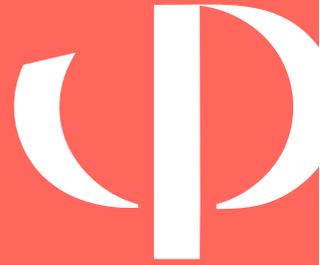


APA Newsletters



FALL 2019

VOLUME 19 | NUMBER 1

ASIAN AND ASIAN-AMERICAN PHILOSOPHERS AND PHILOSOPHIES

FEMINISM AND PHILOSOPHY

HISPANIC/LATINO ISSUES IN PHILOSOPHY

LGBTQ ISSUES IN PHILOSOPHY

PHILOSOPHY AND THE BLACK EXPERIENCE

PHILOSOPHY AND COMPUTERS

Table of Contents

Asian and Asian-American Philosophers and Philosophies	1	What Is Rape? Social Theory and Conceptual Analysis	78
Buddhist Philosophy Today: Theories and Forms	1	Debating Pornography.....	80
Philosophy, Quo Vadis? Buddhism and the Academic Study of Philosophy	3	Foucault and Feminist Philosophy of Disability	84
What/Who Determines the Value of Buddhist Philosophy in Modern Academia?.....	8	Hispanic/Latino Issues in.....	90
Buddhist Philosophy? Arguments from Somewhere	11	Latinx Philosophy Conference	91
Doing Buddhist Philosophy	15	Dying in Detention as an Example of Oppression ..	91
Decolonizing the Buddhist Mind	22	A Negative Path Towards Anti-Racist Immigration Policy	97
Reflecting on Buddhist Philosophy with Pierre Hadot	26	Immigration and International Justice	102
Some Suggestions for Future Directions of the Study of Buddhist Philosophy.....	28	Are Our Racial Concepts Necessarily Essentialist Due to Our Cognitive Nature?	108
Practicing Buddhist Philosophy as Philosophy	31	Epistemic Humility Now!.....	113
Emptiness, Multiverses, and the Conception of a Multi-Entry Philosophy.....	34	Approaching Racial Embodiment, Aesthetics, and Liberation in José Carlos Mariátegui's Seven Essays.....	118
Buddhist Philosophy and the Neuroscientific Study of Meditation: Critical Reflections.....	36	América Tropical, On the Force of Latino/a/x Thought	123
Feminism and Philosophy.....	41	The Marxism of José Revueltas: A Struggle Against Orthodoxy	127
#MeToo and Philosophy	41	Thinking about Exile: Community, Violence, and Law.....	132
Good Survivor, Bad Survivor: #MeToo and the Moralization of Survivorship	45	Book Excerpt: Marx on Bolívar	137
Beyond Silence, Towards Refusal: The Epistemic Possibilities of #MeToo.....	52	Rape and Resistance	138
The Speech Acts of #MeToo	56	LGBTQ Issues in Philosophy	142
#MeToo?	60	Assumptive Care and Futurebound Care in Trans Literature	143
#MeToo vs. Mea Culpa: On the Risks of Public Apologies	64	Participation in Sport Is a Human Right, Even for Trans Women	151
Women, Work, and Power: Envisaging the Radical Potential of #MeToo	69	Conceiving Differently within the Ethics of Assisted Reproduction	155
Field Notes on Conference Climate: A Decade with the Philosophy of Science Association Women's Caucus.....	76	Life and Death of Latisha King: A Critical Phenomenology of Transphobia	158

Table of Contents, continued

Philosophy and the Black Experience 162

Charles Leander Hill (1906–1956)..... 162

Freedom Gaze: Explicating the African Freedom
Aesthetic 164

Philosophy and Computers..... 169

How Many Thoughts Can Fit in the Form of a
Proposition? 172

Consciousness, Engineering, and
Anthropomorphism..... 180

Should Physical Computation Be Understood
Mechanistically?..... 186

Commentary on Gualtiero Piccinini’s Physical
Computation: A Mechanistic Perspective 189

Defending the Mapping Account of Physical
Computation 192

Comments on Gualtiero Piccinini, *Physical
Computation: A Mechanistic Account*..... 193

The Mechanistic Account of Physical
Computation: Some Clarifications..... 195

Philosophical Insights from Computational
Studies: Why Should Computational Thinking
Matter to Philosophers?..... 200

From the Editor 207

DABUS in a Nutshell..... 208

Five Sessions Organized by the APA Committee
on Philosophy and Computers for the 2020 APA
Divisional Meetings 211



APA NEWSLETTER ON

Asian and Asian-American Philosophers and Philosophies

RAFAL STEPIEN, GUEST EDITOR

VOLUME 19 | NUMBER 1 | FALL 2019

FROM THE GUEST EDITOR

Buddhist Philosophy Today: Theories and Forms

Rafal Stepien

NANYANG TECHNOLOGICAL UNIVERSITY, SINGAPORE

This is the second of two special issues of the newsletter dedicated to Buddhist philosophy.¹ My initial intention as guest editor was to prepare a single issue of the newsletter on the topic "Buddhist Philosophy Worldwide: Perspectives and Programs." The idea was to include descriptive and prescriptive/evaluative elements: On the one hand, scholars working on Buddhist philosophy throughout the world were invited to provide a descriptive snapshot of the state of the field in their geographical/disciplinary area; on the other, they could proffer an evaluative appraisal of how Buddhist philosophy has been carried out and/or a prescriptive program of how they feel it should be carried out. This dual remit played out in a foreseeable manner, such that some authors composed largely descriptive pieces, while others took a more methodologically oriented approach in which they outline a vision of what the practice of Buddhist philosophy could or should entail, and/or how it can or could contribute to the practice of academic philosophy *per se*.

Eventually, for both practical and programmatic reasons, the decision was taken to unweave these strands into two separate newsletter issues, with the previous spring 2019 issue remaining devoted to "Buddhist Philosophy Worldwide: Perspectives and Programs," and the current fall 2019 one on "Buddhist Philosophy Today: Theories and Forms." Practically, the total length of the articles submitted by the twenty authors I was able to corral greatly exceeded that typical for a single issue of the newsletter, and the subsequent realization that roughly half of the authors had taken each of the two tracks I had laid led me and the APA to decide upon dividing the articles accordingly. More substantively, upon reading the final products it became clear to me that we were dealing here with two distinct and individually important sets of contributions to the study of Buddhist philosophy. On the one hand, given that the more descriptive articles preponderantly issued from non-Western cultural/national contexts underrepresented within the field at large, and given also that the descriptions provided by these authors were typically accompanied by healthy doses of interpretation, I consider these contributions to constitute a solid bloc of scholarship on

the practice of Buddhist philosophy worldwide. On the other hand, those contributions whose authors took a more evaluative or prescriptive approach likewise taken together comprise a well-rounded collection of articles, in this case one theorizing contemporary Buddhist philosophical scholarship and the future directions it may take.

In preparing the collection as a whole, I was particularly resolute that contributions cover a greater geographical span than that encompassed by the major centers in Europe and North America. For the foregoing survey of "Buddhist Philosophy Worldwide," my insistence on a broad geographical coverage was motivated on the one hand by a methodological impetus to ensure as comprehensive as possible a spectrum of perspectives be included, and on the other hand by the conviction that Buddhist philosophy, being a strikingly multi- and trans-cultural phenomenon itself, could and should be studied, carried out, and put into practice most fruitfully from the widest possible range of vantage points. As such, I actively sought out contributors from a variety of countries in Asia, where Buddhist philosophy has, of course, the longest of intellectual pedigrees, as well as Australasia, Africa, South America, and the Middle East in addition to Europe and North America. Unfortunately, I was unable to locate any scholars based anywhere in Africa, South America, or the Middle East outside of Israel willing to take part.

Interestingly, it so happens that in almost all cases scholars working in European and North American universities where the field's center of gravity lies chose to concentrate on theoretical elaborations of Buddhist philosophical practice; their contributions thus appear in the present issue. Of course, the relatively limited geographical span within which the contributors gathered here work (if a span including Singapore, Taiwan, Thailand, Austria, France, Iceland, the United Kingdom, as well as several North American institutions can be called "limited") has not led to any lack of diversity among the intellectual perspectives expounded in the pages that follow. On the contrary, the present volume includes what I believe is a hitherto unparalleled collection of texts not only detailing and appraising the general state of the scholarly field of Buddhist philosophy today but also proposing ways in which it can flourish further into the future. Brook Ziporyn provides a fitting start to this endeavor, as his "Philosophy, Quo Vadis? Buddhism and the Academic Study of Philosophy" moves from consideration of whether and how Buddhist thinkers could get to use the brand name "philosophy" to a provocative interpretation of Buddhist philosophy as uniquely instantiating the project of radical doubt lying at the threshold of modern Western philosophy. Hans-Rudolf

Kantor's interrogation as to "What/Who Determines the Value of Buddhist Philosophy in Modern Academia?" continues probing the question of the disciplinary status of Buddhist philosophy. This article opens with a sweeping, and not uncontroversial, analysis of the field as practiced across the East-West divide, which Kantor then uses to propose a theoretical distinction between "philosophy in Buddhism" and "Buddhist philosophy." My own contribution, "Buddhist Philosophy? Arguments From Somewhere," continues this line of questioning as to the place of Buddhist philosophy in today's academe, in this case by assembling and critiquing the arguments standardly mobilized to exclude it along with all other non-Western systems of thought. In "Doing Buddhist Philosophy," C. W. Huntington, Jr., then theorizes the field from the perspective of the divergent means and ends of approaches to Buddhist philosophy that foreground reason and logic on the one hand and soteriologically oriented wisdom on the other. Mattia Salvini's account of "Decolonizing the Buddhist Mind" moves along complementary lines to investigate the nature of the institutional space wherein Buddhist philosophy could be the center and life-force of one's enquiry instead of a merely peripheral analytical object. In "Reflecting on Buddhist Philosophy with Pierre Hadot," Matthew T. Kapstein focuses on what he calls "Dharmakīrtian spiritual exercise" to explore the very question of what counts as philosophical progress, and thereby provide a useful pivot toward more explicit discussion of the future of Buddhist philosophy. Jan Westerhoff thus presents "Some Suggestions for Future Directions of the Study of Buddhist Philosophy," which he categorizes under the rubrics of "Editions and Translations," "Integrated Textual and Conceptual Presentation," and "Linkage with Contemporary Philosophical Discussion." In "Practicing Buddhist Philosophy as Philosophy," Pierre-Julien Harter likewise considers the next steps to take so that Buddhist philosophy may evolve from the stage of the recovery of texts and ideas to the stage of participation in the conversation of world philosophy. Gereon Kopf then adumbrates one specific manner in which such participation could take place. His "Emptiness, Multiverses, and the Conception of a Multi-Entry Philosophy" draws on classical Indian, Chinese, and Japanese Mahāyāna Buddhist teachings in conversation with the twentieth-century thought of Gilles Deleuze, Félix Guattari, and Francois Lyotard to propose Buddhism as pre-eminently endowed to forge a new mode of philosophical inquiry according to which not one but multiple equally persuasive and insightful ways of answering the same question may be countenanced. Finally, Birgit Kellner's "Buddhist Philosophy and the Neuroscientific Study of Meditation: Critical Reflections" clarifies some of the points of tension between Buddhist philosophy and the contemplative neurosciences, and thereby seeks to clear the ground for more nuanced future work.

As may transpire from the foregoing account, I have structured this volume in a manner that self-consciously works against any easy compartmentalizations of academic Buddhist philosophy along geographical and/or cultural lines (e.g., South-Asian/East-Asian, Indo-Tibetan/Sino-Japanese, etc.). Instead, and in accordance with the mandate of this special issue, I have foregrounded those pieces which provide generalized accounts of and

responses to the disciplinary status of Buddhist philosophy, before moving to pieces geared more toward the future directions, general and specific, it could or should take. One abiding regret I have to do with the assembled pieces regards the gender representation of the authors, for only three of eleven contributors to the previous issue and only one of ten in the present one are female. This imbalance I readily recognize as problematic, though I can assure the readership that it remains not for any lack of trying to avert or rectify it: In addition to those who did agree to contribute, I invited a further eight female scholars of Buddhism who for various reasons were unable to commit to this project. Had they been able to do so (and I am not trying to make anyone feel guilty!), a more-or-less equal representation of genders would have been ensured; one, it merits mentioning, well in excess of the stubbornly skewed levels of representation in the field (of Buddhist philosophy, to say nothing of philosophy itself) as a whole.

