonstration, as well as Mathematicks: Since the precise real Essence of the Things moral Words stand for, may be perfectly known; and so the Congruity, or Incongruity of the Things themselves, be certainly discovered, in which consists perfect Knowledge" (III.xi.16).

[177] Of ideas of relations and modes, Locke states, "being such collections of simple ideas that the mind itself puts together, and such collections that each of them contains in it precisely all that the mind intends that it should, **they are archetypes and essences**" (II.xxxi.14; bold added). By contrast, the real essence of a substance is unknowable, being "the constitution of the insensible parts of that Body, on which those Qualities, and all the other Properties" of a substance depend (III.vi.2). Proponents of substance-interpretations, including Winkler (1991), Gordon-Roth (2015, 109), and Rickless (2015, 121–22), have pointed to Locke's discussion of the "moral man," where he *seems* to suggest that ideas of substances, like that of a <moral man>, may be used in moral demonstrations (III.xi.16). See LoLordo (2010, 655–56; 2012, 84) for an analysis of this passage in support of a mode-interpretation, where she argues that "moral man" is simply a linguistic signifier of the idea of person, without carrying ontological implications, an argument that a relation-interpreter may also support. After all, "moral man" does not signify man, the single substance, simpliciter. We might unpack "moral man" as glossing "person$_R$" by signifying a man considered vis-à-vis his actions in respect of their moral value.

relational structure of the idea of a person as outlined in section 1, which is the same for all persons. Whether or not an idea of a *particular* person_R is well-founded will not impact the adequacy of that idea but instead bears on its *truth.* As stated, so long as a state of affairs is conformable to a given idea of a relation, whenever that state of affairs is considered under that relation, the idea of that relation is true, and the relation can be said to exist. And it is because actually existent particular relations are necessarily founded in particular, existent substances that the relation-interpretation can also claim the principal advantage of substance-interpretations.

If, as demanded by Locke's forensic intuitions, a person is what links some past agent with some present recipient of judgment, then a person better be the kind of thing that can actually wield agential power. Proponents of substance-interpretations argue that only a substance can possess agential power and tout this as a principal advantage over mode-interpretations.[178] And Locke does indeed state that *"Powers* belong only to *Agents,* and *are Attributes only of Substances*" (II.xxi.16). But as argued in section 1, an agent is properly understood as a relation (i.e., a way of *considering* a substance vis-à-vis an action in respect of generation), and agents are among the relata of person_R. A person_R is thus perfectly suited to possess agential power. But even if one insists that agents *are* substances, given the second metaphysical feature of relations, a person_R can still possess agential power by virtue of the substances in which it is founded, and which are denominated by its relata.

---

[178] See, for instance, Rickless (2015) and Gordon-Roth (2015).

In fact, Locke explicitly attributes power to other kinds of relations that, like person$_R$, are ways of considering a man. For instance, of relations such as dictator or constable, Locke says, "it be certain, that either of them hath a certain Power over some others" (II.xxviii.3). Locke goes on to say,

> Sometimes the foundation of considering Things, with reference to one another, is some act, whereby any one comes by a **Moral Right, Power, or Obligation** to do something. Thus a *General* is one, that hath power to command an Army; and an Army under a General, is a Collection of Men, obliged to obey one Man. A *Citizen,* or a *Burgher*, is one who has a Right to certain Privileges in this or that place. (II.xxviii.3; bold added)

If these relations may have moral rights, power, obligations, and privileges, surely persons as relations enjoy these same endowments.[179] Locke reiterates his claim that relations provide a foundation for moral obligations and duties, stating,

> It is very convenient, that by distinct Names, these Relations should be observed, and marked out in Mankind, there being occasion, both in Laws, and other Communications one with another, to mention and take notice of Men, under these Relations: From whence also arise the Obligations of several Duties amongst Men. (II.xxviii.2)

Person$_R$ is one of the relations under which we take notice of men on occasions of making judgments about moral accountability, from which arises the appropriateness for reward or punishment, similar to the obligations and duties as Locke describes them here. To be

---

[179] According to Gordon-Roth (2015, 104 n. 18), LoLordo has recruited this passage in support of the mode-interpretation, claiming that what is true of relations must also be true of modes, yet this is not necessarily so. In her response in favor of the substance-interpretation, Gordon-Roth acknowledges that relations may in some sense be considered to have powers, though she denies it of modes, and claims that, "if any relation has a power it is because at least one of the entities in the relation is a substance. In other words, any powers a constable has, he has because he is a substance that stands in a particular relation R to something else" (2015, 104). It is not quite right to say of a constable that "he is a substance." Rather, a man who is considered under the relation constable, and who is the primary relatum of the relation constable, is a substance. Still, relations do have power by virtue of the substances in which their relata terminate.

sure, relations have power in virtue of the *things* that they directly or indirectly denominate, yet this does not diminish the fact that constables and persons have agential power. Of persons, Locke states, "The same consciousness uniting those distant Actions into the same *Person,* whatever Substances contributed to their Production" (II.xxvii.10). Here, he suggests that although there may be many substances among the relata *contributing* to the production of actions, one *person* is the agent of these actions, appropriating responsibility for these deeds by means of consciousness.

Conclusion

Given that Locke is not explicit about the ontological classification of persons or the classification of ideas of persons, some rational reconstruction is required to flesh out a metaphysical account of the Lockean person. Some have argued that persons do not fit neatly within the threefold ontological framework that he inherited, and to which he was perhaps not wholeheartedly committed,[180] or even that Locke's inquiry into persons demonstrates the limits of his own empirical project, given that the person eludes classification among the available options. Still others have suggested that metaphysics simply is not what Locke was up to in his account of persons, which is merely an

---

[180] Some have suggested that Locke may admit of additional unknown ontological categories. For instance, Weinberg argues that person is the objective, metaphysical fact of a continued consciousness, which occupies a *sui generis* ontological category that falls outside the above threefold ontology (2016, 160–61). On this interpretation, which amounts to a kind of agnostic account of persons, the idea of person could scarcely always be adequate, as called by Locke's demonstrative science of ethics. Nor could it explain how persons are endowed with agential power. It cannot even afford privileged access to knowledge of our own persons since, according to Weinberg, the person *qua* ongoing consciousness enjoys a "temporal existence through any gaps in my successive, subjective states of awareness of myself" (2016, 153).

epistemological or psychological account of how we *know/take* ourselves to be the same person across time.[181] But there are good reasons to resist these kinds of interpretive moves. Locke is committed to the adequacy of our ideas of persons as well as to "sure" intuitive knowledge of our own person, neither of which accord with agnosticism. At the very least, Locke should wish to assign persons to one of his three kinds of complex ideas of substances, modes, and relations that correspond to his ontological categories. After all, a central part of Locke's project is providing a comprehensive taxonomy of ideas. It would not be an insignificant problem for Locke if the idea of person—a critical idea that grounds moral accountability in his system—both defies categorization within his taxonomy of ideas and has *no* metaphysical foundation.

The fact that Locke presents his account of persons in one of his chapters on relations provides a helpful hint for understanding his intent. Given the demands of Locke's forensic account of persons and the central, driving commitments of his project, Locke *ought* to classify persons as relations. This view makes the most sense of the most text while also conforming to the unique structure of Lockean relations and honoring the distinction between the self and the person. The two metaphysical features of Lockean relations, based on which a person$_R$ is mind-dependent yet well-founded, enable this relation-interpretation to claim the respective principal advantages of the substance- and mode-interpretations: as with modes, the idea of person$_R$, being always adequate, is a suitable Lockean moral idea applicable in his demonstrative science of ethics. And as with substances, the person$_R$ is

---

[181] See, for instance, Newman (2015).

capable of possessing agential power. Moreover, understanding persons as relations yields a number of epistemological and metaphysical advantages, including: (i) meeting Locke's demand for privileged access when it comes to first-personal judgments concerning one's own personal identity, while also (ii) affording non-circular grounds for excluding false memories from constituting persons, (iii) respecting the transitivity of identity, and (iv) according with Locke's non-substantialist intuitions about personal identity.

EPISTEMOLOGICAL PAYOFFS OF A

RELATION-INTERPRETATION OF THE LOCKEAN PERSON

Contemporary cultural and political discourse evidences two competing epistemic intuitions concerning *descriptive* identity claims like those pertaining to gender, ethnicity, sexual orientation, and so on, which bear on the discussion of *personal* identity. On the one hand, there is a growing recognition that the first-person stance affords a uniquely authorized position for judging one's own identity. Those moved by this intuition hold that the first-person stance is sufficient to justify the descriptive identity judgment that, for instance, "I self-identify as female," in the same way that it is sufficient to justify other judgments classified as self-knowledge, like, "I am experiencing a toothache." On the other hand, there is pushback from a competing intuition that claims that introspection from the privileged position of the first-person stance *alone* is insufficient to underwrite descriptive identity claims. Those moved by this intuition maintain that, beyond the fact that, for instance, "I self-identify as Native American," some corroborating evidence—in this case, genetic evidence or membership on a tribal roll—is required to underwrite such an identity claim.

Versions of these same two competing epistemic intuitions may also be observed in the case of *personal* identity. On the one hand, something seems intuitively right about the assertion that I occupy a uniquely authorized position to make judgments about my own

personal identity—after all, I alone have occupied a front-row seat to each moment of my conscious life (omniscient deities excepted). Yet there is a competing intuition that insists that there be some grounding criteria beyond the privileged access of the first-person stance to justify such judgments. While we might be quick to affirm *our own* first-personal authority when it comes to judgments about our personal identity, we might also want there to be an objective fact of the matter when it comes to the question of whether or not, say, Jim Jones is the same person as the Buddha, Jesus Christ, and Vladimir Lenin, as he claimed. In other words, we might think that Jim Jones's privileged epistemic standpoint with respect to his own consciousness and memories should not entail his *infallibility* with respect to judgments concerning his personal identity. These two competing epistemic intuitions, I suggest, must somehow be accommodated and reconciled in any satisfactory account of personal identity.

The first intuition is that I enjoy some kind of privileged access and first-person authority with respect to judgments concerning my own personal identity (hereafter the "Privileged Access Intuition"). The Privileged Access Intuition derives from the recognition that knowledge of our own personal identity is a variety of self-knowledge, coupled with the acknowledgment of the privacy of consciousness, memory, and introspection, which together constitute the means of such knowledge. The Privileged Access Intuition is in *prima facie* tension with a second intuition, which says that there should be some objective fact of the matter concerning personal identity in virtue of which we cannot unconditionally defer to first-person authority to guarantee the veridicality of beliefs concerning personal identity (hereafter the "Fallibility Intuition"). In other words, the fact that I believe or introspect that I am identical with some previous agent *alone* is insufficient to serve as the truthmaker

166

of a personal identity claim. The Fallibility Intuition is, then, a resistance to *infallibility* when it comes to first-personal judgments of personal identity, following from the sentiment that it should be possible that we be deluded about our own identity, as in the case of an individual suffering from dissociative identity disorder or in the case of the occurrence of a false memory. A reconciliation of these two competing intuitions, the Privileged Access Intuition and the Fallibility Intuition, will yield some kind of account of *qualified* privileged access in judgments concerning one's own diachronic personal identity.

A number of classic objections to *Locke's* account of personal identity are related in some way to whether or not Locke is taken to accommodate the Fallibility Intuition, and many of these objections are also based on the prevalent Memory Theory interpretation of Locke's account of personal identity, which says simply that personal identity consists in memory. Locke locates the identity condition for persons in sameness of consciousness, and on the Memory Theory, memory is identical to consciousness. On this interpretation, memory is both necessary and *sufficient* for sameness of person.

There are broadly two different ways that "memory" on this reading of Locke has been interpreted, both resulting in unwanted consequences for his account. The first, which I will refer to as the "Seeming Memory Theory,"[182] says that the "memory" that is taken to constitute personal identity refers to *seeming* memory, viz. any mental state that is *taken to be* a memory by the subject. On the Seeming Memory Theory, Locke is unable to

---

[182] This interpretation of Locke's account of personal identity begins with Joseph Butler (1736) and Thomas Reid (1785), and has been followed by numerous interpreters since then. Reid, for instance, charges that memory, or consciousness (which collapse on his reading), has for Locke "a strange magical power of producing its object" (III.vi). For a recent rendition of this interpretation, see Stuart (2013a).

accommodate the Fallibility Intuition and therefore cannot preclude false memories from *also* constituting the person. This is because seeming memories are inclusive of both veridical memories and pseudo-memories, where veridical memories represent an actual past perception of the subject, while pseudo-memories do not. On this reading, Locke is committed to the undesirable consequence that one may be *justly* punished for deeds that were never committed, but which were psychologically appropriated in virtue of their being presented to consciousness by a pseudo-memory. This does not sit well with Locke's central concern in his forensic account of person that it be the basis for appropriating actions and their merit.

The second version of the Memory Theory, which I will refer to as the "Veridical Memory Theory," says that the "memory" that is taken to constitute personal identity refers only to veridical, truth-tracking memories. On the Veridical Memory Theory, Locke *is* able to accommodate a modest form of the Fallibility Intuition, but the account is circular,[183] since a purported memory can only be certified as veridical by recourse to sameness of person, but sameness of person can only be confirmed by recourse to veridical memory. The interpretive challenge, then, is how Locke is to exclude false memories through an account of *qualified* privileged access that accommodates both the Privileged Access Intuition and the Fallibility Intuition, while at the same time avoiding the pitfall of vicious circularity.

---

[183] Butler (1736, 100) is first credited with this objection.

An account of qualified privileged access calls for a *metaphysical* account of persons. If Locke's person were purely a psychological or moral notion,[184] then it would be difficult to make room for fallibility. The tendency to read the Lockean person as *merely* psychological stems from the understanding of Locke's identity condition for persons—sameness of consciousness—as also being purely psychological. This is exemplified in the Memory Theories, with their reduction of sameness of consciousness—and person—to memory. A metaphysical account of persons is required in order to allow for the kind of appearance-reality distinction that is necessary to account for fallibility in first-personal judgments of personal identity.

If there is to be a metaphysical account of the Lockean person, then, given his threefold classification of complex ideas from which has been derived a corresponding tripartite ontology, Locke's person must be a substance, a mode, or a relation. Since Locke does not explicitly classify person within his taxonomy of ideas or assign it to an ontological category, interpreters have taken up debating whether or not Locke's person is best understood as a substance or a mode.[185] But there is a neglected alternative that has received almost no attention: that Locke's person is a relation.[186] There is, in fact, some

---

[184] Relatedly, some, like Newman, maintain that Locke's account of personal identity is purely *epistemic*: "Locke's broader aim is to clarify the conditions under which we *judge* that we are numerically the same with some earlier person, not the conditions under which we strictly *are* numerically the same person" (2015, 90).

[185] Proponents of the substance-interpretation include Alston and Bennett (1988), Chappell (1990), Winkler (1991), Gordon-Roth (2015), and Rickless (2015), while contemporary proponents of the mode-interpretation include Mattern (1980), Uzgalis (1990), Thiel (2011), and LoLordo (2012), who follow Law (1769).

[186] While this project has been underway, Simendic (2015) published an article making a case for the Lockean person to be read as a relation. Although there is some overlap, his relation-interpretation of

striking *prima facie* evidence for this reading: Locke's presentation of persons turns up in one of his chapters on relations.[187] I suggest that the Lockean person is *neither* a substance, like a body or a soul, *nor* a mode, like a number or an activity, but a *relation,* like a father or a friend. I contend that a relation-interpretation of the Lockean person has more resources than do the substance or mode views for providing a coherent metaphysical account of persons that accords with Locke's forensic intuitions more generally, while at the same time accommodating both the Privileged Access Intuition and the Fallibility Intuition when it comes to judgments of personal identity. For the purposes of this paper, however, I will bracket the substance and mode views, as my principal aim here is to demonstrate the epistemological payoffs that result from pursuing a relation-interpretation of persons.

After providing a sketch of what a relation-interpretation of the Lockean person might look like in section 1, I will turn to the epistemological payoffs of the view in section 2. There I will argue that understanding the Lockean person as a relation affords a uniquely supportive interpretive basis for an account of *qualified* privileged access in first-personal judgments of diachronic personal identity by providing *metaphysical criteria* for persons that: (1) honor Locke's strong Privileged Access Intuition, (2) accommodate the Fallibility Intuition, and thereby preclude false memories from constituting persons, and (3) serve as truth conditions for personal identity judgments that are non-circular.

---

persons differs significantly from that presented here, and the points of agreement and disagreement will be noted throughout.

[187] II.xxvii, "Of Identity and Diversity."

1. A Relation-Interpretation of the Lockean Person

Although in ordinary language we may take the word "person" to be synonymous with "man" or "human," the Lockean "person" is, famously, a technical, forensic term serving the role of identifying some present agent with some past agent for the purposes of tracking the morally or legally responsible party. "Person" is, for Locke,

> a Forensick Term appropriating Actions and their Merit; and so belongs only to intelligent Agents capable of a Law, and Happiness and Misery. This personality extends it self beyond present Existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it self past Actions, just upon the same ground, and for the same reason that it does the present. (II.xxvii.26)[188]

Sameness of consciousness, which serves as the identity condition for the forensic "person," is not coextensive with sameness of *man*[189] or sameness of *substance* (material or immaterial). Locke considers it conceivable, and thus possible, that consciousness need not always be annexed to the same man or substance, potentially transferring bodies and even intellects.[190] Moreover, on Locke's view, given that conscious appropriation of a past agent's action is necessary for that agent to constitute the same person, and at any given moment a man is only able to consciously appropriate a limited subset of the actions he has undertaken throughout his life, it follows that a variety of persons may be indexed to the

---

[188] John Locke, *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch (Oxford: Oxford University Press, 1975). References to the *Essay* are by book, chapter, and section.

[189] Locke assigns man the identity condition of "a participation in the same continued Life" (II.xxvii.6).

[190] Locke famously provides a number of thought experiments in support of this claim, e.g., the prince and cobbler (II.xxvii.15).

same man throughout his lifetime.[191] Thus, sameness of man or substance is not only *insufficient* to link a prior agent with a present recipient of judgment, but it is not even *necessary*. Since neither sameness of man nor sameness of substance can guarantee that reward or punishment is justly meted out, a separate forensic category of "person" is relied upon to *unite* present and past agents partaking in the same consciousness, regardless of the substance(s) in which that consciousness inheres.

If the idea of <person> is, as I suggest, an idea of a relation, then it must in some way answer to Locke's definition of an idea of a relation as a complex idea "which consists in the consideration and comparing one *Idea* with another" (II.xxv.7).[192] Locke states that relations presuppose: "two Ideas, or Things, either in themselves really separate, or considered as distinct, and then a ground or occasion for their comparison" (II.xxv.6). Ideas of relations, then, are ideas that involve considering one idea or thing, *x*, vis-à-vis another idea or thing, *y*, in respect of some ground for comparison, G (xGy). In all of Locke's examples, the idea of the relation denominates more than one relatum, but one relatum is primary ($\boldsymbol{x}Gy$, where the primary relatum is in the first place and bolded). Locke provides a lengthy list of examples of relations that consist in considering one *man* (as the primary relatum) vis-à-vis some secondary relata in respect of some ground:

---

[191] See, for instance, his account of the sleeping and waking Socrates as two distinct persons (II.xxvii.19).

[192] Locke defines ideas of modes as "complex *Ideas*, which however compounded, contain not in them the supposition of subsisting by themselves, but are considered as Dependences on, or Affections of Substances" (II.xii.4). Examples of ideas of modes include simple modes such as the idea of a dozen, and mixed modes, such as the ideas of beauty, theft, and fencing. Ideas of substances are identified as "such combinations of simple *Ideas*, as are taken to represent distinct particular things subsisting by themselves; in which the supposed or confused *Idea* of Substance, such as it is, is always the first and chief" (II.xii.5). Examples of ideas of substances include the ideas of a man and a sheep.

> [T]here is *no one thing*, whether simple *Idea*, Substance, Mode, or Relation, or Name of either of them, *which is not capable of almost an infinite number of* Considerations, in reference to other things: and therefore this makes no small part of Men's Thoughts and Words. *v.g.* One single Man may at once be concerned in, and sustain all these following **Relations**, and many more, *viz.* Father, Brother, Son, Grandfather, Grandson, Father-in-Law, Son-in-Law, Husband, Friend, Enemy, Subject, General, Judge, Patron, Client, Professor, European, English-man, Islander, Servant, Master, Possessor, Captain, Superior, Inferior, Bigger, Less, Older, Younger, Contemporary, Like, Unlike, etc. (II.xxv.7)

Much like father, friend, professor, and possessor, I contend that *person* is just another relation under which we might consider a *man*.