My thanks go first of all to the previous editor of the newsletter, Prasanta Bandyopadhyay, for inviting me to act as guest editor, to the chair of the Committee on Asian and Asian-American Philosophers and Philosophies, Brian Bruya, for supporting my suggestion as to the topic, to Erin Shepherd for her superb skills in coordinating publication, and to my anonymous peer-reviewer for not only agreeing to be involved but for producing such fine reviews at such a speedy rate. I also express my gratitude to the Berggruen Philosophy & Culture Center for funding that enabled initiation of this work while I was the Berggruen Research Fellow in Indian Philosophy at Wolfson College and the Faculty of Philosophy of the University of Oxford, to the Alexander von Humboldt Foundation for funding that enabled completion of this work while I was a Humboldt Research Fellow at the Karl Jaspers Centre for Advanced Transcultural Studies of Heidelberg University, and to the Studies in Inter-Religious Relations in Plural Societies Programme of Nanyang Technological University in Singapore for providing a fitting setting for me, as the incoming Assistant Professor in Comparative Religion, in which to finalize publication of this work. At Oxford and Heidelberg, Richard Sorabji, Jan Westerhoff, and Michael Radich stand out as colleagues and mentors especially supportive of this and like projects in and of Buddhist philosophy. Of course, I reserve my most profound thanks to the contributors themselves, without whose energy and insight none of this could have come to fruition.

NOTES

1. Apart from the summary of contributions comprising this volume, this introduction reproduces (with but minor alterations) that of the preceding volume, on the understanding that this gives readers of each volume access to the overall thrust of both special issues.

SUBMISSION GUIDELINES AND INFORMATION

GOAL OF THE NEWSLETTER ON ASIAN AND ASIAN-AMERICAN PHILOSOPHERS

The *APA Newsletter on Asian and Asian-American Philosophers and Philosophies* is sponsored by the APA Committee on Asian and Asian-American Philosophers and Philosophies to report on the philosophical work of Asian and Asian-American philosophy, to report on new work in Asian philosophy, and to provide a forum for the discussion of topics of importance to Asian and Asian-American philosophers and those engaged with Asian and Asian-American philosophy. We encourage a diversity of views and topics within this broad rubric. None of the varied philosophical views provided by authors of newsletter articles necessarily represents the views of any or all the members of the Committee on Asian and Asian-American Philosophers and Philosophies, including the editor(s) of the newsletter. The committee and the newsletter are committed to advancing Asian and Asian-American philosophical scholarships and bringing this work and this community to the attention of the larger philosophical community; we do not endorse any particular approach to Asian or Asian-American philosophy.

SUBMISSION GUIDELINES

1) **Purpose:** The purpose of the newsletter is to publish information about the status of Asians and Asian Americans and their philosophy and to make the resources of Asians and Asian-American philosophy available to a larger philosophical community. The newsletter presents discussions of recent developments in Asians and Asian-American philosophy (including, for example, both modern and classical East-Asian philosophy, both modern and classical South Asian philosophy, and Asians and Asian Americans doing philosophy in its various forms), related work in other disciplines, literature overviews, reviews of the discipline as a whole, timely book reviews, and suggestions for both spreading and improving the teaching of Asian philosophy in the current curriculum. It also informs the profession about the work of the APA Committee on Asian and Asian-American Philosophers and Philosophies. One way the dissemination of knowledge of the relevant areas occurs is by holding highly visible, interactive sessions on Asian philosophy at the American Philosophical Association's three annual divisional meetings. Potential authors should follow the submission guidelines below:

- i) Please submit essays electronically to the editor(s). Articles submitted to the newsletter should be limited to ten double-spaced pages and must follow the APA submission guidelines.
- ii) All manuscripts should be prepared for anonymous review. Each submission shall be sent to two referees. Reports will be shared with authors. References should follow *The Chicago Manual Style*.

iii) If the paper is accepted, each author is required to sign a copyright transfer form, available on the APA website, prior to publication.

- 2) **Book reviews and reviewers:** If you have published a book that you consider appropriate for review in the newsletter, please ask your publisher to send the editor(s) a copy of your book. Each call for papers may also include a list of books for possible review. To volunteer to review books (or some specific book), kindly send the editor(s) a CV and letter of interest mentioning your areas of research and teaching.
- 3) **Where to send papers/reviews:** Please send all articles, comments, reviews, suggestions, books, and other communications to the editor: Minh Nguyen (atnguyen@fgcu.edu).
- 4) **Submission deadlines:** Submissions for spring issues are due by the preceding November 1, and submissions for fall issues are due by the preceding February 1.
- 5) **Guest editorship:** It is possible that one or more members of the Committee on Asian and Asian American Philosophers and Philosophies could act as guest editors for one of the issues of the newsletter depending on their expertise in the field. To produce a high-quality newsletter, one of the co-editors could even come from outside the members of the committee depending on his/her area of research interest.

ARTICLES

Philosophy, Quo Vadis? Buddhism and the Academic Study of Philosophy

Brook Ziporyn
UNIVERSITY OF CHICAGO

For quite a while now “philosophy” seems to have been undergoing something like what “Pink Floyd” underwent in the late 1980s: a viciously consequential branding dispute. For those of you who don’t obsess over prog rock history, what happened there was that when bassist, lyricist, main writer, and *de facto* creative director Roger Waters left the band around the middle of that decade, the remaining members—David Gilmour, Richard Wright, and Nick Mason—wished to continue using the copyrighted name “Pink Floyd” for their future musical projects. Waters objected, alternately claiming the name for himself on the grounds of his dominant role during the period of the band’s greatest successes, or else moving to have the name retired altogether, asserting that Pink Floyd as such was already by that time “a spent force creatively.” Years of litigation ensued. And we can understand why: a lot of money, and a lot of continuity and prestige and cultural attention and fanbase connection, were at stake. The words “Pink Floyd” were a brand, a recognizable name, which automatically brought with it millions of fans and a certain cultural position and resonance. “Pink Floyd” was a

bank account full of cultural capital. A concert played with hired backup musicians by Waters under his own name, recognized only by aficionados, and the exact same concert with the exact same band labeled "Pink Floyd," would have hugely different consequences, ticket sales, cultural impact. A Gilmour-Wright-Mason concert—under the name, say, GWM or The Space Cadets—could be expected to do a lot less well than the megaconcerts of the new Watersless Pink Floyd in the 1990s in fact did do. It didn't matter that more than half of the material played at those Watersless Pink Floyd concerts was penned by Waters, and reflected his own very personal and specific obsessions. What mattered culturally was who got to use the precious name. (For the record, I take no sides in the Pink Floyd branding dispute. May Gilmour and Waters both live and flourish.)

And something similar seems to be the case with the word "philosophy." Whatever account you favor for what defines the analytic/continental divide in philosophy, and whatever theory you might embrace about its causes, it is clear that "creative differences" have made it impossible for these bandmates to work together. The incompatibility has reached the point of breakup, whether for institutional or substantive reasons. It would be folly, I think, to try to adjudicate which party has a more legitimate claim to the brand name "philosophy" on the basis of historical precedent or traditional usage. Indeed, if we look to the history of the word, we find ourselves the more bemused as we descend into the Syd Barrett-era-esque morass of natural scientists and sage-figures who were among the undisputed claimants of the term in centuries past. The cynical, probably correct, view is that this is really about lucrative jobs at prestigious institutions, and the cultural cachet of the word, just as in the Pink Floyd case. But in any case, it is now indisputable that somehow, in Anglophone institutions of higher learning and academic presses, what "philosophy" has come to mean is a cluster of practices oriented toward certain forms of inquiry, methodologies, and areas of concern that most of us would not hesitate to call unambiguously "analytic." To those who do not participate in these practices, or perhaps have not been trained in them or are simply not good at them, they are easy to despise: they can appear to be a tragic narrowing of what philosophy meant in the good old heydays, a professionalized form of nitpicking designed for maximal convenience of professional assessment, or even a vipers' nest of mediocrity abdicating the glorious high calling of the philosophical demi-gods of old in favor of the rising dominance of a very different kind of human being, the plodding antihero—the reasonable jigsaw puzzle hobbyist rather than the raving tortured Prometheus. If the old philosophers were one-man bands, simultaneously playing bass drum with one foot and cymbal with another, a harmonica wired around the neck and a three-necked guitar strapped on the back and an accordion under the arm—ontology, metaphysics, epistemology, logic, and ethics all jumbled together and careening onward at once in a messy cacophony of virtuoso pyrotechnics—these new philosophers look like a man doing physical therapy after a serious car accident, relearning step by minute step how to use the fine muscles in his fingers so that he can one day again hold a pen: now bend the index finger at the first joint, now the second, now bring in the thumb, good,

good, now slowly lower it toward the paper. . . . To a neutral and uninformed observer this may indeed look pretty ridiculous. But it has its function, doesn't it? Some people need just this physical therapy, and it is important to relearn to do so, so it must be good that there is a method, a well-tested and effective and responsible method, to do it. I don't myself participate in these practices, and I am certainly (obviously) not immune to this temptation to disparage them, particularly keen when forced to witness the painful spectacle of well-meaning and often brilliant young students, afire with a passion for "philosophy" inspired by random and naturally quite superficial high school readings, having the gumption knocked out of them by their first "philosophy" class in a university: *this* is what philosophy is supposed to be? From there it's sink or swim: either they learn this new jargon and mode of procedure—the *responsible* way, the *conscientious* way, the *careful* way—or they change majors. The waste and moral destruction of these high-spirited talents is a subject almost worthy of a classical tragedy. As "thinkers," let's say, rather than the contested "philosophers," one often has the impression that these naïve and unbridled young minds are clearly superior to—or less contentiously, at least more interesting than—the professors who upbraid and regulate and re-educate them, who take it as their duty to whip them into line; it is not a pleasant thing to have to see. As a non-practitioner, though, and thus as unskilled labor in this realm, I not only must keep this sentiment to myself, but I must admit that it is probably unjustified. I am in no position to judge this set of practices, and could and should rationally give it the benefit of the doubt: probably they are doing something worthwhile, to someone, over there. But it remains a problem, though not one that can be readily blamed on any single set of bad actors, that they have won the branding battle and monopolized the holy and powerful name "Pink Floyd"—I mean "philosophy."

One of the effects of this outcome is the fate of prospective new members of the guild. Who gets to play with this band? Who gets to be part of the new lineup? On our analogy, perhaps the answer is "no one": the maximal membership of Pink Floyd is the four members during the 1973–1981 period of the band's greatest successes, when the brand acquired its market value. Therefore, the prestige follows either all four of them, or some subset of them; anyone else who plays on albums or concerts is a hired gun, a studio or road musician in the role of independent contractor. They can later go out and form their own bands, perhaps parlaying the prestige of having played with "Pink Floyd" for so long. Is this the case for the barbarians at the gate of philosophy? I mean, of course, those forms of thinking that are neither "analytic" nor "continental," that come from neither part of the contested European terrain. I mean non-European thinkers and traditions—my personal concerns lie mainly in Chinese traditions, i.e., Confucianism, Daoism, and Buddhism, but the same question can be asked of traditions of thought originating in India or Africa or anywhere else. Do they get to use the brand name "philosophy"? Let's take the case of Buddhism: Should there be more of it taught in philosophy departments? Should Buddhist thinkers get to use the brand name "philosophy"?

Some try to adduce a criterion that would serve as a shibboleth for membership in the guild: as if to say, if you can play the bassline of “Money” the way Roger did, perhaps, you can be the new Pink Floyd bassist. Analogously, if you do not appeal to authority, with the exception of the absolute authority of reason as defined by Aristotle and his legitimate successors, if you offer evidence and arguments for your assertions and can defend them in the agreed-upon format of logical dispute, and do it exceptionally well, then you can be granted admission to the guild, irrespective of the content of the claims that you are arguing for or against. As long as you hold to the accepted method, all claims are welcome. This is already, I think, a quite generous and open-minded sentiment on the part of the gatekeepers, and one that should be applauded. To put it another way: Should the defining and identifying and qualifying characteristic of philosophy be the scope of *subject matters* with which it has, until quite recently, mainly concerned itself—what exists, what does it mean, what can and should we do, what can and should we know?—or with the *methods* by which answers to those questions were sought: not revelations from the gods, not fiat from a prophet, not catchy aphorisms from a charismatic sage, not poetic riffs from an inspired improviser, but reasoned argumentation? Much in these foreign traditions arguably addresses the areas that have been of greatest concern to many European philosophers of the past. But the method of presentation and establishment of their claims may not always accord in any obvious way with the accepted philosophical methods. Nor, however, do they seem to align easily with the foil of the philosophical methods in Europe, the anti-philosophical authoritarianism of monotheistic religious revelation. We begin to see a deeper problem here, in that the accepted methods embraced by European philosophies have themselves been forged on certain assumptions and presuppositions, not to mention the specific contours of the historical European case, particularly its highly abrasive love-hate relationship to its highly abrasive religion. For myself, I think the really interesting work to be done lies precisely here, in excavating those presuppositions and exploring the alternatives, which may also open up new methodological vistas that fit neither the philosophical nor the non-philosophical as generally understood in Europe. But it is not so unreasonable for the guild to want to reserve its imprimatur for the primacy of philosophical method as previously understood, given the historical divisions of labor that obtained in the context of European intellectual history, as the criterion for what gets to use the brand name philosophy. And on either grounds, in both method and content, Buddhist traditions certainly have much to offer, much that is recognizably within the fold of philosophical method understood in this relatively narrow way. This is not all Buddhism is, and in some ways, for me personally, it is not what is most interesting or intellectually exciting about Buddhist traditions, nor for that matter what is most intellectually thoroughgoing and rigorous in their thinking through of the reconfiguration of premises. This means that their full potential to dialogue and interbreed and cross-germinate with European thought, to generate new ways of thinking, and to shake up and/or invigorate European thought, is going to be severely crippled by the application of this (reasonable) criterion, and we’re going to end up with a Buddhist philosophy that is constrained to playing

by an alien set of rules. But there is much in Buddhism that has decent grounds to claim membership for those parts of the tradition that meet this narrow definition. Even hobbling around as the visiting team on this alien terrain, under the paranoid constitution of a manically litigious state, wearing this constricting gear required by the stringent safety regulations enforced by the hosts of this mutant form of away game, Buddhism is a pretty formidable player.