If the idea of person is an idea of a relation (hereafter <person$_R$>), what are the <person$_R$>'s relata, and what ground are the relata considered in respect of? Locke states that "person" stands for:

> a thinking, intelligent Being, that has reason and reflection, and can consider **it self as it self,** the **same** thinking thing **in different times and places**; which it does only by that consciousness, which is inseparable from thinking . . . (II.xxvii.9; bold added)

A person, then, is firstly something that considers itself *as* itself. In other words, <person$_R$> involves a relation of *synchronic* identity. Furthermore, even when considered independently from the enterprise of determining personal identity, the <person$_R$> involves the relation of sameness "in different times and places." Thus, the <person$_R$> also involves a *diachronic* identity relation; <person$_R$> signifies a being who not only considers herself as herself in the present but who also considers her *present* self to be the same as some "thinking thing" in another time and place. Finally, <person$_R$> also signifies relations of agency, both in the present and the past, as indicated by the fact that the present self and past self are each described as a "thinking, intelligent Being," or "thinking thing," that is, an *agent* of

173

thought.[193] In other words, the <person$_R$> involves the relation of being the agent, and thus appropriator of a set of actions, past and present, that are indexed to the same consciousness.

Locke further explains how the <person$_R$> signifies two kinds of constitutive relations: identity and agency—when he states,

> For as far as any intelligent Being can repeat the *Idea* of any past Action with the same consciousness it had of it at first, and with the same consciousness it has of any present Action; so far it is the same *personal self.* For it is by the consciousness it has of its present Thoughts and Actions, that it is *self* to it *self* now, and so will be the same *self* as far as the same consciousness can extend to actions past or to come. (II.xxvii.10)

I suggest that the <person$_R$> necessarily involves a *minimum* of six relations in order to carry out its forensic role of underwriting moral accountability: two synchronic identity relations, one diachronic identity relation, and three agential relations.[194]

Relation 1, "Present Synchronic Identity Relation": *person involves the synchronic identity relation of being oneself to oneself in the present.* This relation is the present synchronic identity of the *self*.[195] Although Locke often uses the terms "self" and "person"

---

[193] That "thinking thing" here refers to an agent is clarified elsewhere in this chapter; e.g., person is "a Forensick Term appropriating Actions and their Merit; and so belongs only to **intelligent Agents** capable of a Law, and Happiness and Misery" (II.xxvii.26; emphasis added).

[194] Cf. Simendic's relation-interpretation, wherein he maintains that person *qua* relation is an idea "comprised of two principal elements—substance and personal identity" (2015, 80). Simendic regards personal identity as "the relation between multiple diachronic iterations of the same consciousness" (ibid.). He takes person to be a complex idea involving some kind of relation between the *substance*, which is taken to be the person by substance-interpretation proponents, and the *mode* of the continuation of consciousness, taken to be the person by mode-interpretation proponents. It is, in effect, a hybrid of the substance- and mode-interpretations.

[195] Locke asserts, "we have an intuitive Knowledge of our own *Existence*" (IV.iii.21), and, "it is evident, that, considered in any instant of its existence, it is in that instant the same with itself" (II.xxvii.3); in this consists synchronic identity.

interchangeably, they are not strictly equivalent; "person" always signifies a relation of *diachronic* identity, whereas "self" need not.[196] It is the person—not the self—that is a forensic entity linking some present subject of judgment with some past agent for the purpose of tracking moral accountability (II.xxvii.26), and it is self—not person—that is the synchronic object of knowledge of the cogito (IV.iii.21). Locke defines "self" as a synchronic, self-conscious agent:

> When we see, hear, smell, taste, feel, meditate, or will any thing, we know that we do so. Thus it is always as to our **present** Sensations and Perceptions: And by this every one is **to himself**, that which he calls *self*. (II.xxvii.9; bold added)

By contrast, he consistently describes the person in diachronic terms, as that which "extends itself beyond present existence to what is past" (II.xxvii.26). Moreover, as cited above, Locke uses "self" and relations to the self as *constituents* of the definition of "person" (II.xxvii.9). If "self" and "person" were strictly equivalent, then Locke's definition of person would be circular. The Present Synchronic Identity Relation may be signified by xSx, where S stands for a relation of synchronic identity, and x stands for the self of the present moment. The present self is the *primary relatum* of the person$_R$, just as a woman is the primary relatum of the relation mother. Where mother is one relation under which we might consider a woman vis-à-vis some child in respect of generation, person$_R$ is one way in which we might consider the present self vis-à-vis a set of secondary relata in respect of a set of

---

[196] Lähteenmäki (2018) also distinguishes the Lockean "self" from "person," though in a different way: "a person is the sensible self taken together with those thoughts and actions the sensible self attributes to itself as its own by being conscious of them." Cf. Simendic (2015, 81), who takes "self" and "person" to be first-person and third-person equivalents. Simendic does not accept any synchronic relations as constituents of person, since he does not appear to acknowledge the existence of a synchronic self.

grounds. To unpack these secondary relata and their respective grounds, we turn to the remaining constituent relations of the $<person_R>$.

Relation 2, "Diachronic Identity Relation": *the $<person_R>$ involves the present self standing in a relation of diachronic identity with (at least one) past self*.[197] The Diachronic Identity Relation is signified by xCy, where C stands for the ground for comparison, viz. sameness of consciousness, *x* stands for the present self as its primary relatum, and *y* stands for a past self as its secondary relatum.[198] Because forensics—whether literally a legal matter or a question of moral accountability on Judgment Day—is necessarily pastward looking, the $<person_R>$ always involves an idea of a Diachronic Identity Relation, uniting the present self with a past self-conscious agent.[199] This relation is a prerequisite for the conscious

---

[197] Simendic also includes a version of the Diachronic Identity Relation as a constituent of his relational person, insofar as he regards personal identity to be a constituent of the person, where personal identity amounts to the diachronic "identity of consciousness" (2015, 92).

[198] This account seems to entail that the expression "same person" is redundant, and I think that in a sense it is. Although Locke frequently uses the expression "same person," what I suggest he means, strictly speaking, is "same self," which is a constituent of the idea of person. This follows from Locke's definition of what a person stands for, which involves the identification of the present self with a past self-conscious agent (II.xxvii.9). The relation-interpretation also seems to entail the collapse of the ideas of "person" and "personal identity," though considered more precisely, "personal identity" is actually an abstract term descriptive of the *kind* of relation signified by "person," just as "fatherhood" is to the relation "father." This represents another point of departure between Simendic's interpretation and that presented here, since, as noted earlier, Simendic regards personal identity to be a *constituent* of person.

[199] Although Locke also gestures towards the future in his discussion of personal identity, counting as the same person whoever's pain and happiness we are concerned with as our own, the future self indicates a relatum that is no more than an *idea*, being presently only potential. Given the forensic context for Lockean personal identity, however, *judgments* of sameness of person are always made pastward to determine moral accountability, and never forward. Still, so long as some present agent is concerned for his future self, the *idea* of the future self might constitute one of the relata of the person, despite the fact that it cannot denominate a particular substance. That relational component of the person, however, is not used to underwrite actual moral accountability, but only to anticipate a *potential* future subject who would be responsible for the present agent's actions.

appropriation of past actions that is necessary for the Lockean person to underwrite moral accountability.

Relation 3, "Past Synchronic Identity Relation": *<person_R> involves the synchronic identity relation of being oneself to oneself at some point in the past.* It is in virtue of sameness of consciousness that the present self extends pastward to identify with some previous agent, and this self-identification is made possible by the reflexivity of Lockean consciousness, both present and past. In other words, some past agent *y* is capable of being consciously identified with some present agent *x* only if *y,* like *x,* was also self-conscious, or was "itself to itself." The Diachronic Identity Relation, then, depends not only on a Present Synchronic Identity Relation but also on a Past Synchronic Identity Relation, $\textbf{\textit{y}}Sy$. This subjectivity, or synchronic sense of self-identity, is also necessary for regarding *oneself* as standing in the relation of agency in respect of an action, which is another component essential to the forensic <person_R>.

The <person_R> also involves a minimum of three agential relations:[200] **Relation 4, "Present Synchronic Agential Relation"**: *<person_R> involves a present self standing in a relation of agency in respect of some present (mental or physical) action.* Locke's forensic "person" is fundamentally a diachronic moral agent, with the present self described as a thinking being, or an agent of thought. The Present Synchronic Agential Relation consists in the consideration of the present self (*x*) vis-à-vis some present action (*z*), in respect of generation (G), signified by $\textbf{\textit{x}}Gz$.

---

[200] Simendic does not include agential relations as constituents of his Lockean person.

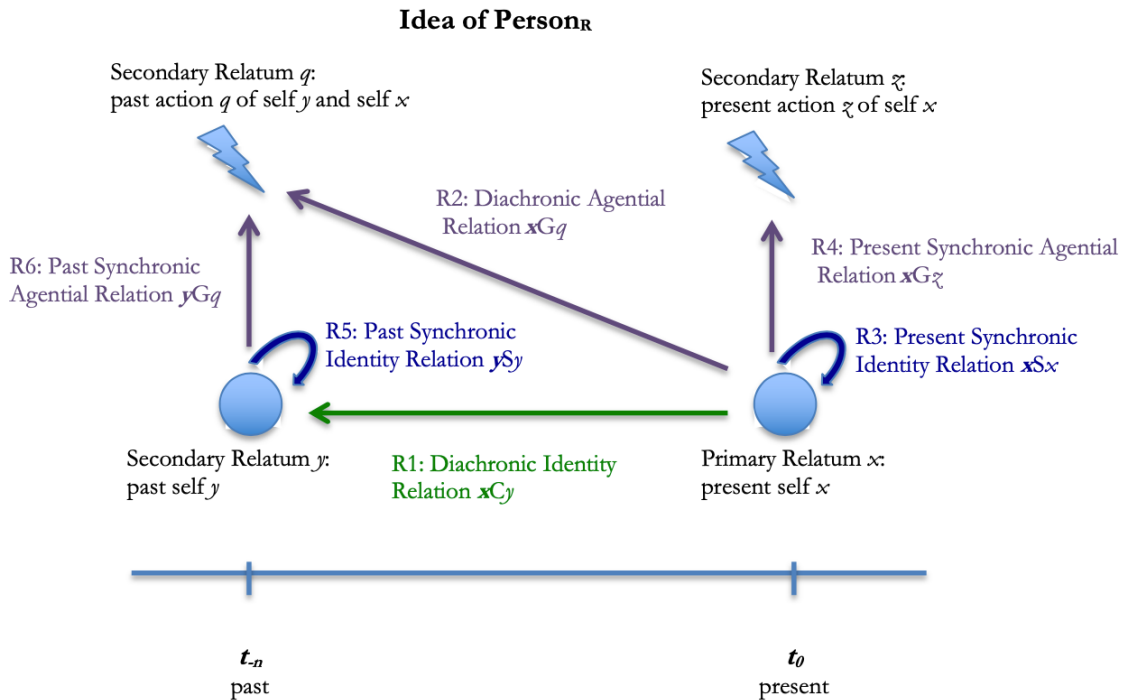Given that the Lockean "person" is a forensic term, concerned foremost with tracking accountability for some *past* action, a second agential relation is required: **Relation 5, "Past Synchronic Agential Relation"**: *<person_R> involves a past self standing in a relation of agency in respect of some past (mental or physical) action.*[201] The Past Synchronic Agential Relation consists in the consideration of some past self y vis-à-vis some past action q in respect of generation G, that is, $\boldsymbol{y}Gq$.