Indeed, one can imagine quite strong claims being made for the role even this restricted form of Buddhist thinking might play in modern philosophy. For modern philosophy begins, according to the standard textbook account, when Descartes declares, “If you would be a real seeker of truth, it is necessary that at least once in your life you doubt, as far as possible, all things.” Thinking was at last to stand on its own feet, without support from the unexamined presuppositions that tend to lie low beneath our explicit ruminations; the ancient Socratic experiment which had been so long sidetracked or bought off or domesticated or enslaved as a handmaiden was now to be free of any heteronomous constraints, to throw off its demented master and run its own household once again. The unnamed cracker of the whip that had put philosophy as autonomous thinking in chains for so long here was, of course, theology, in particular the theology of a *revealed religion*. A supernatural revelation of this kind by definition claims for itself exemption from all doubt, because it makes no bones about placing itself above the realm of confirmable premises: in enforcing its claim that it is “revealed” at a particular time and place rather than deduced from or discovered in universally available premises, it admits and even brags that there is no other way to know that it is true other than by accepting what is revealed as coming from an unimpeachably authoritative supernatural source—and doubting it, the very mental state of doubting per se, in fact puts the fate of your soul at dire risk. Descartes flies in the face of this entire ethos, so it seems: doubt is necessary, doubt is good, doubt is an irreplaceable condition for the revelation of truth.

We know, of course, that if we read onward in the *Meditations*, this turns out to be largely much ado about nothing, a false alarm, maybe even a bit of a bait and switch: Descartes does want his moment of doubt—but only as a (very) temporary means not only to the most apodictic certainty possible, complete freedom from doubt, but also one that ends up doing just what reason was supposed to do in the subordinate role assigned it by the revealed religion: confirm the existence of God as necessary and also *necessarily extrinsic* guarantor of truth, not to mention free will and immaterial soul more or less as revealed back there in the revealed religion that had been with him since the nursery. And, to a Buddhist thinker looking at this spectacle from afar, in this Descartes really is the father of modern philosophy, for this same story seems to repeat *mutatis mutandis* again and again in the subsequent history of Western thought: someone makes an attempt to doubt accepted certainties so as to light out for new territories, but again and again they arrive back at their oldest presuppositions. Indeed, when viewed through Buddhist eyes, it might be claimed more broadly that we can scratch the word “subsequent” here and see

this pattern throughout Western thought both before and after Descartes. And this is what invites me to make a provocative claim about the irreplaceable necessity for anyone who claims to be a “seeker of truth” in Descartes’s sense to throw herself wholeheartedly into the study of Buddhist thought: for Buddhism, even in its restricted form, can make some claim to have broken through this impasse: perhaps it was over there that the program of doubt was really accomplished.

How can I make this absurd claim, given the fact that Buddhism too is a “religion,” and one which depends for its doctrinal claims on authority—if not exactly on the authority of a revelation from a supernatural being who created the world, something pretty close: the revelations of someone who does claim to be a pretty supernatural being (not the creator or judge, but at least a wildly above-it-all best possible knower of the world) with a very special experience of ultimate truths, unavailable (for now) to the recipient of these claims, which cannot be justified independently of those (for now) unavailable experiences. Buddhist traditions do not “doubt all things,” because they accept the authority of the Buddha’s experience of enlightenment—not to mention plenty of other fanciful tales of supernatural comings and goings in an extravagantly mythical cosmos.

But nevertheless, there is room to assert that it may be only in Buddhism that a philosopher, Buddhist or not, can find true doubt. I say this although Buddhists never employ anything like the Cartesian method, and rarely even anything like the Socratic method. But Buddhism has special claims to define the heart of the philosophic quest because, for whatever reason, more or less every other form of thinking known to us seems to be unable to doubt *all five* of the following five things I am about to name. That is, whenever thinking starts to become active and autonomous and alive to its own critical power and self-authorizing claims in Descartes’s sense, able to doubt its own presuppositions and stand on its own to find new truths, the history of non-Buddhist thought (outside of China, may I add) seems to be unable to undermine any one or two or three or four of these five things without immediately, or even thereby, further embracing an unshakeable belief in the fifth, as if there is nowhere else to go besides these five, as if the undermining of any four of them proves the necessity of the fifth. The five are:

- 1) God
- 2) Mental Substance
- 3) Physical Substance
- 4) The Law of Non-Contradiction
- 5) Absolute Ethics

By “God” I mean any *intelligent* source of the world that either plans and designs it, or an incomprehensible ground of the world, beyond intelligence but with a specially favorable relationship to intelligence as opposed to its opposite, grounding it and for us exemplifying some kind of hyperintelligence, that stands as a supernatural guarantor and ground of the world’s consistency and/or reality.

By “mental substance” I mean self-sufficient and indivisible souls of individuals or a single world soul serving as the

ground of experience, and/or of material reality or the appearance of perceived material reality, or even any *irreducible or uncaused* units of thought or experience.

By “physical substance” I mean either atoms or forces or energies or fields or one or many indivisible masses that takes up space and stands as a substrate or as a concomitant for all mental experience.

By “the law of non-contradiction” I mean the unsurpassable logical structure construed as ontological information about the world, such that entities are real if and only if they cannot be both P and non-P at the same time and in the same respect.

By “Absolute Ethics” I mean ethics as first philosophy and last philosophy, either pragmatically or epistemologically defining what can and cannot, or must and must not, be thought, concluded, desired, surmised, assumed, but itself unsusceptible to critique or surpassing by means of anything thought, concluded, desired, surmised, or assumed.

I’m here making a controversial claim both about Western thought and about Buddhism, and there will be legitimate doubts on both sides of this assertion, from students of Western thought on the one hand and from students of Buddhist thought on the other. It should be clear from the above that I mean all of these five items in the broadest possible sense, allowing for a wide range of variation in the details, and without wanting to quibble about whether, on some elaborate interpretation, some exception can be found here and there in some marginal or very modern Western thinker. I submit nonetheless that when the dust settles, one of these five is always found holding up the tent at the end of the day. But I’m willing to simply assert that here as a research agenda: let’s go look and see if this is true or not. My own impression, after many years of concern over this issue, is that again and again, in one form or another, if God is rejected, matter is affirmed; if matter is negated, soul is affirmed; if God and matter and soul are rejected, logical absolutism is affirmed, if all of these are rejected, absolute ethics is assumed, and so on, round and round. It’s *The Cat in the Hat Comes Back* all over again. In that story, the cat promises to clean a stain out of the bathtub, but does so using mother’s best towel. Now the stain is on the towel, so he cleans it off with the curtains. Now the stain is on mom’s beautiful curtains, and so on: the stain keeps getting transferred from one place to another, but mom is on the way home and the stain is still there, somewhere.

The “stain” here, of course, is foundationalism, absolutism, dogmatism, the uncaused as distinct from the caused, the unconditional as absolute other to the conditional, a one-direction chain of grounds that must end in a primary ungrounded ground of some sort: the very idea of “self-nature,” a self-standing, non-relational entity possessing fully on its own account and its own power its own determinate essence, being just what it is *simpliciter* and without qualification, which can be seen as the source or ground of all other composite or derivative epiphenomena. I will not go into the details of the critique of this idea in

Mahāyāna Buddhism (which is mainly what I have in mind here) in its many forms and with its multifarious implications. Nor will I delve into the possible responses Buddhist thinkers have and could have had to the objection that thinking is never presuppositionless, that the adoption of something unquestioned and unquestionable is inevitable and not to be shunned in a quixotic quest for Cartesian presuppositionlessness—but I will note, for interested readers, that those responses are well worth exploring. But I will say that the fact that Mahāyāna Buddhism from beginning to end identifies itself as a tradition that stands or falls with the nonreliance on substances, even very much in spite of apparent exceptions in seemingly backsliding concepts like ālaya-vijñāna or buddha-nature or dharmā-nature. The thorough chasing down and weeding out of all such irreducible first-cause notions, of anything both unconditional and determinate, even determinate to the extent of not-being-the-conditional, is its flagship doctrine, its *raison d'être*, its defining agenda.

Is there nothing like this in European thought? Yes, we have Pyrrhonian skepticism, with fine arguments to neutralize even the absolute authority of logical method and its presuppositions, and a close tracking of the influences of this unique incident in Western thinking throughout subsequent European traditions, pro and con, perhaps does some of the same heavy lifting, and it is true that this indubitably plays a part in modern philosophical curricula. But though we have here a similarly thorough countercommonsensical skepticism, in comparison to Buddhist commitments and elaborations we find it only in a soon-truncated form, roundly batted down by its respondents, extremely modest in its aims, a torso with no head. The systematic and practical application of Buddhist anti-foundationalism is different. Exactly how and on what premises, and to what effect, is precisely what we miss the opportunity to explore when we leave it out of account in our study of philosophy.

But do the Buddhists really succeed in their dismissal and doubt of these five items? Do they really want it quite as thoroughlygoingly as I'm suggesting? Here is where the other side may raise some objections, for some interpreters of Buddhist thought will also find at the end of the day that what Buddhist thinkers *really* mean cannot do without resting firmly on one or two of these items—most commonly, a highly unusual version of either physical substances or mental substances (or perhaps a neither-physical-nor-mental substance or substances—dharmas, dharma-nature, buddha-nature, etc.), and above all the law of non-contradiction, or the pragmatic ethics of a specific form of Buddhist life. I take this to be an interpretative misunderstanding of the entailments of the Buddhist sources, or perhaps just another example of the very lack of philosophical imagination I am trying to spotlight here. But in both the Western and the Buddhist cases, the conversation must proceed from this point through careful analysis of individual philosophers and texts. My own view is that in many if not most Buddhist systems *all five* of these items are not only doubted, but subjected to a principled and sustained rejection, at times in the form of an argued refutation, at times in the form of the premises of further developments that proceed in their absence.

This is not the same as doubt in Descartes's sense, in that precisely *this undermining of the five* is the very content of the undoubted doctrines delivered by the authority of the Buddha or, later, by philosophers who are themselves regarded as infallible bodhisattvas. Dogmatic nihilism! some might cry, and this is perhaps one of the reasons for that aversion to Buddhism—above and beyond the blasé neglect or condescending dismissal that philosophers have otherwise shown for this world of thought—lately given the handy name *Buddhaphobia*. And it might be argued, I think somewhat convincingly, that "seeker of truth" is not quite the right name for what a Buddhist thinker is: he is a seeker of liberation first and foremost, and that quest, while it perhaps must get its hands as dirty as possible in the thickest and most intricate tangles of thinking, does not do so in order to arrive at a thought which contains true propositions once and for all. If the latter is what Descartes or Socrates means by "truth" as what is sought or loved by a philosopher *qua* philosopher, then the Buddhist thinkers are not philosophers. But even if we grant this highly debatable description of the meaning of "truth" as it pertains to Western philosophical practice, and even if we bend over backwards to grant the forefronting of the soteriological framing and authority structures informing the practices of Buddhist thought (in my view rarely relevant to the on-the-ground proceedings of the actual thinking), we might still insist that the study of Buddhist thought is essential to every philosopher. For in Buddhism we see at least the possibility for a living, breathing human being to reject all five of the items listed above, and to the great profit of his thinking and living: to live and think and flourish in *the absence of God, soul, matter, absolute morality, and absolute logic*. Here alone can we envision what it would mean to doubt all things and to stay there, at last—and, not least, why anyone would want to do that.

What is left for a worldview that accepts no physical realities, no absolute logical laws, no absolute morality, no controlling creator God, and no souls? Buddhism. A philosopher who wishes to stand thinking on its own two feet, to *be able* to doubt all things, must become aware of what things there are to doubt, and how difficult it is to doubt them, how easily they slip back in unnoticed after being briefly doubted, how cunningly they replace each other and compensate for one another, how doubting one unjustified assumption tends to bolster an alternate one, what happens when they are doubted, whether it is possible to doubt *all* of them at the same time. Buddhists arguably fail to doubt the reliability of the Buddha, and in this they would to be judged to have failed Descartes's mission. Perhaps they therefore fail to doubt the key doctrines of Buddhism: non-self, anti-substantiality, radical atheism, translogicism, and so on. But without Buddhism, we would have no example in the history of the world of a well-thought-through, systematic, elaborate, rigorous, multi-millennial tradition of thought that functions and flourishes without falling back on one or more of the five. And perhaps to be *able* to doubt these five, to *succeed* in doubting them, it is necessary to know and understand this, and to see what the implications of that have been in all their multifaceted and varied forms throughout the histories of diverse Buddhist traditions of thought.

The above would be the case for a robust and even militant advocacy of Buddhism as a necessary component in the academic study of philosophy. But actually, for my part, I think the only reason this would matter one way or another is for the sake of those tragic freshmen I alluded to earlier. I'd like them to get the nurturing of their innate speculative talents that they yearn for, and I think it would be good for the world to see their thinking brought to term rather than aborted by the pressures of "philosophy" as now defined. Other than that, I would be very content to leave "philosophy" to its own devices, to disgrace itself or find new glories as it pursues its present narrow course—as long as those of us who do not fit the agenda can find the resources and the place in the world to flourish and to develop our alternate modes and methods and concerns. It is a pity, perhaps, that it gets to use the brand and wield the prestige of its noble ancestor to the exclusion of rival claimants. But which of us is better serving the spirit of philosophy as it was once known and lived is perhaps something only the future can judge.

What/Who Determines the Value of Buddhist Philosophy in Modern Academia?

Hans-Rudolf Kantor

HUAFAN UNIVERSITY, GRADUATE INSTITUTE OF ASIAN HUMANITIES, TAIPEI

BUDDHIST PHILOSOPHY AND MODERN ACADEMIA

"Philosophy" is not an expression coined by Buddhists in Asia, just as Buddhism is not a teaching developed by philosophers in the West. However, for many in the field of philosophy in the social system of modern academia, "Buddhist philosophy" is a term which is substantiated by their perception that traditional Buddhist thinking deals with subject matters which belong to the concerns of traditional Western philosophy. But who is it, then, who says what "Buddhist philosophy" is? And who or what determines the value of Buddhist philosophy in modern academia?

Modern scholarship would hardly deny that interests, content, topics, and approaches developed in philosophical discourse are historically grown, as it seems to be rather unlikely that thinkers of different times, eras, as well as socio-cultural and linguistic backgrounds, share the same concerns. This, of course, does not exclude an understanding across the temporal, cultural, or linguistic divides which philosophers might experience in their attempt to explore a source of inspiration that has developed independently from their own backgrounds. Nevertheless, philosophical questions and interests are distinctly informed by the temporal, social, cultural, and linguistic contexts in which they occur.

Hence, philosophical discourse across the East-West divide, as it is practiced not only in analytical philosophy but also in post-modern approaches, or by East Asian thinkers, tends

to adduce Buddhist views in order to tackle questions and problems that typically characterize the domain of thought in the social system of modern academic philosophy, which is deeply rooted in the cultural practices of Western traditions. For instance, currently discussed themes in that system are the question of free will in respect to Buddhist thought; the logical implications concerning paradoxes in Mahāyāna seen from the viewpoint of non-classical logic; Buddhist contributions to metaphysics and ontology; epistemological reflections of Buddhist thinkers in the light of cognitive science, phenomenology, idealism, etc.; Buddhist ethical values related to the question of human rights; Buddhist views on personhood and identity; Buddhist treatments of linguistic concerns seen from the viewpoints of recent philosophy of language; Buddhism in relation to post-modern gender issues; and the extent to which Buddhism coincides with environmental concerns and inspires eco-criticism, or might overlap with natural science.

The same applies to those East Asian thinkers from the twentieth century who considered themselves philosophers in a newly established modern academic system rather than literati or intellectuals in a traditional East Asian society. The most well-known figures are Mou Zongsan (1909–1995), Tang Junyi (1909–1978), Nishida Kitarō (1870–1945), and Nishitani Keiji (1900–1990). All of these figures were significantly inspired by Buddhist thought, and yet the philosophical problems that they addressed in their examination of Buddhist sources reveal very strong ties both to the topics discussed in the modern academic system and to the related traditions of Western thought, such as Buddhism in respect to the epistemological function of intuition, to the status of metaphysics and ontology, to ethics linked to the notion of the free will, to the role of religion in modern society, and to political theory and the building of modern society.

In all the accounts produced in this system, the evaluation of Buddhist teachings depends, then, on the amount of benefit which modern philosophical discourse expects to gain from including them. This becomes evident, for instance, if we look at Mou Zongsan's philosophical project of modernity. His Chinese work *Buddhanature and Prajñā* (佛性與般若. Taipei: Xuesheng shuju, 1982) is the first monograph which recounts the formation of East Asian Buddhist philosophy from a systematic point of view, as distinct from the usual chronological approaches. Together with his other works, he presents and evaluates pre-modern Buddhist thought within an expanded discussion about the relevance of Confucian values for his vision of modernity into which he further integrates Kantian ideas and those of other Western thinkers. Buddhist thinking plays a role only as a conceptual instrument that he borrows to construe his syncretism of values, based on which a global sense of modernity across the East-West divide is to be anticipated. Moreover, his systematic account selects only sources and thoughts which fit his overarching conceptual framework, while other traditionally important positions and thinkers are not even mentioned. This illustrates that what philosophy in Buddhism means to those who operate in a modern world and its academic system is inevitably defined by the selective perception informed by certain and particular interests immanent to that system.

In other words, we observe that academic interests, often enough, only marginally overlap with the seminal concerns that have been effective in shaping the traditions in which Buddhist doctrinal thought has been developed and transmitted. The most crucial one, certainly, is the soteriological concern of liberating the minds of sentient beings from their self-induced deceptions and the suffering that is rooted therein, as many traditional sources which focus on doctrinal exegesis explicitly hint at this purpose of their production. Of course, modern academia is aware of the discrepancy which such prioritizing of differing concerns entails. But again, at this point, it is important to mention that what Buddhist philosophy means and what its value signifies to those who develop an interest in it is tantamount to its potential contribution in solving an agenda of philosophical questions relevant to and typical of the discourse which the social system of modern academia has generated.

All this is not really a criticism, since the system as a whole can hardly operate in a different way. By virtue of its very nature, its functioning is selective. However, what matters and makes a difference for the agents in it is their awareness of this. A similar discrepancy might probably also occur when we look at this problematic from the opposite point of view, addressing the same question to the tradition of Buddhist thought. Only selected parts of the philosophical discourses which, in terms of origination, we would attribute to the traditions in the West would be considered beneficial to what is at stake and of particular interest in the world of thought that Buddhists have traditionally been committed to.

Therefore, in its attempt to capture the type of thinking that has independently developed within the Buddhist traditions, modern discourse would need to rephrase the question of the title in this way: To what degree can emic and traditional concerns, which are the constitutive factors in the formation, development, and transmission of Buddhist doctrines, be considered relevant to philosophical discourse in the system of modern academia? Seen from the viewpoint of philosophical hermeneutics, the modern academic understanding of pre-modern Buddhism would need to reflect on its own background of interests, agenda of questions, and also implicit pre-occupations informing its approaches, in order to be then capable of identifying and describing those seminal concerns. This is not at all solely a problem of historical discourse. To look at doctrinal contents apart from their formation within the traditional patterns of transmission means to miss the manner in which Buddhist thought really is, or could be, an independent source of inspiration for philosophical discourse in modern academia.

“PHILOSOPHY IN BUDDHISM” AND “BUDDHIST PHILOSOPHY”

Perhaps the distinction between “philosophy in Buddhism” and “Buddhist philosophy”—which is a distinction only between ideal types (*Idealtypus*)—might help to delineate the peculiarity and the value that philosophical thought in this particular tradition might imply. “Philosophy in Buddhism” would denote and signify the contributions of traditional

Buddhist thought to philosophical discussions in modern academia, while “Buddhist philosophy” would designate and underscore the particular type of thinking characteristic of the traditions wherein it has been developed. The value of “philosophy in Buddhism” depends then on the extent and significance of such contribution, whereas “Buddhist philosophy” would rather display an independent value of its own, and yet remain within philosophical discourse not strictly committed to ranking the relevance of its topics.

These are two different ways of looking at philosophical thought developed and transmitted in the Buddhist traditions. Most importantly, “Buddhist philosophy” in the context of modern academia would be more committed to philosophical hermeneutics, as it would aspire to understand Buddhist thought in its formation within the traditional patterns of transmission, without excluding the concerns of “philosophy in Buddhism.” In accordance with the insights of philosophical hermeneutics, “Buddhist philosophy” would need to cultivate an awareness of the historicity and contingency of its own interests and concerns which determine the selection of its themes and the perspectives on them.

Consequently, “philosophy in Buddhism” must then rely on the understanding of “Buddhist philosophy,” if it is to be deemed indispensable and unique in its contribution to modern academic discourse. If, however, “philosophy in Buddhism” is given priority so that the value of “Buddhist philosophy” depends on it, or the two are no more different from one another, the evaluation of Buddhist thought would then solely be determined by parameters which completely disregard the independence of the tradition that has generated it. This would hardly match the ethical standards which Bryan v. Norden and Jay Garfield have outlined in their conception of philosophical discourse in modern academia—a discourse of diversity, devoid of racism, or superiority complexes of Western thinkers, or their tendencies toward discrimination and exclusion.¹ Therefore, such a discourse should integrate these two ideal types—that is, the two should be related to one another in a fashion in which each takes the other into account, yet without denying the difference that persists between them.

Again, it is important to note that this difference is not one between philosophical and historical discourse. “Buddhist philosophy” deals not solely with the chronological order or diachronic development of thoughts. It focuses on thoughts in their formative and contiguous process because it intends to describe the specific dynamic wherein they persist. However, the reason why a certain thematic object is selected and appears to be relevant has to do with what constitutes the perspective of its description. It is the discourse of certain topics in modern academia which is the condition that enables “Buddhist philosophy” to emerge from its traditional background, and this must be explicated within that description. This approach would need to become aware of the contingency and historicity of its own concerns, and this is the specific manner in which “Buddhist philosophy” takes “philosophy in Buddhism” into consideration. Again, two aspects are important in this approach: (1) the specific dynamic of the

process wherein Buddhist thoughts have evolved; and (2) the specific conditions of modern discourse which enable “Buddhist philosophy” to emerge in it.

As “Buddhist philosophy” accounts for the approach which does not veil the facts of its own temporality and contingency, it has the capacity to embrace not only the method of modern philosophical hermeneutics but also the perspective that many traditional Buddhist thinkers have adopted when they have expressed what truth has meant to them. For many traditional Buddhist thinkers, such as the Mahāyāna masters, “ultimate truth” is what never separates from its opposite, “conventional truth.” This is provisional and does not cease to alter, as it is responsive and adaptable to the constantly changing circumstances which each instance of evidence that reveals a sense of truth is inextricably bound up with.

That view can indeed be considered as an important point of intersection between Mahāyāna thought and philosophical hermeneutics. The concept of the two truths—conventional truth and ultimate truth which persist only in correlation, as many Buddhist sources emphasize—inspired the pre-modern protagonists of the formative process of the Chinese Buddhist schools to construe doxographies with the interpretative purpose of (1) outlining the hidden coherence between all the numerous doctrines in the bulk of texts transmitted from India and translated into Chinese, and (2) showing that all these manifold expressions of Buddhist teaching are congruent with the ultimate meaning of awakened liberation which itself, however, is independent from any speech and thus reaches beyond language.

Hence, the Chinese practitioners and interpreters developed an ambiguous stance toward the textual transmission of Buddhist doctrine: independence of its ultimate meaning from speech on the one hand, and indispensability of the canonical word in the understanding of that meaning on the other. This indicates an awareness of the hermeneutical situation and its temporality which the accomplished understanding must realize in respect to its own exegetical activity. Such awareness promoted practices of self-referential observation and became an influential factor in the formation and transmission of Buddhist thought in East Asia. It shaped an approach which many of the traditional Chinese Buddhist masters, such as Tiantai master Zhiyi (538–597), Sanlun master Jizang (549–623), Huayan master Fazang (643–712), and also the masters of the Chan schools pursued—which could be referred to as “practice qua exegesis”: To practice the path to liberation is to specialize in doctrinal exegesis. “Practice qua exegesis” paradoxically culminates in accomplishing detachment from textual expression via such expression, or via other performances. Whereas Tiantai interpretation taught the inseparability of construction from deconstruction, the Chan masters tended to shift their focus on the performative aspect in their practices.

In a modified manner, the universal point of this stance might even concern the hermeneutical situation of modern academic philosophy, which asks for the role that the textual heritage of traditional Buddhist thoughts might play

for its discourse. Therefore, far from being just an object of chronological description, “Buddhist philosophy” itself would adopt, or be inspired by, essential considerations or positions developed and transmitted in the traditions that it refers to and, at the same time, would also follow the approaches of philosophical hermeneutics, which would fit the requirements of the modern academic system.