Yet the relation of central concern to forensics is that of agency born by the *present* self in respect of some *past* action. This is **Relation 6, "Diachronic Agential Relation"**: *<person_R> involves the present self standing in a relation of agency in respect of a past (mental or physical) action.* This relation is necessary for the "appropriation" of past actions by a present agent. The formation of the idea of the Diachronic Agential Relation is directly based on two of the earlier constitutive relations: the Past Synchronic Agential Relation taken together with the Diachronic Identity Relation. If the past self y stands in a relation of agency in respect of some past action q ($\boldsymbol{y}Gq$), and the present self x stands in a relation of diachronic identity with y ($\boldsymbol{x}Cy$), then x is in a position to appropriate the actions of y as her own: the <person_R>, then, also involves the consideration of present self x vis-à-vis some past action q in respect of generation G, $\boldsymbol{x}Gq$.

---

[201] This is intended to be the minimally complex structure of the idea of the relation <person>. Of course, the instances of past self-conscious agency with which a present individual identifies are numerous, and, as noted above, the present self may also look forward with concernment for the future self, which may constitute further relata.

Figure 9: The Idea of a Person as an Idea of a Sixfold Relation

Idea of Person$_R$

Secondary Relatum q:
past action q of self y and self x

Secondary Relatum z:
present action z of self x

R2: Diachronic Agential
Relation xGq

R4: Present Synchronic Agential
Relation xGz

R6: Past Synchronic
Agential Relation yGq

R5: Past Synchronic
Identity Relation ySy

R3: Present Synchronic
Identity Relation xSx

Secondary Relatum y:
past self y

R1: Diachronic Identity
Relation xCy

Primary Relatum x:
present self x

t_{-n}
past

t_0
present

The present self, x, is the primary relatum of the person$_R$, standing as the head vis-à-vis itself (R3) as well as vis-à-vis each of the secondary relata: the past self y (R1, xCy), the present action z (R4, xGz), and the past action q (R2, xGq).

The fact that the person as a relation itself involves a number of constitutive relations does not make it a strange or unusually complex relation. Many other Lockean relations can also be analyzed into multiple relational constituents. For instance, the relation <grandfather> signifies a man S who stands in the relation of begetter to a child P (Relation 1), where P stands in relation of begetter to a child Q (Relation 2), and Q stands in relation of grandchild to S (Relation 3) (II.xxv.7). Similarly, the relation <king> may be analyzed into an immense number of constitutive relations, equal in number to the population of subjects in respect of whom the man *qua* king stands as ruler.

If the *idea* of person is an *idea* of a relation, then from a metaphysical perspective, persons *are* relations. which involves two key claims: (1) relations are mind-dependent insofar as they involve the mental act of comparison,[202] and yet (2) the mind-independent world of particular substances provides a foundation for relations and justification for the formation of our ideas of and judgments about relations.[203] Persons as relations, then, are in some sense mind-dependent insofar as the person$_R$ consists in the consideration of the present self-conscious agent vis-à-vis some past self-conscious agent(s) and action(s) in respect of sameness of consciousness and generation, respectively. But just as whether or not a given woman is *actually* a mother depends upon the existence of the substances denominated by the relata of mother (viz. a woman and her child), persons *qua* relations must also answer to the mind-independent world of particular substances as their foundation and truthmakers. Each of the relata of an actually existent person$_R$ must

---

[202] Locke states, "This farther may be considered concerning *Relation*, That though it be not contained in the real existence of Things, but something extraneous, and superinduced: yet the *Ideas* which relative Words stand for, are often clearer, and more distinct, than of those Substances to which they do belong" (II.xxv.8). What precisely Locke means here has fostered much debate among interpreters. For instance, Bennett (1971, 253–54) takes Locke to mean that relations are all reducible in some sense; Langton (2000) endorses the interpretation that relations are in some sense irreducible, mind-independent entities; Green (1885, 35) and Gibson (1917, 193–95) both interpret this passage as a denial of the reality of relations; and Stuart takes this to mean that "we project our ideas of relations on to the world rather than copying them from it," and concludes that for Locke, "there are ideas of relations, but no relations" (2013a, 31, 30). But to say in an unqualified way that there are no relations and only ideas of relations seems unsatisfying, for as Stuart also acknowledges, Locke wants relations to play an important and ubiquitous role in the world. Locke identifies powers as relations (II.xxi.19; II.xxiii.37), or sometimes as including relations (II.xxi.3), and states that most of our simple ideas from which we form complex ideas of substances (excepting figure and bulk) are powers, which are "Relations to other Substances" (II.xxxi.8).

[203] Locke describes the substances in which the relata terminate as the foundation and truthmakers of the relations in the following passage: "*The nature* therefore *of Relation* consists in the referring, or comparing two things, one to another; from which comparison, one or both comes to be denominated. And if either of those things be removed, or cease to be, the relation ceases, and the denomination consequent to it, though the other receive in itself no alteration at all" (II.xxv.5).

terminate in some actually present or past existent particular substance. Consciousness can occur only as a mode of a (thinking) substance, and actions too—whether mental or physical—are modes, or affections, of substances.[204] In this way, the constituents of the person$_R$ that are "thinking, intelligent Being[s]" both present and past are denominated by the relational idea of person. Importantly for Locke's account of personal identity, the substance denominated by the present relatum may be otherwise *entirely unrelated* to the substance(s) denominated by the past relata. Like the relation king, a person$_R$ may link multiple distinct substances. The person$_R$ is the relational bridge uniting present and past self-conscious agents, regardless of the substance or substanc*es* in which each inheres/inhered, for the purpose of tracking moral responsibility.[205] It is the metaphysics

---

[204] Of course, as with all Lockean complex ideas, when fully analyzed, ideas of relations, their relata, and their grounds for comparison ultimately terminate in simple ideas: "*all Relation terminates in*, and is ultimately founded on those *simple Ideas*, we have *got from Sensation*, *or Reflection*: So that all that we have in our Thoughts ourselves, (if we think of any thing, or have any meaning,) or would signify to others, when we use Words, standing for Relations, is nothing but some simple *Ideas*, or Collections of simple *Ideas* compared one with another" (II.xxviii.18). Locke illustrates this point using the example of "father": "*v.g.* when the Word Father is mentioned: *First*, There is meant that particular Species, or collective *Idea*, signified by the Word Man; *Secondly*, Those sensible simple *Ideas*, signified by the Word Generation; and *Thirdly*, The Effects of it, and all the simple *Ideas*, signified by the Word Child" (II.xxviii.18).

[205] Locke claims, "'tis plain consciousness, as far as ever it can be extended, should it be to Ages past, unites Existences, and Actions, very remote in time, into the same Person, as well as it does the Existence and Actions of the immediately preceding moment: So that whatever has the consciousness of present and past Actions, is the same Person to whom they both belong. Had I the same consciousness, that I saw the Ark and *Noah*'s Flood, as that I saw an overflowing of the *Thames* last Winter, or as that I write now, I could no more doubt that I, who write this now, that saw the *Thames* overflow'd last Winter, and that view'd the Flood at the general Deluge, was the same *self*, place that *self* in what Substance you please, than that I that write this am the same *my self* now whilst I write (whether I consist of all the same Substance, material or immaterial, or no) that I was Yesterday. For as to this point of being the same *self*, it matters not whether this present *self* be made up of the same or other Substances, I being as much concern'd, and as justly accountable for any Action was done a thousand Years since, appropriated to me now by this self-consciousness, as I am, for what I did the last moment" (II.xxvii.16).

of Lockean relations, I will argue, that enables the relation-interpretation of persons to support a significant set of epistemological payoffs for Locke's account of personal identity.

2. The Relation-Interpretation of Persons and Qualified Privileged Access

I have suggested that the relation-interpretation provides a set of metaphysical criteria for the person$_R$ that enables an account of *qualified* privileged access with respect to first-personal judgments of diachronic personal identity, accommodating both the Privileged Access Intuition and the Fallibility Intuition by means of excluding infallibility. In making this case, I will begin in section 2.1 with an analysis of privileged access in the context of first-personal judgments of diachronic personal identity and demonstrate the ways in which Locke endorses a strong Privileged Access Intuition. In section 2.2, I will present the metaphysical criteria for the person$_R$ afforded by the relation-interpretation that provide the basis for the qualification of privileged access through the exclusion of infallibility. This qualification will provide the grounds for precluding false memories from also constituting the person$_R$. Finally, in section 2.3 I will clarify the distinction between the metaphysical criteria for the person$_R$ on the relation-interpretation and the psychological process of forming first-personal judgments concerning diachronic personal identity. Here it will be necessary to distinguish consciousness from memory, and to spell out the roles that each plays in these two domains. The differentiation between the psychological and metaphysical domains will make clear that, on the relation-interpretation, Locke's account of personal identity does *not* suffer from vicious circularity, or so I will argue.

2.1 Privileged Access

In order to forward the account of *qualified* privileged access afforded by a relation-interpretation of the Lockean person, it is first necessary to spell out what, exactly, I take privileged access to entail in first-personal judgments of diachronic personal identity. I will consider five claims commonly associated with privileged access, which have been adjusted for the case of personal identity:[206]

(1) **Epistemic asymmetry**: *S may have a type of warrant for the belief that she is identical with P that is unavailable to others*. On Locke's view, I enjoy epistemic asymmetry in judgments concerning my diachronic personal identity since I alone possess non-inferential justification for the belief that I am identical with some past agent, namely, by means of introspection, or for Locke, "intuitive knowledge." Given that Locke's identity condition for persons is sameness of consciousness, which is imperceptible by the senses, and which is accessible by means of intuition and only by the subject to whom the consciousness belongs, *third*-personal beliefs concerning personal identity can be justified only by means of inference, or for Locke "demonstrative knowledge," or "reasoning." As Locke asks rhetorically, "Can another Man perceive that I am conscious of anything, when I perceive it not my self? No Man's Knowledge here, can go beyond his Experience" (I.i.19).

(2) **Epistemic privilege**: *S's belief that he is identical with P can, in principle, achieve a higher degree of epistemic certainty than others' beliefs that S is identical with P*. Locke would have it that I also enjoy epistemic privilege, or a higher degree of epistemic certainty than others,

---

[206] Derived from claims associated with privileged access as compiled and formulated by Gertler (2003, xii).

in judgments concerning my diachronic personal identity, since, having epistemic asymmetry, I alone have *direct*, unmediated access to the identity condition for persons, viz. consciousness, by way of the epistemic instrument of intuition. Others have only indirect access to my consciousness by means of inferential reasoning, yet inferences about sameness of consciousness cannot afford certainty, since the basis for such inferences, like past and present sense perceptions of the body and actions of the agent in question, are not necessarily invariably connected with a given consciousness on Locke's non-substantialist account of personal identity. For this reason, properly speaking, others cannot *reliably* have even *indirect* access to another individual's consciousness. Moreover, Locke's account of the different means of knowledge is structured in terms of "*degrees* of knowledge," where intuitive knowledge, demonstrative knowledge, and sensitive knowledge are hierarchically structured in descending degrees of epistemic certainty. So even if a third-personal judgment of diachronic personal identity were based on the most reliable inference achievable, it would still be, in principle, inferior in its degree of certainty to the intuitive knowledge available only to the first-person stance.

(3) **Incorrigibility** (or "first-person authority"): *S's claim that she is identical with P cannot be justifiably disputed by others, or cannot be shown by others to be mistaken.* On Locke's account, it also seems as if I also enjoy some form of incorrigibility, or first-person authority, such that my belief in my diachronic personal identity cannot be justifiably disputed by others, or demonstrated to be mistaken. Given epistemic privilege, third-personal judgments concerning diachronic personal identity are possible only on the basis of inferential reasoning, which definitionally obtains a lesser degree of epistemic certainty than first-

personal intuitive knowledge for Locke. Moreover, given epistemic asymmetry, such inferences are not even based on direct, sensitive knowledge of the "evidence" or identity condition for persons, sameness of consciousness, which is in principle not achievable. Thus, the third-person stance looks incapable of justifiably disputing first-personal beliefs concerning diachronic personal identity. Locke famously acknowledges the shortcomings of the human justice system in this regard, noting that, "want of consciousness cannot be proved" (II.xxvii.22).

(4) **Self-intimation**: *If S is identical with P, then he believes that he is identical with P.* Locke would also endorse the claim that I enjoy self-intimation such that if I am identical with some past agent, then I believe that I am identical with that past agent.[207] Or, at the very least, Locke would endorse this weaker construal: *If S considers whether he is identical with P, then S will have an occurrent, episodic belief that he is identical with P.*[208] This accords with one of Locke's central intuitions regarding justice, that no one should be held accountable for a deed that he is not cognizant of having committed. As Locke puts it, "no one shall be made to answer for what he knows nothing of; but shall receive his Doom, his Conscience accusing or excusing him" (II.xxvii.22). Self-intimation follows from understanding two of Locke's guiding principles in his forensic account of persons: (1) it is just that I be held accountable for the deeds of some past agent *if and only if* I am the same *personal self* as that

---

[207] Analogously, just as we could potentially summon any subject and recognize a man as king in relation to him or her, one may summon any memory of past self-conscious agency and recognize it as denominated by one's person.

[208] This weakened form is modified from the case of first-personal knowledge of mental states in Gertler (2003, xiii). Of course, "self-intimation" as used here differs from the "self-intimation" generally associated with consciousness and self-knowledge.

past agent, and (2) I am the same personal self as some past agent *only if* I consciously

identify with that past agent and appropriate her deeds. Thus, if I am identical with some

past agent, it must be the case that I *believe* that I am the same personal self as that past

agent.

Finally, (5) **Infallibility**: *If S believes that she is identical with P, then she is, in fact, identical with P.* Given Locke's strong account of privileged access by way of his endorsement of

some form of epistemic asymmetry, epistemic privilege, incorrigibility, and self-intimation,

how might he manage to exclude infallibility, so as to accommodate the Fallibility Intuition?

2.2 Qualified Privileged Access: Excluding Infallibility and False Memories

The *qualification* of privileged access in first-personal judgments of diachronic

personal identity accords well with Locke's project more generally of delimiting the scope

of human knowledge. To be sure, for Locke, we are the (mortal) authority when it comes

to our own personal identity, but we are not infallible. This is because conscious

appropriation of a past action that is presented by one's memory is *necessary* for its

punishment/reward to be just, but it is not *sufficient.*[209] As noted, on the Seeming Memory

Theory interpretation of Locke's account of personal identity, in which seeming memory

constitutes personal identity, infallibility looks inevitable, and the Fallibility Intuition must

be sacrificed. This is because on this interpretation, memory is both necessary and *sufficient*

---

[209] Given that conscious appropriation is a necessary (though not sufficient) condition for a past agent and action to constitute the person, infallibility must also be rejected in order to exclude cases where first-personal judgments concerning one's personal identity are formed on the basis of testimony or inference rather than introspection.

for sameness of person$_R$. And on the Veridical Memory Theory interpretation, in which only veridical memory constitutes personal identity, the Fallibility Intuition is modestly accommodated, but at the cost of circularity, resulting in incoherence. In limiting personal identity to being constituted by either seeming memories or veridical memories, the Memory Theories reduce Locke's account of personal identity to merely a psychological account. If the problematic consequences that follow from these interpretations are to be sidestepped, a metaphysical account of persons is also required. Given the ontology of Lockean relations as mind-dependent and yet also beholden to the existence of particular substances for their foundation and justification, the relation-interpretation is uniquely capable of supporting both Locke's strong Privileged Access Intuition while also honoring the Fallibility Intuition.

2.2.1 The Metaphysical Criteria for Person qua Relation

As I have suggested, the means by which the relation-interpretation is able to exclude infallibility hinges on the very nature of relations, which provides a pair of metaphysical features for the person$_R$: The first is sameness of consciousness, which is widely accepted as the identity condition for the Lockean person (hereafter the "Consciousness Criterion"). I take sameness of consciousness to be not merely a psychological criterion, as the Memory Theories would have it, but a *metaphysical* criterion (more on what this means later). If the relation-interpretation is correct and persons are relations, then there is a second condition that is necessary to underwrite a justified belief concerning diachronic personal identity. Any person *qua* relation that picks out a real

187

human agent in the world must also meet the well-foundedness criterion for a Lockean relation: it must terminate in existent particular substance(s) denominated by its relata, and which serve as its truthmaker(s) (hereafter the "Well-founded Criterion"). While ideas of relations are always adequate, they are not always *true*. The truth value of ideas, Locke claims, is assigned when there is a "secret or tacit Proposition" (II.xxxii.1):

> When-ever the Mind refers any of its *Ideas* to any thing extraneous to them, they are then *capable to be called true or false*. (II.xxxii.4)

Whenever the idea of a given person$_R$ denominates actual relata "extraneous to it," it may be assigned a truth value. The nature of relations is such that if any of the purported relata do *not* terminate in existent particular substances, then the relation in question is unfounded; if my child does not exist, I am not a parent. As Locke states,

> *The nature* therefore *of Relation* consists in the referring, or comparing two things, one to another; from which comparison, one or both comes to be denominated. And **if either of those things be removed, or cease to be, the Relation ceases**, and the Denomination consequent to it, though the other receive in itself no alteration at all. *v.g. Cajus*, whom I consider to day as a Father, ceases to be so to morrow, only by the death of his Son, without any alteration made in himself. (II.xxv.5; bold added)

In other words, as soon as *any* of the purported relata of a relation "father" fail to terminate in an existent particular substance, that relation father ceases to exist in the world, and its application is unfounded. Similarly, for a correct judgment of the identity of any purported person$_R$ (which involves the relation between a present self-conscious agent and some past self-conscious agent), if any of the relata fail to terminate in a particular substance that either really *did* or *does* exist, then that person *qua* relation does not exist, and its application is unfounded. Of course, the *diachronic* relation person$_R$ differs from the *synchronic* relation

"father" in that some of the person$_R$'s relata, viz. a past self-conscious agent and her action, *never* exist in the present, though they must have exist*ed* in the past.[210]

In sum, persons as relations (1) are mind-dependent insofar as they involve the mental act of comparison, but (2) the mind-independent particular substances in which their relata terminate provide (i) truthmakers for our judgments about relations, (ii) truthmakers for our ideas of relations in the Lockean sense of true/false ideas, and (iii) a metaphysical foundation for relations in the world.[211]

Recall that the person *qua* relation can be analyzed into (at least) six constituent relations: Present and Past Synchronic Identity Relations, a Diachronic Identity Relation, Present and Past Synchronic Agential Relations, and a Diachronic Agential Relation. The relata of the person$_R$ thus include (at least) a present self-conscious agent, a past self-conscious agent, and present and past (physical or mental) actions. Given self-intimation, any purported *past* relata must be presented by a seeming memory as a necessary prerequisite for their conscious appropriation in the present. But their presentation by a seeming memory is not sufficient to certify the purported relata as genuine. If the person$_R$ to whom these purported relata belong is to be well-founded, then all of its relata must denominate, or terminate in, particular substance(s) that actually do, or did, exist.