INDO-TIBETAN AND SINO-JAPANESE BUDDHIST PHILOSOPHY

Of course, in the world of Buddhist thought there are various traditions of transmission. Apart from few exceptions (Steven Heine, Brook Ziporyn, Graham Parks, Dan Lusthaus, etc.), the majority of academics in the West (Mark Siderits, Jay Garfield, Tom Tillemans, Georges Dreyfus, Guy Newland, Graham Priest, Jan Westerhoff, Jonardon Ganeri, Dan Arnold, Mario D’Amato, Sara McClintock, Karen Lang, Christian Coseru, David Eckel, Joseph Walser, Jose Ignacio Cabezon, Thomas Wood, etc.) emphasize and consider Indo-Tibetan Buddhist traditions, while modern scholars in East Asia, (Xiong Shili, Mou Zongsan, Tang Junyi, Wu Rujun, Nishida Kitarō, Keiji Nishitani, Masao Abe, Hajime Tanabe, Shizuteru Ueda, Hisamatsu Shinichi, etc.) primarily elaborate on Buddhist sources in Chinese or Japanese. This, of course, has to do with the fact that the Mahāyāna sources in Sanskrit and the three major traditions of textual transmission—the two Northern traditions in Chinese and Tibetan, as well as the Southeast Asian Theravāda tradition in Pāli, along with the commentarial and scholastic elaborations in each of these three regions—have developed and coexisted almost without mutual influence, although quite a few of the Indian root texts relevant to the formation of these three transmissions overlap to a certain extent. Studies which compare the peculiar features and philosophical characteristics of thought in these three have barely evolved yet. One of the few exceptions is Nakamura Hajime, yet his work does not really venture into a deep reading of the indigenous Chinese schools, which have been immensely important for the East Asian development of Buddhist thought.

One of the major problems which besets the attempt to make Buddhist thought accessible to philosophical discourse in modern academia consists of the self-referential or closed system of doctrinal concepts and idiosyncratic terminology in Buddhism. This is particularly true with regard to doctrinal exegesis in the Chinese Song (960–1279), Yuan (1279–1368), and Ming (1368–1644) dynasty sources. At that time debates in Chinese Buddhist exegesis had become very convoluted and specialized due to the long history of scholastic and sectarian ramifications. A coded idiom was used to communicate the complexity of references and contexts which the contentious issues in these debates between the different exegetical traditions implied.

Therefore, Western studies in Buddhism which deal with philosophical issues seldom reference those sources. For instance, the Cowherds’ (self-designation chosen by the authors: Jay Garfield et al.) book about the two truths in Madhyamaka thought, *Moonshadows—Conventional Truth in Buddhist Philosophy*,² does not contain a single chapter

about that topic in the Buddhist philosophy of China, although it is this doctrine of the two truths which was one of the most debated issues in the exegetical traditions of medieval China, leading to the formation of the first indigenous Buddhist schools, those of the Sanlun and Tiantai, at the end of the sixth century. Sanlun master Jizang even composed a lengthy treatise on this Buddhist topic. Similarly, Zhiyi's complex Tiantai teaching is construed on the basis of this doctrine.

Moreover, seen from the background of the extremely abundant Chinese literature on this Buddhist topic, it is very surprising that the Cowherds focused only on "conventional truth," although this truth persists only in correlation to/with "ultimate truth" according to the textual transmission in Chinese. This is to say, understanding conventional truth consists in comprehending ultimate truth and vice versa, which also applies to *nirvāṇa* and *saṃsāra*. Apart from its opposite, neither of the two can be adequately comprehended according to the pre-modern Buddhist texts translated into Chinese and those composed by the indigenous interpreters of them. It would have been necessary to explain why the modern discussion of this subject must deviate from the views in traditional Chinese Buddhism. All this illustrates that the content of Buddhist philosophy seems to exceed the scope which the recent promoters of this field within institutionalized academia in the West have, so far, been aware of.

Hence, the process of integrating Buddhist philosophy into the institutional and educational framework of modern academia in the West must be further advanced, as it is far from being completed. But most importantly, this process must overcome its strong tendency to prioritize "philosophy in Buddhism" over "Buddhist philosophy." The relevance of "Buddhist philosophy" does not need to be evaluated (solely) through an assessment of its meritorious contribution to philosophical discourse in modern academia, because this would be the value that it has for something other than itself. It would be more intriguing and enriching, also for the sake of plurality and diversity to which North American universities are so much committed, to recognize and display the value that it has in itself.

NOTES

1. Bryan van Norden and Jay Garfield, "If Philosophy Won't Diversify, Let's Call It What It Really Is," *New York Times*, May 11, 2016; *Taking Back Philosophy—A Multicultural Manifesto* (New York: Columbia University Press, 2017).
2. Jay Garfield et al., *Moonshadows—Conventional Truth in Buddhist Philosophy* (New York: Oxford University Press, 2011).

Buddhist Philosophy? Arguments From Somewhere

Rafal Stepień

NANYANG TECHNOLOGICAL UNIVERSITY, SINGAPORE

INTRODUCTION

This article is principally a critique of arguments purporting to demonstrate that Buddhist philosophy is *not* philosophy. Now, the reasons as to why such arguments have been and in fact continue to be deployed (and thus stand in need of critiquing), and more generally as to why analytic philosophers (who continue to dominate the profession far beyond the Anglo-American axis at the analytic core) are preponderantly inimical (rather than sympathetic) to the idea of Buddhist thought being taken seriously as philosophy, doubtless have far more to do with personal biases and ignorances as well as with the systemic conservatism (and consequently, given the particular past they strive to conserve, the racism) of academic institutions than they do with reasoned argumentation (that is, if that's its definition, with philosophy). But reasons there are, and these tend to coalesce around the purported areligiosity of analysis and the related refusal to engage in metaphysics on the part of analytic philosophers (as opposed to the imputed religiosity and metaphysical nature of Buddhist arguments), as well as reference to (largely mythical) historical and linguistic threads linking all Western philosophers but alien to their brethren to the East or South. In discussing these reasons, I will leave aside the (to my mind obvious) fact that the practice of analytic philosophy as a whole (its areas of concern, its very identification of and manners of framing these topics, its criteria for adjudicating success, not to mention the personal convictions of many of its adherents, not ever wholly extricable from professional practice . . .) is governed by a prior intellectual history (including a prior *philosophical* history) utterly infused with religious presuppositions (which, for that matter, go well beyond what Heidegger called ontotheological ones). I leave aside also the related point that all philosophizing, down to the most nitpickety punctilios of the seemingly "hard" philosophies of logic, language, and mind (the preferred fields of traditional analytics) are, whatever their practitioners may claim to the contrary, suffused with metaphysical premises. Rather (that is, instead of pointing out their self-deception and double standards), I want for the moment to grant analytic philosophers their cake so as to see (not how it may taste—that particular fusion of horizons must await another symposium—but) whether it is reasonable to swallow it.

It should be clear that the *unargued* dismissal of any intellectual position or tradition is not conducive to disputation and refutation. Such unargued dicta as would simply deny the status of "philosophy" to non-Western thought traditions (including the Buddhist) are therefore not themselves philosophical; as such, I will not bother considering them here. Instead, I will draw liberally from Roy Perrett's classification of the four reasons standardly given to justify the divide, which I will call the Historicist

Argument, Terminological Argument, Argument Argument, and Religion Argument.¹ Although I augment Perrett's critical rejoinders with observations and examples of my own, nevertheless it should be clear to readers familiar with the relevant debates and diatribes that my descriptions and discussions of the arguments/conditions in the paragraphs that follow is heavily indebted to extant publications. I mention some of the more direct borrowings *ad locum*, but I hasten to admit that I do not claim to be making any major original contribution to what I see as an already done-and-dusted matter—at least philosophically, and there's the rub. For although scholars of Buddhist philosophy (or of many another non-Western philosophy) may well consider that "critique of the narrowness, arbitrariness, and ethnocentrism of this characterization [according to which only Western philosophy is properly to be accounted philosophy] is too easy and too boring to undertake,"² unfortunately, we are still a long way indeed, practically speaking, from actually having the conclusion—viz. that non-Western philosophies are philosophies—incorporated into institutional academic practice. As such, and in addition to the suggestion I make in the final portion of this piece as to philosophy being properly located *nowhere*, I hope that the present contribution may at least serve to assemble in one place (or better: one other place) the relevant arguments and counter-arguments, and thereby to act as a(nother) convenient gathering point for those of various persuasions along the route toward mutual comprehension, if not agreement.

PHILOSOPHY & HISTORY

Put succinctly, the Historicist Argument goes as follows: Philosophy has historically been practiced in the West; therefore, philosophy is a Western phenomenon (or, more strongly, a Western phenomenon *alone*). Heidegger provides a classic formulation of the strong version of this stance in stating:

The often heard expression "Western-European philosophy" is, in truth, a tautology. Why? Because philosophy is Greek in its nature. . . . The statement that philosophy is in its nature Greek says nothing more than that the West and Europe, and only these, are, in the innermost course of their history, originally "philosophical."³

In assessing the strength of this argument, note firstly the conflation Heidegger enacts between "Western-European" and "Greek." It is an obvious point, but "Greece" is not co-extensive with "Europe," much less with "the West." If we were to accept Heidegger's premise that "philosophy is Greek in its nature," the valid conclusion would be not that philosophy is an exclusively "Western-European" enterprise but an exclusively Greek one. This, however, would of course have the rather undesirable consequence that, depending on how "Greece" is defined, either philosophy stopped with the demise of ancient Greece or continues to this day only within Greek national or linguistic borders.

But let's grant for a moment Heidegger's conflation, accepting that philosophy has indeed been practiced in the West broadly understood. Is this a valid reason to deny that the same activity has been undertaken in other parts of the

world? Do we claim that music, art, or poetry are practiced only in Western civilization (whatever that is . . .) on the grounds that these domains are conceptually grounded, for "us" Westerners, in their Western histories? Even if we were to accept the claim that the first people to ever argue about the nature of reality were the Greeks—which is historically just plain false—would this suffice to restrict philosophy (if it is understood to be just such argument) to the West? After all, Indian linguists, for example, elaborated highly sophisticated analyses of grammar, phonetics, and semantics well before Europeans did: Does this mean that Europeans cannot legitimately speak of "linguistics"? And let us not forget the fact that Greek philosophy was itself largely forgotten for several centuries in the West, and only reintroduced via the translations and commentaries of Islamic scholars writing largely in Arabic. Perhaps, then, we should conclude that philosophy is an exclusively Arabic or Islamic endeavor, since there is no unbroken historical thread linking ancient Greek and post-Renaissance European thought? Hopefully, the various versions of a *reductio ad absurdum* I have proposed will suffice to persuade you that these and other such Historicist Arguments are patently false. They rest on a combination of conceptual conflation and historical fallacy.

PHILOSOPHY & TERMINOLOGY

This applies, *mutatis mutandis*, also to a second argument, sometimes taken as a linguistic sub-version of the Historicist Argument, which I call the Terminological Argument.⁴ This states that since there is no (exactly) equivalent term for "philosophy" in Sanskrit, Chinese, Tibetan, or take your pick of any other non-European language, therefore there is no such thing as philosophy practiced in any of these language communities. This I find to be the most peculiar of the four arguments I am summarily working through, and one subject to what are surely two inter-related and fatal objections. Firstly, if the exact denotative and connotative force of a given term is what encloses the space, as it were, of the activity defined by that term, then the ancient Greek term *philosophia* is surely as distant and distinct from the contemporary English word "philosophy" as any potential equivalent in another language. After all, no one would claim that the cultural context informing the assumptions and presuppositions—the very intellectual frameworks—of the paradigmatic ancient Greek philosophers are similar to, let alone exactly equivalent to, those of today's English-language philosophers. Indeed, the fact that the overarching worldviews of members of these cultural worlds are vastly divergent itself reduces the Terminological Argument to the absurd conclusion that only the ancient Greeks practiced "philosophy." For even if we manage, through historical scholarship, to excavate, even to some extent comprehend, the manner in which a Plato or an Aristotle saw the world and his part in it, nevertheless we do not, and cannot, *share* that view. As such, to follow the logic of the Terminological Argument rigorously is to deny "philosophy" to any but the original practitioners of the activity denoted by the etymologically originary term. Indeed, we could force the point even further, and argue that, if we accept that no term means quite the same thing to any two language users, then strictly speaking we should restrict "philosophy" to the individual originator of the term itself—mythical though such a speaker be.

Secondly, and relatedly, why should the absence of an (exactly) equivalent term in one language to a term in another language (assuming for the moment that such trans-lingual exact equivalents exist . . .) entail that no such concept or activity exists at all in that language community? To advocate such a position would be to adhere to a very strong version of the Sapir-Whorf hypothesis, and would likewise result in absurd consequences. Thus, applying the same logic, we would be led to conclude that there is no such thing as poetry in anything but Greek (or its direct etymological descendants) since the term itself derives from the Greek *poiesis*. Likewise, there can be no algebra in the West because the word used in English and numerous other European languages is derived from the Arabic *al-jubrā*,⁵ and included fields of mathematics we would now classify under different branches. The denotative range (and therefore the very meaning) of the term in its original Arabic being quite different from that of its European-language etymological descendants, we are left to conclude that, while Europeans working in these languages may well engage in algebra-like thinking, “algebra” as such must be reserved for Arabic language users . . . just as non-Westerners may engage in philosophy-like thinking but not “philosophy” itself.

Or let me give a counter-example that remains within the Western tradition. We have all heard of the old sign standing atop Plato’s Academy, according to which no one ignorant of geometry was to be allowed entry. It would seem that, for Plato, philosophy without geometry was not philosophy at all—and if anyone should know what philosophy is, surely it should be Plato! Following the logic of the Terminological Argument, then, anyone not working in Philosophy and Mathematics (with a specialization in geometry no less) is not a real philosopher: Sorry! An acquaintance with the history of philosophy closer to our own time teaches us similarly that what we think of as philosophy today used to include fields now called physics, biology, astronomy, geology, etc. So again, either we conclude that Westerners active prior to the inauguration of these (now distinct) disciplines were not philosophers, or conversely we conclude that they were . . . but we cannot be. In either case, the Terminological Argument, like the Historicist one, entails conclusions we are all, I take it, unwilling to accept.

PHILOSOPHY & ARGUMENT

Turning to what I call the Argument Argument, this proposes that philosophy is an activity defined by the use of *argument* for or against a given claim.⁶ There may be plenty of wisdom in world literature, but only philosophy (that is, Western philosophy) deploys arguments in support of its conclusions. There are two ways we may defuse this argument. On the one hand, as with the Religion Argument (see below), it soon becomes clear that it leads us to a conclusion we are in all likelihood unwilling to accept; that is, that a great deal of the core canon of Western philosophy does not count as philosophy at all. Have you been taught that the Pre-Socratics were the first philosophers? Forget it. Throw them out: there are no arguments there, just gnomic pronouncements. Throw out a whole lot of Plato too, as well as much of Nietzsche, Kierkegaard, even Wittgenstein.⁷ The point is that philosophers have used a great variety of literary forms in which to couch their works,

from myths to dialogues, poems, parables, epigrams, etc. If we are serious about the Argument Argument, we will have to ditch all of this.⁸

On the other hand, perhaps we do want to bite the bullet. Let’s excise all that dross from philosophy; it will be the purer and better for it. So be it: Analytic philosophy has been taken by many to have originated in “a philosophical revolution on the grand scale—not merely in a revolt against British Idealism, but against traditional philosophy on the whole” (Preston, Introduction). However, this will in any case not lead to the conclusion that only Western philosophy is philosophy, for as even cursory acquaintance with relevant Buddhist (or Hindu, or Islamic, or Chinese, or African . . .) philosophical texts demonstrates, there are a great many works from within these traditions that utilize argument. Indeed, the logical rigor, technical sophistication, and intellectual ambit of the arguments deployed by Buddhist philosophers (to limit myself to the particular case under discussion here, but to in no way imply any lesser status to members of other non-Western philosophical traditions) are, in fact, formidable. If we are content to toss out the non-argumentative bathwater, the Buddhist baby will still remain, just as argumentative as any child.

PHILOSOPHY & RELIGION

And so we arrive at the fourth and final argument in support of philosophy as an exclusively Western activity to be discussed here. This I call the Religion Argument, and it is perhaps the most commonly heard of the four. Here it is as formulated by Jay Garfield:

There is a world of difference between philosophy and religion, and what passes for “Eastern philosophy” is in fact religion misnamed. Western philosophy is independent of religion, and is a rational, religiously disinterested inquiry into fundamental questions about the nature of reality, human life, and so on.¹⁰

The idea here is that philosophy as “we” (Westerners) understand it, as we practice it, is *not* religion, whereas what some call non-Western philosophies are in fact inveterately religious thought-traditions. Thus (in a wondrously circular argument), Buddhism being a religion, it is strictly speaking a misnomer to speak of Buddhist “philosophy” for, being a religion, it cannot count as philosophy. Garfield continues:

But this distinction is supposed to deliver the result that St Thomas Aquinas’s *Summa Theologica*, Descartes’s *Meditations*, including the proofs of the existence of God, and Leibniz’s discussion of theodicy are philosophical, while Dharmakīrti’s investigations of the structure of induction and of the ontological status of universals, Tsongkhapa’s account of reference and meaning, and Nāgārjuna’s critique of essence and analysis of the causal relation are religious. Anyone who has a passing familiarity with all of the relevant texts will agree that something has gone seriously wrong if this distinction is taken seriously.¹¹

Now, it does not take a deep knowledge of the history of European philosophy to realize that it is permeated through and through with Christian doctrines and premises. If we take the Religion Argument seriously, we would have to jettison just about every Western philosopher from the canon of Western philosophy, which is surely (again) not a conclusion most Western philosophers are willing to countenance. Given that there is no justification for applying a condition of strict secularity discriminately (i.e., to non-Western philosophies, but not to Western ones), this argument also fails.

PHILOSOPHY FROM NOWHERE

Having surveyed and dismissed four arguments standardly put forward against treating Buddhist (and more broadly non-Western) philosophies as “philosophy,” I want to make one final point about these and any other potential arguments in support of the idea that philosophy is an exclusively Western enterprise, or indeed in support of the idea that philosophy is grounded in, and therefore limited to, any one geographic-cultural region—be it on the basis of (stipulatively) definitive necessity or (selectively) historical contingency. Contrary to Heidegger, I want to suggest that philosophy—philosophy above all—is *not* Greek, *not* European, *not* Western. Don’t get me wrong: The last thing I want to claim is that only non-Western philosophy is philosophy. I am thinking instead of something A. L. Motzkin, a prominent Jewish philosopher, proposes in his piece on “What is Philosophy?” In the context of a discussion of Socrates, he writes:

In a word, and in a most fundamental way, the philosopher is unGreek. . . . In other words [and, be it noted, in outright contradistinction to Heidegger’s “tautology”], the phrase *Greek Philosophy* is an *oxymoron*. The first thing which Plato and his descendants throughout the ages, of whom the foremost was Aristotle, would like to call to your attention is, philosophy, is *not* Greek, *not* Roman, *not* French, *not* Russian. . . . The philosopher is neither child of his times nor even stepchild of his times.¹²

In other words, if we do take seriously the idea that philosophy (even, if philosophers are to be believed, *quintessentially* philosophy) is the unbiased pursuit of truth, then we are methodologically obliged to discard any and all of our biases, be these nationalist, racist, ethnic, religious, political, etc., to the extent possible. Only thus will what I would like to coin here the “solocidentary” view, according to which solely the Occident has philosophy, be seen to be but prejudice. However much it may be true that we are always already grounded in what Gadamer, borrowing (or perhaps stealing) from Hölderlin, called the “conversation that we ourselves are,”¹³ nevertheless, if it is truth we seek, wisdom we love, then I propose that it is only by viewing from both “here” and “there,” from “within” and “without” simultaneously;¹⁴ in other words, only by going beyond ourselves, beyond conversation *cum* internalized monologue into conversation *cum* other-oriented polylogue, that we may break out of the paradigms of our times, the “common sense” around us, and think outside those boxes; in other words, actually think.¹⁵

NOTES

1. Perrett, *An Introduction to Indian Philosophy*, 4–6. Perrett himself speaks of the historicist condition, lexical equivalence condition, argumentation condition, and secularity condition respectively.
2. D’Amato et al., *Pointing at the Moon: Buddhism, Logic, Analytic Philosophy*. Also cited, alongside Jay Garfield’s lambasting of the “postcolonial racism” (Garfield, *Empty Words: Buddhist Philosophy and Cross-Cultural Interpretation*, 253) characteristic of the “intellectually and morally indefensible” (Garfield, *Engaging Buddhism: Why It Matters to Philosophy*, ix) nature of the mainstream view, and in the context of a broader theoretic-methodological discussion regarding the study of Buddhism as philosophy and/or/as religion, in Stepien, “Orienting Reason: A Religious Critique of Philosophizing Nāgārjuna,” 1081.
3. Heidegger, *What Is Philosophy?*, 29–30. This passage is also quoted by Van Norden (*Taking Back Philosophy: A Multicultural Manifesto*: 25; “Western Philosophy Is Racist”), together with a not dissimilar pronouncement by Jacques Derrida (regarding which, see also Perrett, *An Introduction to Indian Philosophy*, 5).
4. The amalgamated Historical-Terminological Argument, as propounded by Nicholas Tampio (“Not All Things Wise and Good Are Philosophy”) is discussed, and comprehensively (even hilariously) disposed of, by Van Norden (*Taking Back Philosophy*: 16–17; “Western Philosophy Is Racist”), Garfield (“Foreword”: xiv–xix), and Olberding (“DoD named ‘Wise and Good!’”).
5. Cf. Van Norden, “Western Philosophy Is Racist.”
6. For a more extended discussion on argument and its spouse rationality, see in particular the section on “Philosophy as love of reason,” including Robert Nozick’s definition of philosophy there, in the contribution to this volume by C. W. Huntington.
7. Cf. Perrett, *An Introduction to Indian Philosophy*, 4.
8. In “Philosophy, Literature, Religion: Buddhism as Transdisciplinary Intervention,” I discuss Arthur Danto’s (“Philosophy and/as/of Literature,” 8) embrace of philosophy as a literary activity (“I cannot think of a field of writing as fertile as philosophy has been in generating forms of literary expression”), and attempt to expand his list of philosophical-literary forms to include those characteristic of Buddhist discourse.
9. Preston, Introduction, “Analytic Philosophy.”
10. Garfield, *Empty Words: Buddhist Philosophy and Cross-Cultural Interpretation*, 252.
11. *Ibid.*
12. Motzkin, “What Is Philosophy? Speech to the Graduating Class of 2000, The Hebrew University of Jerusalem,” 68; emphases in original.
13. Gadamer, *Truth and Method*, 386. For a fine discussion of Gadamer’s use of this expression in the context of “authentic conversation,” see Figal (“The Doing of the Thing Itself: Gadamer’s Hermeneutic Ontology of Language,” esp. 106ff). For references to the same Gadamerian position with specific application to Buddhist philosophy, see Avramides (“Engaging with Buddhism,” 548) and Garfield (“Engaging Engagements with Engaging Buddhism,” 583).
14. I am obliquely referring here (and in the title to this article), of course, to Thomas Nagel’s celebrated “deliberate effort to juxtapose the internal and external or subjective and objective views at full strength . . . [such that,] [i]nstead of a unified world view, we get the interplay of these two uneasily related types of conception, and the essentially incompletable effort to reconcile them” (Nagel, *The View From Nowhere*, 4).
15. I am reminded here of Arindam Chakrabarti and Ralph Weber’s “fourth stage” of what their book titles *Comparative Philosophy without Borders*; a stage “beyond comparative philosophy” which “would amount to just doing philosophy as one thinks fit for getting to the truth about an issue or set of issues, by appropriating elements from all philosophical views and traditions one knows of but making no claim of ‘correct exposition’, but just solving hitherto unsolved problems possibly raising issues never raised before anywhere” (Chakrabarti and Weber, *Comparative Philosophy without Borders*, 22).

BIBLIOGRAPHY

- Avramides, Anita. "Engaging with Buddhism." *Sophia* 57, no. 4 (2018): 547–58.
- Chakrabarti, Arindam, and Ralph Weber (eds.). *Comparative Philosophy without Borders*. London and New York: Bloomsbury, 2016.
- D'Amato, Mario, Jay L. Garfield, and Tom J. F. Tillemans (eds.). *Pointing at the Moon: Buddhism, Logic, Analytic Philosophy*. New York: Oxford University Press, 2009.
- Danto, Arthur C. "Philosophy and/as/of Literature." *Proceedings and Addresses of the American Philosophical Association* 58, no. 1 (September 1984): 5–20.
- Dostal, Robert J. (ed.). *The Cambridge Companion to Gadamer*. Cambridge: Cambridge University Press, 2002.
- Figal, Günter. "The Doing of the Thing Itself: Gadamer's Hermeneutic Ontology of Language." In *The Cambridge Companion to Gadamer*, edited and translated by Robert J. Dostal. Cambridge: Cambridge University Press, 2002.
- Gadamer, Hans-Georg. *Truth and Method*. Translation revised by Joel Weinsheimer and Donald G. Marshall. London and New York: Bloomsbury, 2013.
- Garfield, Jay L. *Empty Words: Buddhist Philosophy and Cross-Cultural Interpretation*. New York: Oxford University Press, 2002.
- . *Engaging Buddhism: Why It Matters to Philosophy*. New York: Oxford University Press, 2015.
- . "Foreword." In *Taking Back Philosophy: A Multicultural Manifesto*, by Bryan W. Van Norden. New York: Columbia University Press, 2017.
- . "Engaging Engagements with Engaging Buddhism." *Sophia* 57, no. 4 (2018): 581–90.
- Heidegger, Martin. *What Is Philosophy?* Translated by William Kluback and Jean T. Wilde. New York: Twayne Publishers, 1958.
- Motzkin, Aryeh Leo. "What Is Philosophy? Speech to the Graduating Class of 2000, The Hebrew University of Jerusalem." *Philosophia* 28 (2001).
- Nagel, Thomas. *The View From Nowhere*. New York: Oxford University Press, 1986.
- Olberding, Amy. "DoD named 'Wise and Good!'" *Department of Deviance*, September 16, 2016. Available at <http://departmentofdeviance.blogspot.com/2016/09/dod-named-wise-and-interesting.html>. Last accessed July 9, 2019.
- Perrett, Roy W. *An Introduction to Indian Philosophy*. Cambridge: Cambridge University Press, 2016.
- Preston, Aaron. "Analytic Philosophy." *The Internet Encyclopedia of Philosophy*. Available at <https://www.iep.utm.edu/analytic/>. Last accessed July 9, 2019.
- Stepien, Rafal K. "Orienting Reason: A Religious Critique of Philosophizing Nāgārjuna." *Journal of the American Academy of Religion* 86, no. 4 (2018): 1072–1106.
- . "Philosophy, Literature, Religion: Buddhism as Transdisciplinary Intervention." In *Buddhist Literature as Philosophy, Buddhist Philosophy as Literature*, edited by Rafal K. Stepien. Albany: State University of New York Press, forthcoming 2020.
- Tampio, Nicholas. "Not All Things Wise and Good are Philosophy." *Aeon* (2016). Available at <https://aeon.co/ideas/not-all-things-wise-and-good-are-philosophy>. Last accessed July 9, 2019.
- Van Norden, Bryan W. *Taking Back Philosophy: A Multicultural Manifesto*. New York: Columbia University Press, 2017a.
- . "Western Philosophy Is Racist." *Aeon* (2017b) <https://aeon.co/essays/why-the-western-philosophical-canon-is-xenophobic-and-racist>. October 31, 2016. Last accessed July 9, 2019.