---

[210] This is similar in some respects to a relation like "result." For example, when considering a sprout under the relation "result," at least one of its denominated relata, viz. the seed *qua* cause, no longer exists at the time of its result, but it must have exist*ed* in the past.

[211] Ott's (2009, 159–69) foundational conceptualism interpretation of Lockean relations nicely captures these two dimensions of the ontology of Lockean relations. Ott sums up his foundational conceptualist account of Lockean relations as follows: "[W]hile relations are fully mind-dependent and have no real being, it remains the case that the mind-independent world provides a foundation (and a justification) for us to form the ideas of relations that we do" (2009, 167).

To summarize, the relation-interpretation provides *two metaphysical criteria*, both necessary, only jointly sufficient, to underwrite judgments concerning personal identity:

1. **Consciousness Criterion**: Sameness of consciousness

2. **Well-founded Criterion**: A person$_R$ must terminate in actually existent particular substance(s) as its truthmaker(s) in virtue of its relata.

Taken together, these two metaphysical criteria for the person$_R$ provide the grounds for the qualification of privileged access to knowledge of diachronic personal identity through providing a basis for the exclusion of infallibility. To see how, it may be helpful to consider a few examples.

2.2.2 Applying the Metaphysical Criteria as Truth Conditions to Exclude False Memories

Case Study 1: Purely fictitious false memory

> *I have been implanted with a false memory of having shoplifted this morning.*

According to Locke, a *necessary* condition for my being held accountable for this act is that I have consciously appropriated it through an intuition-based judgment of sameness of consciousness with the agent of that act:

> [N]o one shall be made to answer for what he knows nothing of; but shall receive his Doom, his Conscience accusing or excusing him. (II.xxvii.22)

My conscience accuses me of shoplifting this morning. Is my conscious appropriation of this fictitious action *sufficient* to justify my being punished for it? On the Seeming Memory Theory, it would appear so. This poses a problem, for surely Locke would not want his forensic account of person to entail the justification of punishment for deeds that were never

190

committed. In this example, it *seems* as if, in virtue of intuitive knowledge, the Consciousness Criterion is met. The Well-founded Criterion, as uniquely afforded by the relation-interpretation, however, does *not* obtain. The pseudo-memory of my act of shoplifting, being purely fictitious, does not terminate in a particular substance. There was no such past shoplifting human agent in which my consciousness could have inhered to justify this judgment. For this reason, the fact that the Well-founded Criterion was *not* met functions as a defeater of this judgment of personal identity.

The relation-interpretation is, then, also capable of excluding *incorrigibility* as formulated in the above list of claims associated with privileged access, allowing for the further qualification of privileged access and accommodation of the Fallibility Intuition. With the Consciousness Criterion alone, as maintained by other interpretations of Lockean personal identity, the *third*-person stance cannot afford a position to justifiably dispute first-personal identity claims since consciousness is accessible only by intuition and only by the subject to whom the consciousness belongs. But given the Well-founded Criterion unique to the relation-interpretation, if a purported relatum of a given person$_R$ were somehow known by some third party to be a fictitious, nonexistent substance, then that third party *would* be in a position to reasonably dispute that first-personal identity claim. Although "want of consciousness cannot be proved," the absence of a given material substance at a certain time and place *can,* in some instances, be proved (II.xxvii.22). Locke might still endorse this weakened version of the incorrigibility: *S's claim that she is identical with P cannot be justifiably disputed by others, or cannot be shown by others to be mistaken, so long as the substance denominated by P existed at the indicated time and place.*

Already, we have seen how the Well-founded Criterion manages to preclude purely fictitious false memories from constituting the person$_R$ through providing the basis for the exclusion of infallibility and even the stronger incorrigibility claim. However, it is not enough to say that the purported relata of a person$_R$ must terminate in just *any* particular substance(s). More must be said about these substances to fill out the picture of how their denomination functions as a metaphysical criterion for persons *qua* relations, and thus as a truth condition for judgments concerning personal identity. To this end, let us consider a second example.

Case Study 2: Mismatched false memory

Before considering the scenario of mismatched false memory in the case of the person$_R$, it may be helpful to first consider an analogous scenario of mismatch in the case of another relation, that of "biological mother".

A. Analogous case: Mismatch in the case of the relation "biological mother"

> *Anna, an adopted child, mistakenly believes her adoptive mother, Bertha, is her biological mother,*
> *when in fact Cathy is her biological mother.*

Anna's judgment of Bertha as her biological mother is mistaken. What accounts for the mistake? It is *not* mistaken because (a) there is no existent substance that could *rightly* be denominated by the relation "Anna's biological mother" (Cathy *does* exist). *Nor* is it mistaken because (b) Anna's judgment of Bertha as her biological mother did not denominate an existent particular substance (Bertha—a real, particular substance—*was* denominated).

192

Rather, the judgment is mistaken because one of the relata denominated the *wrong* particular substance, a substance that is incapable of providing a foundation for the relation "biological mother of Anna." It is a case of a mismatch. Here, the Well-founded Criterion, some version of which is common to all relations, *seems* to obtain. The relation "biological mother" seems to terminate in particular substances in virtue of its purported relata— Bertha and Anna. However, the Well-founded Criterion is checked by the unique, defining criterion for the relation "biological mother," which is a Causal Criterion that requires that the substances denominated by the purported relata stand in the requisite causal relation of female begetter and begotten child. This explains how we can guarantee the veridicality of judgments in the case of the relation "biological mother," but how is the Well-founded Criterion to be checked in the case of the person$_R$? How are we to be sure that the *right* substance(s) have been denominated?

B. The relation person: Mismatched false memory

> *I have been implanted with a memory of shoplifting this morning that belongs to my friend, who did in fact undertake the action of shoplifting this morning.*

In this scenario, it appears that the Well-founded Criterion obtains, since there *does* exist a particular substance that can be denominated by the relatum "past shoplifting agent" as presented by the pseudo-memory; my friend is indeed a human agent who engaged in this very act of shoplifting. The defining criterion for the person$_R$ should do the work of checking that the *right* kind of substance has been denominated by the purported relata, just as the Causal Criterion did in the case of the biological mother. But it *seems* as if the

Consciousness Criterion is also met, based on intuition: in virtue of the presentation of a seeming memory, it appears as though my consciousness stretches back to an agent who shoplifted. The two metaphysical criteria for the person$_R$ as formulated above, which are intended to serve as truth conditions for judgments of personal identity, look incapable of excluding cases of mismatch.

How are we to determine the right kind of substance that may adequately qualify as a foundation for the person$_R$? In the case of making a correct judgment about the relation "biological mother," the relata must terminate in substances that are *causally* related in the requisite way, but in the case of "person," a hallmark of Locke's account of personal identity is that, not only does sameness of person *not* require sameness of substance, but any two substances connected under the relation of person may be otherwise *entirely* unrelated. These substances need not bear an identity relation, a causal relation, or even a relation of spatiotemporal contiguity. Sameness of consciousness is the only necessary condition that Locke explicitly places on personal identity, and thus the only necessary link between the substances denominated by a given person. Locke famously provides a variety of thought experiments to demonstrate that very point, like that in which the consciousness of a prince transfers substances to inhere in the body of a cobbler (II.xxvii.15). Here, a single person$_R$ terminates in two otherwise unrelated substances in virtue of sameness of consciousness.

Yet such a scenario of the transference of consciousness is not the same as the case of an implanted memory. In the transference of consciousness scenario, since the earlier moment of consciousness that inhered in the prince was indeed the *same consciousness* as that which later inhered in the cobbler, the person$_R$ may denominate differing substances in

194

such a way that they serve as truthmakers for the judgment of personal identity. In the case of an implanted memory, however, my present consciousness is not the same as the consciousness of my past shoplifting friend. Despite my present seeming memory, this morning, my consciousness did not perceive my undertaking the act of shoplifting. So, just as in the case of the misidentified biological mother, in my judgment of self-identification with the shoplifting agent, the *wrong* substance is denominated. The substance is "wrong" because it is not capable of providing a foundation for the person$_R$; it was not the bearer of the same consciousness.

If sameness of consciousness were reducible to memory, and Lockean personal identity were merely a psychological account, then the Consciousness Criterion could not check the Well-founded Criterion to ensure that the *right* kind of substance is denominated. As the Consciousness Criterion was phrased earlier simply as "sameness of consciousness," it was ambiguous between being a psychological and metaphysical criterion. Its apparent failure to exclude cases of mismatching is just another example of the problems that result when sameness of consciousness is misunderstood as being purely psychological, rather than a metaphysical fact.[212] The Consciousness Criterion must be disambiguated through

---

[212] I suggest that the much-disputed "fatal error" passage (II.xxvii.13) supports understanding sameness of consciousness as a metaphysical fact. Locke concludes in this passage that consciousness could conceivably transfer intellectual substances, whereupon "two thinking Substances may make but one Person." But what he deems *impossible* is the "fatal error" in which consciousness, "which draws Reward or Punishment with it," transfers *persons* or "agents," whereupon it presents to an agent an act for which she is not morally responsible, and which is "without reality of Matter of Fact" like a dream. This Locke regards as inconceivable due to God's goodness. Some have understood this as excluding the possibility of false memories; however, I contend that it underscores the distinction between memory and consciousness. There is an objective matter of fact about consciousness that carries with it reward or punishment, and it is that which may not through some error transfer agents. Yet memory, as to be discussed further below, is indeed capable of error. Though Locke might not accept the possibility of an *implanted* memory, the thought experiment is useful for fleshing out the metaphysical criteria for persons

195

the following emendation, in order to make explicit that it is a metaphysical—not psychological—criterion:

1. **Consciousness Criterion***: The same consciousness must have inhered in the particular substances in which the relata of the person$_R$ terminate.

2. **Well-founded Criterion**: A person$_R$ must terminate in actually existent particular substance(s) as its truthmaker(s) in virtue of its relata.[213]

---

afforded by the relation-interpretation. For relevant alternative views on this passage, see Garrett (2003), LoLordo (2012), Stuart (2013a), and Weinberg (2016).