Doing Buddhist Philosophy

C. W. Huntington, Jr.

HARTWICK COLLEGE

We are getting pretty good at debating on Madhyamaka but this is not the real understanding of emptiness, for it is bound by conceptual elaborations (*prapañca*, *sproś pa*). We could even defeat a person who had realized emptiness! Such a person would be able to see through conceptual elaborations but could not answer our questions.

– Gen Nyi-ma¹

The *Journal of Indian Philosophy* commenced publication almost fifty years ago, in October of 1970, and immediately established itself as a premier academic setting for discussion of South Asian intellectual history. The first issue included a three-page editorial that defined the journal's objectives in the following terms:

The field of our contributions will be bound by the limits of rational inquiry; we will avoid questions that lie in the fields of theology and mystical experience. Our method will be, in a very general sense, analytical and comparative, and we will aim at rigorous precision in the translation of terms and statements. Our aim will be to attract professional philosophers rather than professional internationalists.²

The editors go on to lament what they see as the shortcomings of a discipline dominated for too long by scholars trained only, or primarily, as Sanskritists:

The philologists, interested always in the oldest texts of their chosen language, produced good translations of the ancient texts of Vedism and Hinduism. Unfortunately, the public applied to those texts the term philosophy, a term justifiable in such an application only if taken in its etymological sense. The *Upaniṣads* and the *Gītā* exhibit a 'love of wisdom', even as the Gospel of John and the Book of Revelations exhibit the same love. But their aim is a wisdom beyond reason and logic.³

The editorial concedes that such texts—and such wisdom—may have great human value, but it has nothing to do with philosophy as that is presently understood in the West: "For philosophy in the English-speaking world since somewhat before the time of the French revolution has expended its intellectual effort in fields where some degree of certainty may be achieved without recourse to religious or ontological commitment."⁴ Philosophy in this sense of the word aspires to certainties not founded on any a priori faith or belief, but rather on universally acceptable, rationally justified premises. It has no pretensions to anything so grand or nebulous as "wisdom."

Perhaps the field of our labor grows poorer crops than the fields of the theologians and the mystics. Certainly our food lacks the ambrosia that they find in theirs. But it is a field that many farmers have plowed and we wish to emphasize that Indian farmers have been as diligent as Europeans and Americans. They too have spent lifetimes defining the epistemological process, framing systems of logic, analyzing language, speculating on the relationship between syntax and semantics.⁵

Having discussed the general outlines of their purpose and method, in the final paragraph the editors define the journal's mission in more exact terms:

Specifically, we intend to offer a medium where philosophers, using the word to mean those who pursue rationally demonstrable answers to meaningful questions, both Indian and Western, may converse. . . . We shall guide the conversation toward questions that seem to us important. But we will remember the value of precision, whatever the question on which it is achieved, and we will keep in mind the fact that the most important questions may be, for the time at least, insoluble by the methods we have chosen.⁶

In the case of Buddhism this approach to Indic texts has a long history. Efforts to rationalize or naturalize Buddhist doctrine have characterized a good deal of writing both in the West and among Western-influenced intellectuals writing in Asia. Such writing is one element of what has been referred to as "Protestant Buddhism" or "Buddhist modernism"—a progressive movement with roots in nineteenth-century colonial Sri Lanka that constituted from the beginning an effort to make Buddhist thought and practice palatable to modern Western sensibilities by looking for ways to make it conform to enlightenment and scientific ideals of rationality and empiricism. A characteristic feature of Buddhist modernism is its tendency to selectively cull the Indian texts so as to view Buddhism more as a "philosophy" than a "religion." Walpola Rahula's *What the Buddha Taught* is a classic of the genre; it exemplifies this strategy not only in the interpretive and exegetical chapters of the book, but also in the English translations of Pāli scriptures included as appendices, which have been bowdlerized so as to eliminate all reference to divinity, magic, or supernatural attainments (*siddhis*). As Robert Sharf puts it in speaking of Japan's New Buddhism, "'indigenous' traditions were reconstituted so as to appropriate the perceived strengths of the Occident . . . Buddhist apologists sought to secure the integrity of Buddhism by grounding it in a trans-cultural, trans-historical reality immune to the relativist critique."⁷ A similar observation has recently been brought to bear on the effort by a number of Western scholars to define and legitimize Nāgārjuna's *Madhyamaka* in strict rationalist terms.⁸

PHILOSOPHY AS LOVE OF REASON

The definition of philosophy advocated by the editors of the *Journal of Indian Philosophy* is by no means exceptional. "The word philosophy means the love of wisdom," writes Robert Nozick in *The Nature of Rationality*, "but what philosophers really love is reason. . . .":

They formulate theories and marshal reasons to support them, they consider objections and try to meet them, they construct arguments against other views. Even philosophers who proclaim the limitations of reason—the Greek skeptics, David Hume, doubters of the objectivity of science—all adduce reasons for their views and present difficulties for opposing ones. Proclamations or aphorisms are not considered philosophy unless they also enshrine and delineate reasoning."⁹

To advocate for a particular way of doing philosophy is to express an opinion about what constitutes valid philosophical method and what purpose is served by philosophizing in this or that way. As Nozick makes clear, when philosophy is grounded in love of reason its method is argument and its purpose is to compel others to accept one's own opinion, belief, or view, which is itself construed as the outcome of right reason:

The terminology of philosophical art is coercive: arguments are *powerful* and best when they are *knockdown*, arguments *force* you to a conclusion, if you believe the premises you *have to* or *must* believe the conclusion, some arguments do not carry much *punch*, and so forth. A philosophical argument is an attempt to get someone to believe something, whether he wants to believe it or not. A successful philosophical argument, a strong argument, *forces* someone to a belief.¹⁰

The catch here, of course, lies in the protasis: *if you believe the premises*. And there is a single premise that underwrites the entire enterprise: In order to do philosophy in this way everyone involved must maintain an unquestioned faith in the value of rationality as the only reliable guide to any truth worth knowing. Without this faith, philosophical argumentation has no clout.

Though philosophy is carried on as a coercive activity, the penalty philosophers wield is, after all, rather weak. If the other person is willing to bear the label of "irrational" or "having the worse argument" then he can skip away happily maintaining his previous belief. He will be trailed, of course, by the philosopher furiously hurling philosophical imprecations: "What do you mean, you're willing to be irrational? You shouldn't be irrational because..." And although the philosopher is embarrassed by his inability to complete this sentence in a noncircular fashion—he can only produce reasons for accepting reasons—still, he is unwilling to let his adversary go.¹¹

"Wouldn't it be better," Nozick suggests, "if the philosophical arguments left the person no possible answer at all. . . ?" Silence is certainly preferable to unreason. Or even better:

Perhaps philosophers need arguments so powerful they set up reverberations in the brain: if the person refuses to accept the conclusion, he *dies*.¹²

The problem for philosophers is that people are not always able, willing, or even interested in offering a reasoned defense for what they know to be true from their own experience. As my high school girlfriend once remarked, after listening patiently while I unspooled an elaborate rationalization for some bit of typically inconsiderate behavior: *you may win every argument, but you're still wrong.*

Arguing to prove one's point and so force one's interlocutor to accept one's own view, Nozick concludes, is an uncivil way of doing philosophy;¹³ he proposes explanation as an alternative method more in harmony with philosophy's quest to achieve mutual understanding through the congenial exchange of ideas. However, setting aside the issue of incivility, there are other matters that warrant our attention in this context. As Nozick's discussion reveals (and as I have already noted), both argument and explanation are rooted in a single, a priori commitment to reason as the final arbiter of any truth worth knowing. Moreover, regardless of whether it is achieved through argument or explanation, the purpose of philosophizing in this way is to present and defend one's own theories, opinions, beliefs, and views as the best possible theories, opinions, beliefs, and views. As Nozick has it, in characterizing his own preferred approach to philosophizing: "On the view presented here, philosophical work aspires to produce a highest ranked view, at least an illuminating one, without attempting to knock all other theories out as inadmissible."¹⁴ Philosophy of this kind—whether argumentative or exegetical—is all about holding and defending views, and what is of primary meaning and value turns out to be, for the most part, views that have to do with "defining the epistemological process, framing systems of logic, analyzing language, speculating on the relationship between syntax and semantics" and etc. This way of doing philosophy is a profoundly theoretical affair, and its altogether modest certainties—"rationally demonstrable answers to meaningful questions"—are at once both precise and entirely speculative and abstract.

Like Nozick, I cannot see how disputes about the nature of the philosophical enterprise—"how to satisfy the desire to start philosophy in a neutral way"¹⁵—can ever be resolved. Nor—again, like Nozick—do I see this as a particularly troubling problem. What counts as legitimate philosophical purpose and method is and should be (in my view) open to discussion. In the words of the editorial I quoted above: "[W]e will keep in mind the fact that the most important questions may be . . . insoluble by the methods we have chosen."¹⁶

PHILOSOPHY AS LOVE OF WISDOM

What, then, are the most important questions for Buddhist philosophers, and how are they addressed?

Rather than deal in generalizations, I suggest we look closely at the work of one particular individual: the Mahāyāna philosopher Nāgārjuna (c. 150–250 CE), arguably the most renowned author in the Indian Buddhist tradition. As it happens, the third issue of the *Journal of Indian Philosophy* included an article by Frederick Streng, titled "Buddhist Doctrine as Religious Philosophy," that directly addresses Nāgārjuna's central composition, the

Mūlamadhyamakakārika (or *Madhymakaśāstra*). In this article Streng proposed that the text was informed by two distinct methodological assumptions:

The first is that the statements are made in the context of religious philosophy; that is, the highest purpose for formulating any statement is soteriological, not speculative. Thus "truth" can refer to a development of an attitude as well as a judgment about a proposition. The second is that the meaning or significance of the most profound religious statements includes reference to, though not limited to, commonly recognized experience of reality.¹⁷

He adds, "To the extent that the latter assumption is correct we can understand the religious significance of the statements without claiming that we have attained the highest spiritual insight."¹⁸

And so, in a single stroke, Streng appears to have subverted—or at least reframed—the stated intentions of the journal's editorial board to "avoid questions that lie in the fields of theology and mystical experience." This was inevitable, it seems to me, since Indian intellectuals writing in Sanskrit have traditionally been preoccupied by precisely these sorts of questions. In South Asia, concern with theology and mystical experience is like the ceaseless drone of the tanpura in Indian classical music—it permeates the atmosphere in which discourse on issues of epistemology, ontology, and ethics takes place. By obviating the need for any claim to "spiritual insight," Streng has legitimized his interest in theology and mysticism; as a modern analytically inclined philosopher he is free to discuss these dimensions of the Sanskrit texts without compromising his faith that rationally demonstrable truth is the only kind of truth worth pursuing.

Be this as it may, Streng nevertheless introduced the expression "religious philosophy" and thereby implicitly sanctioned a properly philosophical enterprise grounded in a methodology and purpose that do not prioritize the truth claims of reasoned argumentation—much less certify it as the only acceptable model of philosophical discourse. According to Streng, the highest meaning and value for the practitioners of so-called religious philosophy is not found in "the conceptual content of statements, nor the accuracy with which one identifies notions with their assumed objective referents," but rather in "a difference of perspective of the user of truth statements: whether he is attached to, or free from, the distinctions he makes for the purposes of communication."¹⁹ "Absolute truth is not only a definition of 'the way things are'; it is a situation of freedom, of health, of joy."²⁰

In other words, for Nāgārjuna, genuine understanding of emptiness (*śūnyatā*) is not about holding or defending a theoretical position, nor is philosophical work a matter of producing "the highest ranked view." Rather, to genuinely understand the truth of emptiness is to be free from attachment to the method and goals of reasoned argumentation and therefore no longer susceptible to the frustration and anxiety associated with this way of thinking,

speaking, and writing. This “attitude”—to use Streng’s term—constitutes a “transformation in self-awareness”,²¹ the freedom it embodies is captured above all in the Sanskrit word *prajñā*, commonly translated into English as “wisdom.”