       Weinberg has recently argued that there is an "objective fact of an ongoing consciousness," which has a "temporal existence through any gaps in my successive, subjective states of awareness of myself," and which is "epistemically available from a third-personal (maybe only God's) point of view" (2016, 153). She calls on this "metaphysical fact . . . of my diachronic existence" to metaphysically ground personal identity, allowing it to occupy a *sui generis* ontological category outside what is commonly taken to be Locke's threefold ontology of substances, modes, and relations (2016, 153, 160–61). Weinberg's consciousness has an ongoing temporal existence, which has *duration* not only between one moment of conscious awareness to the next, but "even as measured between any gaps in my awareness of myself as thinking—say between the last idea I had prior to falling into a dreamless sleep and the first idea I had upon waking. And even though there are gaps in my awareness of my own continued existence, the full extent of my duration is known by God" (2016, 157). There are certain aspects of Weinberg's account that seem right, and that accord well with the relation-interpretation, but some qualifications must be made. Firstly, it seems unnecessary to posit a unique ontological category for consciousness. Although Locke is prone to agnosticism in the domain of substances, he relies on consciousness rather than a substantial identity condition precisely because we have intuitive access to it; it can afford us a judgment of which we can be "sure" (II.xxvii.23). This does not accord with the notion that consciousness is such an obscure entity that it demands a mysterious *sui generis* ontological category. Secondly, I do think that Weinberg is right to claim that just because there are gaps in awareness, that does not discount *some kind* of continuation, or duration, of the same consciousness. But, if we, as Weinberg seems to do, affirm that somehow (God only knows) consciousness enjoys a continued existence even during gaps in its operation, this threatens to endorse a picture that Locke regards as incoherent: conscious unconscious states (II.i.10). Simply because consciousness does not continue in an *uninterrupted* continuum does not mean that it does not in some sense continue. Consider other modes that are activities, like a song. I may pause in the middle of a song and restart it, and still understand it to be the same song. Much more can be said here, though due to the limitations of space, it lies beyond the scope of this dissertation.

[213] My thanks to Donald Ainslie for pushing me to modify the Well-founded Criterion in a way that more clearly distinguishes the work it is doing from that of Consciousness Criterion*.

Just as the Causal Criterion did for the relation <biological mother>, Consciousness Criterion* specifies the defining, requisite connection that must obtain between substances denominated by the person$_R$. This revised formulation of the twofold criteria excludes cases of mismatched false memories, in reliance on there being some objective fact of the matter concerning sameness of consciousness, such that the presentation of a seeming memory is not sufficient to infallibly determine whether or not Consciousness Criterion* obtains. The Well-founded Criterion guarantees that the substances denominated by a given relation in virtue of its relata are *actual*, existent substances, present or past, and Consciousness Criterion* guarantees that a purported person$_R$ does not simply denominate *any* particular substances in virtue of its relata, but that it denominates the *right* kind of particular substances that are capable of providing a foundation for the person$_R$.

2.3 Avoiding Circularity: Disentangling Consciousness and Memory and Their Psychological and Metaphysical Roles

The success of the *qualified* privileged access account just presented, including its capacity to exclude *all* varieties of pseudo-memories from constituting the person$_R$, demands an account of the metaphysical sameness of consciousness that is differentiated from memory. Sidestepping the circularity objection also calls for this distinction to be made clear. As I claimed earlier, the relation-interpretation provides the basis for differentiating between the (1) psychological criteria necessary for being in a position to make first-personal judgments about personal identity and (2) metaphysical criteria for person$_R$ that serve as truth conditions for judgments about personal identity. Consciousness

197

and memory are involved in *both* of these domains, so in order to demonstrate that the metaphysical account is *not* reducible to the psychological (and thus, that vicious circularity is *not* entailed), it is necessary to clarify the respective roles of consciousness and memory within these two domains.

What is it that distinguishes consciousness from memory? Although Locke has been charged by many interpreters, beginning with Butler and Reid, of conflating consciousness and memory, and while he does seem to use them interchangeably at times, nevertheless each plays a distinct role within his system. Let us first consider memory. Retention is one of the most salient faculties of the mind, making possible the sense of continuity that we have concerning the ideas that constitute the objects of our experience. As Locke defines it, retention is the "keeping of those simple *Ideas*, which from Sensation or Reflection it hath perceived" (II.x.1). Retention operates in two different ways: contemplation and memory, where contemplation maintains some previously perceived idea, and memory is "the Power to revive again in our Minds those *Ideas*, which after imprinting have disappeared, or have been as it were laid aside out of Sight" (II.x.2). Locke goes on to describe memory as follows:

> This is *Memory*, which is as it were the Store-house of our *Ideas*. . . But our *Ideas* being nothing but actual Perceptions in the Mind, which cease to be any thing, when there is no perception of them, this *laying up* of our *Ideas* in the Repository of the Memory, signifies no more but this, that the Mind has a Power, in many cases, to revive Perceptions, which it once had, with this additional Perception annexed to them, that it has had them before. And in this Sense it is, that our *Ideas* are said to be in our Memories, when indeed, they are actually no where, but only there is an ability in the Mind, when it will, to revive them again; and as it were paint them anew on it self, though some with more, some with less difficulty; some more lively, and others more obscurely. (II.x.2)

Memory, although described as a metaphorical storehouse of impressions of past perceptions, is actually simply a power, capacity, or operation of the mind, by which it re-

198

presents to the mind an idea, together with an "additional perception annexed" to it of a certain fact, viz. that one has previously perceived the revived idea.

This should be differentiated from consciousness, which Locke defines as simply "the perception of what passes in a Man's own mind" (I.i.19). Consciousness is, thus, a perception that invariably co-occurs with thought,[214] and that takes as its simultaneous percepts both the subject and object of whatever mental processes occur.[215] Although memory and consciousness are certainly interrelated, and while they both are described as involving reflexivity in some sense,[216] consciousness and memory are indeed distinct mental activities. Consciousness is a passive faculty, while memory is active (II.x.7). Consciousness witnesses, while memory revives. Memory, like Lockean knowledge, comes in degrees. Some previously perceived ideas fade, while others are longer lasting (II.x.4). The endurance and clarity of revived ideas depends on a number of factors, including their repetition, their being conjoined with attention, pleasure, pain, etc. (II.x.6). Locke also identifies two prevalent *defects* in the function of memory as oblivion and slowness (II.x.8). All this should be contrasted with consciousness, which is omnipresent whenever there is

---

[214] "[C]onsciousness always accompanies thinking"; it is "inseparable from thinking, and it seems to me essential to it" (II.xxvii.9). Locke even defines "thinking" in terms of consciousness: "thinking consists in being conscious that one thinks" (II.i.19).

[215] For a careful analysis of what it means for Lockean consciousness to perceive, see Lähteenmäki (2011), where he details how consciousness perceives simultaneously not only the subject and object *qua* ideas, but also the *act* of thinking. For the sake of simplicity, I will regard both the object/ideas and the actions of the mind to be subsumed under the "object-percept" of consciousness.

[216] As Garrett puts it, both memory and consciousness have a "Self-Representation Requirement" (2003, 100, 103). In the case of memory, the self must be represented as having been the past subject of some perception perceived again in the present. Consciousness too, being always reflexive, necessarily involves the representation of the self as the subject of any occurrent perception.

thinking, and cannot be slow or allow its objects to fall into oblivion. Consciousness is never "defective"; it is incapable of losing track of its percepts, the invariable possession of which defines it. "Sameness of *consciousness*," where consciousness is understood as the passive, witnessing, omnipresent mental operation of the simultaneous perception of both the subject and object of each thought, is *not* reducible to memory.

I have proposed that Consciousness Criterion* and the Well-founded Criterion might be understood as both (a) the metaphysical criteria for the person$_R$ and (b) the truth conditions for any given judgment of personal identity. They do not, however, fully explain the requisite psychological conditions for making such a first-personal judgment in the first place. As claimed earlier, seeming memory is a *necessary* psychological condition for making first-personal judgments of personal identity in virtue of its presenting candidate, purported relata with which we might identify. It does this by (at least purportedly) reviving a past perception of a self-conscious agent and an action. Consciousness too, being invariably co-occurrent with all forms of thinking, is a necessary psychological condition for such judgments in at least three ways: (a) The reflexivity of consciousness affords the judgment of the *synchronic* identity of the past and present self, which constitute relata of the person$_R$. (b) Consciousness witnesses the perception revived by the power of memory of a past self-conscious agent and action, while providing the reflexivity necessary for the awareness of the fact that the perception was previously perceived by the same subject, and is thus a (seeming) memory. (c) Consciousness also witnesses and provides the reflexivity necessary for the mental acts of *identification* with some past agent and the *appropriation* of some past action presented by the memory. The cooperation of these two faculties of memory and

200

consciousness on the occasion of making such a judgment of personal identity is indeed necessary, but it is *not* sufficient to guarantee the veridicality of such judgments. Nor can personal identity be reduced to, or said to *merely* consist in, the cooperation of these two mental activities, which would entail circularity.

The psychological roles of memory and consciousness in personal identity judgments must be differentiated from their respective manners of *metaphysically contributing* to the person. Seeming memory not only puts us in a position to judge personal identity, but it also plays a metaphysical role in constituting the person$_R$. Seeming memory both provides and also *delimits* the set of eligible, purported relata that may be assessed in respect of the two metaphysical criteria, viz. the Well-founded Criterion and Consciousness Criterion*. Given self-intimation, without the seeming memory of some purported relatum x, x is ineligible to constitute the person$_R$ in question. The metaphysical role of consciousness in constituting the person$_R$ is expressed by Consciousness Criterion*, which says that the sameness of consciousness understood as an objective, metaphysical fact is the unique criterion for the person$_R$.

Conclusion

A purely psychological interpretation of Locke's account of personal identity, as exemplified by the Memory Theories, poses a dilemma: either Locke's account is viciously circular on the Veridical Memory Theory or he must accept false memories as also constituting the person on the Seeming Memory Theory. I have argued that a metaphysical account of persons is necessary to allow for the appearance-reality distinction required to

arbitrate between veridical and false memories and to provide truth conditions for our personal identity judgments that are non-circular. Among the available ontological options, I have argued that a *relation*-interpretation of the Lockean person provides a uniquely supportive interpretive basis for resolving the apparent dilemma posed by the Memory Theories.

The fact that persons *qua* relations are mind-dependent safeguards Locke's strong Privileged Access Intuition, while also according well with Locke's non-substantialist account of personal identity. Yet the fact that persons *qua* relations are beholden to the existence of particular substances as their truthmakers affords a basis for the *qualification* of privileged access, accommodating the Fallibility Intuition and excluding false memories from constituting the person$_R$. The relation-interpretation achieves this by providing, what I have termed the Well-founded Criterion and Consciousness Criterion*, which serve as the jointly sufficient metaphysical criteria and truth conditions guaranteeing that a purported person$_R$ is actual.

Finally, through disambiguating the roles of consciousness and memory in the psychological process of how we make personal identity judgments and the metaphysical domain of how the person *qua* relation is constituted, I have clarified that the resulting account is non-circular. In sum, understanding the Lockean person as a relation provides an interpretive framework that not only affords promising new strategies for defending his account against historically intransigent objections but that also yields a coherent picture of personhood that supports the guiding forensic intuitions motivating Locke's influential theory of personal identity.

PART II BIBLIOGRAPHY

Alston, William, and Jonathan Bennett. 1988. "Locke on People and Substances." *Philosophical Review* 97: 25–46.

Atherton, Margaret. 1983. "Locke's Theory of Personal Identity." *Midwest Studies in Philosophy* 8: 273–93.

Baker, Lynn Rudder. 2000. *Persons and Bodies: A Constitution View*. Cambridge: Cambridge University Press.

Bennett, Jonathan. 1971. *Locke, Berkeley, Hume: Central Themes*. Oxford: Clarendon Press.

Boeker, Ruth. 2014. "The Moral Dimension in Locke's Account of Persons and Personal Identity." *History of Philosophy Quarterly* 31, no. 3: 229–47.

———. 2015. "Locke and Hume on Personal Identity: Moral and Religious Differences." *Hume Studies* 41, no. 2: 105–35.

———. 2016. "The Role of Appropriation in Locke's Account of Persons and Personal Identity." *Locke Studies* 16: 3–39.

———. 2017. "Locke on Personal Identity: A Response to the Problems of His Predecessors." *Journal of the History of Philosophy* 55, no. 3: 407–34.

Bolton, Martha Brandt. 1994. "Locke on Identity: The Scheme of Simple and Compounded Things." In *Individuation and Identity in Early Modern Philosophy: Descartes to Kant*, edited by Kenneth F. Barber and Jorge J. E. Gracia, 103–31. Albany: State University of New York Press.

———. 2007. "The Taxonomy of Ideas in Locke's *Essay*." In *The Cambridge Companion to Locke's "Essay Concerning Human Understanding,"* edited by Lex Newman, 67–100. Cambridge: Cambridge University Press.

Brody, Baruch. 1972. "Locke on the Identity of Persons." *American Philosophical Quarterly* 9, no. 4: 327–34.

Butler, Joseph. 1736. "Of Personal Identity." In *The Analogy of Religion*, reprinted in Perry (1975, 99–105).

Chappell, Vere. 1989. "Locke and Relative Identity." *History of Philosophy Quarterly* 6, no. 1: 69–83.

———. 1990. "Locke on the Ontology of Matter, Living Things and Persons." *Philosophical Studies* 60: 19–32.

Curley, Edwin. 1982. "Leibniz on Locke on Personal Identity." In *Leibniz: Critical and Interpretive Essays*, edited by Michael Hooker, 302–326. Manchester: University of Minnesota Press.

Garrett, Don. 2003. "Locke on Personal Identity, Consciousness, and 'Fatal Errors.'" *Philosophical Topics* 31, nos. 1–2: 95–125.

Gertler, Brie. 2003. "Philosophical Issues about Self-Knowledge." In *Privileged Access: Philosophical Accounts of Self-Knowledge*, edited by Brie Gertler, xi–xxii. Burlington, VT: Ashgate.

———. 2011. *Self-Knowledge*. New York: Routledge.

Gibson, James. 1917. *Locke's Theory of Knowledge and Its Historical Relations*. Cambridge: Cambridge University Press.

Gordon-Roth, Jessica. 2015. "Locke on the Ontology of Persons." *The Southern Journal of Philosophy* 53, no. 1: 97–123.

Green, Thomas Hill. 1885. *The Works of Thomas Hill Green*. Vol. 1. Edited by R. L. Nettleship. London: Longmans, Green and Co.

Kaufman, Dan. 2007. "Locke on Individuation and the Corpuscular Basis of Kinds." *Philosophy and Phenomenological Research* 75, no. 3: 499–534.

Kierkegaard, Søren. *The Sickness Unto Death: A Christian Psychological Exposition for Upbuilding and Awakening*, edited and translated by Howard V. Hong and Edna H. Hong. Princeton, NJ: Princeton University Press, 1980.

Langton, Rae. 2000. "Locke's Relations and God's Good Pleasure." *Proceedings of the Aristotelian Society* 100: 75–91.

Law, Edmund. 1769. *A Defence of Mr. Locke's Opinion Concerning Personal Identity; in Answer to the First Part of a Late Essay on That Subject*. Cambridge.

Lähteenmäki, Vili. 2011. "Locke on Consciousness and What It Is About." *Studia Leibnitiana* 43, no. 2: 160–78.

———. 2018. "Locke and Metaphysics of 'State of Sensibility.'" In *Philosophy of Mind in the Early Modern Age and in the Enlightenment*, vol. 4 of *History of the Philosophy of Mind*, edited by Rebecca Copenhaver and Christopher Shields, 157–173. New York: Routledge.

Locke, John. 1975. *An Essay Concerning Human Understanding*. Edited by Peter Nidditch. Oxford: Oxford University Press.

LoLordo, Antonia. 2010. "Person, Substance, Mode and 'the *Moral Man*' in Locke's Philosophy." *Canadian Journal of Philosophy* 40: 643–68.

———. 2012. *Locke's Moral Man*. Oxford: Oxford University Press.

Mattern, Ruth. 1980. "Moral Science and the Concept of Persons in Locke." *The Philosophical Review* 89: 24–45.

Newman, Lex. 2015. "Locke on Substance, Consciousness, and Personal Identity." In *Locke and Leibniz on Substance*, edited by Paul Lodge and Tom Stoneham, 89–109. New York: Routledge.

Odegard, Douglas. 1969. "Locke and the Unreality of Relations." *Theoria* 35: 147–52.

Ott, Walter. 2009. *Causation and Laws of Nature in Early Modern Philosophy*. Oxford: Oxford University Press.

———. 2017. "Archetypes without Patterns: Locke on Relations and Mixed Modes." *Archiv für Geschichte der Philosophie* 99, no. 3: 300–325.

Pasnau, Robert. 2011. *Metaphysical Themes 1274–1671*. Oxford: Oxford University Press.

Perry, John, ed. 1975. *Personal Identity*. Berkeley: University of California Press.

Reid, Thomas. 1785. "Of Memory." In *Essays on the Intellectual Powers of Man*, reprinted in Perry (1975) as "Of Identity" (107–12) and "Of Mr. Locke's Account of Our Personal Identity" (113–18).

Rickless, Samuel C. 2015. "Are Locke's Persons Modes or Substances?" In *Locke and Leibniz on Substance*, edited by Paul Lodge and Tom Stoneham, 111–27. New York: Routledge.

———. 2018. "Locke's Ontology of Relations." *Locke Studies* 17: 61–86.

Shoemaker, Sydney. 1963. *Self-Knowledge and Self-Identity*. Ithaca, NY: Cornell University Press.

———. 1984. *Personal Identity*. Oxford: Blackwell.

———. 1996. *The First Person Perspective and Other Essays*. Cambridge: Cambridge University Press.

Simendic, Marko. 2015. "Locke's Person Is a Relation." *Locke Studies* 15: 79–97.

Strawson, Galen. 2011. *Locke on Personal Identity: Consciousness and Concernment*. Princeton, NJ: Princeton University Press.

Stuart, Matthew. 2013a. *Locke's Metaphysics*. Oxford: Oxford University Press.

———. 2013b. "Revisiting People and Substances." In *Debates in Modern Philosophy*, edited by Stewart Duncan and Antonia LoLordo, 186–96. New York: Routledge.

Tabb, Kathryn. 2018, "Madness as Method: On Locke's Thought Experiments about Personal Identity." *British Journal for the History of Philosophy* 26, no. 5: 871–89.

Thiel, Udo. 1981. "Locke's Concept of Person." In *John Locke: Symposium Wolfenbüttel 1979*, edited by Reinhard Brandt, 181–92. Berlin: de Gruyter.

———. 1998. "Personal Identity." In *The Cambridge History of Seventeenth-Century Philosophy*, vol. 1, edited by Daniel Garber and Michael Ayers, 868–912. Cambridge: Cambridge University Press.

———. 2006. "Self-Consciousness and Personal Identity." In *The Cambridge History of Eighteenth-Century Philosophy*, vol. 1, edited by Knud Haakonssen, 286–318. Cambridge: Cambridge University Press.

———. 2011. *The Early Modern Subject: Self-Consciousness and Personal Identity from Descartes to Hume.* Oxford: Oxford University Press.

Uzgalis, William. 1988. "The Anti-Essential Locke and Natural Kinds." *The Philosophical Quarterly* 38, no. 152: 330–39.

———. 1990. "Relative Identity and Locke's Principle of Individuation." *History of Philosophy Quarterly* 7: 283–97.

———. 2008. "Locke and Collins, Clarke and Butler, on Successive Persons." In *Personal Identity*, 2nd ed., edited by John Perry, 315–326. Berkeley: University of California Press.

Weinberg, Shelley. 2008. "The Coherence of Consciousness in Locke's *Essay*." *History of Philosophy Quarterly* 25, no. 1: 21–39.

———. 2012a. *Consciousness in Locke*. Oxford: Oxford University Press.

———. 2012b. "The Metaphysical Fact of Consciousness in Locke's Theory of Personal Identity." *Journal of the History of Philosophy* 50, no. 3: 387–415.

———. 2016. *Consciousness in Locke*. Oxford: Oxford University Press.

Wilkes, Kathleen. 1988. *Real People: Personal Identity without Thought Experiments*. Oxford: Oxford University Press.

Wilson, Fred. 1994. "Substance and Self in Locke and Hume." In *Individuation and Identity in Early Modern Philosophy*, edited by Kenneth F. Barber and Jorge J. E. Gracia, 155–99. Albany: State University of New York Press.

Winkler, Kenneth P. 1991. "Locke on Personal Identity." *Journal of the History of Philosophy* 29: 201–26.

Yaffee, Gideon. 2011. "Locke on Consciousness, Personal Identity and the Idea of Duration." *Noûs* 45, no. 3: 387–408.