THE SOTERIOLOGICAL METHOD AND PURPOSE OF BUDDHIST TEXTS

Where reason is employed in Nāgārjuna’s philosophy it is subordinated to a soteriological end conceived as freedom from suffering. Suffering (*duḥkha*)—defined as a fundamental, existential dis-ease—is rooted in both intellectual and emotional soil; it manifests in ignorance or misunderstanding of the nature of reality (*avidyā*) and in a variety of negative emotions (*kleṣa-s*)—notably “thirst” or “clinging” (*trṣṇa/upādāna*). The absence of suffering—*nirvāṇa*—is similarly conceived as two-fold, characterized by the presence of wisdom (*prajñā*) and compassion (*karuṇā*). Within Nāgārjuna’s *Madhyamaka*, reasoned argument serves as a methodological strategy (*upāya*) for the cultivation of wisdom, but argumentation has in itself no ultimate value and is moreover explicitly recognized as risky medicine, precisely because it can so easily promote, rather than diminish, attachment to views, opinions, beliefs, and other conceptual abstractions (*prapañca*). In doing so, it can as easily foster intellectual hubris as compassionate wisdom:

Emptiness is proclaimed by the conqueror as the unmooring of all views; those who hold emptiness as a view are deemed incurable.²²

Given this inherent danger, the use of rational argumentation is traditionally circumscribed within a Buddhist context that includes the practice of morality and meditation. These three—morality (*śīla*), meditation (*samādhi*), and “wisdom” (*prajñā*)—constitute the ancient components of the Buddhist soteriological path as described, for example, in Buddhaghosa’s *Visuddhimagga*. Wisdom is further broken down into three developmental stages: (1) familiarizing oneself with basic Buddhist doctrine (*śrotamayī-prajñā*); (2) critically reflecting on this doctrine (*cintamayī-prajñā*); and (3) actively cultivating what has been understood conceptually through a variety of ethical and ritual practices (*bhāvanāmayī-prajñā*) that culminate in the “transformation of self-awareness” mentioned above. Strictly speaking, the use of reason is confined to the first and second stages, but primarily to the second. In this respect, although Nāgārjuna makes an appeal to rational argument, as a Buddhist (“religious”) philosopher, he does so within a broader context where argument is not, in and of itself, considered an entirely reliable strategy; nor is it the fundamental or sole defining characteristic of his philosophy.

UNDERSTANDING NĀGĀRJUNA OUTSIDE ANY TRADITIONAL BUDDHIST CONTEXT

In a piece written for *The New Yorker* on Stanley Kubrick’s enigmatic masterpiece *2001: A Space Odyssey*, Dan Chiasson observes that the film baffled early audiences because it had yet to create the taste it required to be appreciated. Nevertheless, repeated viewing of Kubrick’s

work forged its own context for interpretation. “You didn’t solve it by watching it a second time,” Chiasson comments, “but you did settle into its mysteries.”²³ One might similarly argue that Nāgārjuna’s writing has yet, in the West, to forge its own interpretive context. Certainly Western scholars have had some difficulty settling into its mysteries.

Nāgārjuna situated himself resolutely within the Mahāyāna Buddhist lineage, as a commentator on the *prajñāpāramitā* literature; his highest commitment was to a soteriological purpose that explicitly repudiated all interest in propositional truth claims. Modern scholars and professional philosophers engaged with his writing have for the most part been bound to a set of goals and methods shared by the academic community; we deploy the tools of philology, historiography, and rational exegesis, positioning ourselves securely outside any Buddhist tradition while attempting to speak for Nāgārjuna in a way that does not compromise whatever meaning and value his work may possess on its own terms. The extent to which such an undertaking can hope to succeed is debatable. At best it’s like the difference between, on the one hand, telling a joke, and on the other, explaining the same joke or trying to prove that it’s funny; no matter how adroit the explanation or how convincing the proof, no one’s going to laugh.

In any event, at this point Anglo-American analytic philosophy has emerged as the only widely accepted model for understanding Nāgārjuna as *philosophy*. In practice this has meant that his writing is viewed almost exclusively as a species of rational argumentation. In recent years forms of non-classical logic—in particular paraconsistent logic and dialetheism—have attracted the attention of scholars in the field, but unexamined assumptions about the priority of reason in Nāgārjuna’s writing persist.

No doubt, reading Nagarjuna through an analytic lens has the advantage of making him seem appealing to modern intellectuals and in particular to analytical philosophers, but it has tended to marginalize his profound distrust of theoretical, propositional truth and his central and overriding concern with soteriology. Nāgārjuna’s writing, as a form of religious philosophy, is not designed to advocate for one view over another. Rather—as Streng pointed out—its objective is to provoke a fundamental shift in one’s attitude toward thinking and reasoning, a shift in attitude that goes hand in hand with the development of nonattachment and compassion. In this respect, analytic philosophy may not be the best model for appreciating what Nāgārjuna is about, and indeed, other models have been proposed. Nāgārjuna’s approach to philosophizing bears comparison with Rorty’s “edifying philosophy,” with Wittgenstein’s notion of philosophy as therapy, and with Derrida’s deconstruction. It shares the general concern of Continental philosophers with the unification of theory and practice and the goal of cultural and personal transformation, and is in this respect in accord with Pierre Hadot’s characterization of ancient Greek philosophy as a genre of spiritual exercise aimed toward a “transformation of self-awareness” marked by freedom from anxiety (*ataraxia*). There are especially interesting parallels between Nāgārjuna and Pyrrhonian skepticism. There are as well important affinities between Nāgārjuna

and contemporary Western ethicists like Martha Nussbaum, Alasdair MacIntyre, Cora Diamond, and Christopher Hamilton, all of whom maintain in one way or another that its preoccupation with reason and rational discourse makes analytic philosophy incapable of engaging with the deeper recesses of human moral experience. "As it seems to me," writes Hamilton,

the truth that a poem or other work of art can tell about the world seems to leave philosophy behind and to show its inadequacy. This is no doubt connected with the typical failings of (much) philosophy: its conviction that all that is true can be stated clearly and without self-contradiction; its hostility to imagination and its literal-mindedness; its reduced conceptual repertoire; and so on.²⁴

By shifting the conversation from talk of logic and reason to discussion of literature, poetry, and metaphor, Hamilton directs our attention to a form of lived truth that cannot be captured through explanation of or arguments for a particular point of view—even if (or especially if) the truth represented in that point of view is considered in some way to transcend the arguments constructed to reveal it. In this respect he offers the possibility of a radically different approach to interpretation of Nāgārjuna's use of language in the service of soteriological ends.

Understood from Hamilton's perspective, the appearance of rational or logical argumentation in Nāgārjuna's writing can be interpreted not as a means to secure propositional truth (e.g., the theoretical truth of emptiness), but rather as a literary device—a *semblance* of reason—that serves to illustrate an attitude of nonattachment with respect to the distinctions fostered by literal, discursive thought and associated forms of discourse. Georges Dreyfus seems to suggest as much when he writes,

Because emptiness requires the overcoming of any reification and thus of any differentiation, it cannot be understood conceptually in the same way as other points of Buddhist doctrine such as impermanence. Applying a concept will not do, for emptiness is not a thing or an object that can be captured by distinguishing it from other things. Hence, it cannot be said (i.e. directly signified); it can only be skillfully shown (i.e. indirectly indicated).²⁵

This way of reading Nāgārjuna—as the skillful display of a particular attitude of non-clinging—is in broad accord with traditional Mahāyāna Buddhist practice, where the *Madhyamakaśāstra* is studied in the context of a community of like-minded *religiosi* engaged in ritual activities of various kinds, including forms of devotional prayer and meditation, and governed by a shared commitment to a set of ethical principles—a radically different environment from the arid, hyper-rationalist atmosphere of a seminar room at Oxford or Berkeley. Philosophical argumentation is framed within this traditional Buddhist context not as a strategy for establishing some recondite version of Nozick's "highest ranked view," but rather in such a way as to exhibit the futility of reason's endless, tortured striving to overcome its own self-imposed limitations. I want to suggest that

for someone properly situated within this ritual frame, Nāgārjuna reads more like literary fiction or poetry than analytic philosophy. Which is to say, first of all, that for such a reader Nāgārjuna's writing is inextricably caught up in its own fictive ontology.²⁶ But there is yet another sense in which the *Madhyamakaśāstra* functions more like a work of narrative literature than a logical tract. I am thinking here of Jerome Bruner's point that narratives "recruit the reader's imagination" by offering "'performances' of meaning rather than actually formulating meanings themselves."²⁷

Understood within a traditional Mahāyāna Buddhist milieu, it becomes clear that Nāgārjuna is not ensnared in the fiercely theoretical contest of reason versus unreason, and is not bound by any version of its rules. Rather, his writing evokes an empty, insubstantial world—an utterly contingent "cultural universe"²⁸—to be imaginatively entered and inhabited, so that we might learn to live more authentically in the face of impermanence and death. It seems to me that this is a useful and appropriate perspective from which to assess the significance of the *catuṣkoṭi*:

When all things are empty,
What is endless and what is not?
What is both endless and not endless?
What is neither endless nor not endless?

What is one thing and what is another?
What is immortal and what must die?
What is both, and what is neither?²⁹

The appearance of rational argumentation has an important role to play in Nāgārjuna's soteriological project—just as it does in the famous scene from Fyodor Dostoyevsky's novel *The Brothers Karamazov*, where, in response to the force of the Grand Inquisitor's subtle logic, Jesus offers a single, silent kiss; or in Harper Lee's *To Kill a Mockingbird*, where Atticus Finch argues in vain before the Maycomb County court on behalf of his client. In these groundless worlds—where words take their meaning not through reference to extra-linguistic, intrinsically real people, things, and places, but rather only in relationship to each other—what is proven and what is refuted? What is justified or forgiven and what is not? What innocent life is either saved or lost?

It is as if a skilled magician or a magician's apprentice were to conjure up a large group of people at a crossroads, and then, having conjured them, he were to make them disappear. What do you think, Subhūti, has anyone been slain or killed, destroyed or made to disappear? In just this way a bodhisattva leads measureless, countless beings to nirvana, and yet there is not a single being led to nirvana nor anyone who does the leading.³⁰

PREACHING THE DHARMA/DOING BUDDHIST PHILOSOPHY

In a review of Francisca Cho's book *Seeing Like the Buddha*, Lina Verchery writes:

The book is a welcome expansion on a compelling claim Cho previously made in an essay for *Imag(in)*

ing Otherness: “I turn to Buddhism as the *source* of my theory making rather than as the *object* of ideological clarification. . . . Buddhism is the instrument of my analysis rather than its target.” In *Seeing Like the Buddha*, Cho models what such a project—namely, the deployment of Buddhism as a source rather than an object of theoretical investigation—might look like. This is valuable for . . . as many have noted, the traditional divide between “theory” and “data”—which makes the former the exclusive purview of scholarly authorities and relegates informants and religious “insiders” to the latter—has perpetuated precisely the kinds of problematic power dynamics that so much work in postcolonial, post-orientalist Buddhist studies, and in the humanities more broadly, has sought to dismantle.³¹

To take Nāgārjuna’s writing as the object of theoretical investigation is to prioritize the methods and goals of modern Western academic discourse. But to write as he writes—to make use of the language of reason while no longer being bound by it—is to subvert those same standards. Streng made it possible for analytic philosophers to acknowledge Buddhist religious philosophy as philosophy and to take an active interest in it, but he did not clear the way for such philosophy to be practiced within the pages of academic journals. As Jay Garfield has pointedly observed, “doing the history of Buddhist philosophy” is one thing; “doing Buddhist philosophy” is another.³² For a contemporary Buddhologist to argue for any particular understanding of Nāgārjuna’s project—how he writes and for what soteriological purpose—falls well within the accepted methodological parameters of academic scholarship. For that same person to attempt to take up in earnest Nāgārjuna’s philosophical task—to write as I claim he did, for the express purpose of exhibiting an attitude of nonattachment to the distinctions fostered by rational, discursive language—is another thing altogether. What is jettisoned in such writing is not reason *per se*, but rather an implicit loyalty to reason, an undergirding faith in rationality as the essential guide to any truth worth knowing. What remains when this faith has gone is what I refer to as the mere *appearance* of rational argumentation.³³ To abandon faith in reason and do philosophy as Nāgārjuna does means to write in a way that favors figurative, tropological language over a style of discourse that is forthrightly analytic, literal, or referential. Writing like this means that any overt appeal to reason is necessarily, even if not obviously, ironic. That is to say, in Nāgārjuna’s hands rational argumentation maintains the pretense, but not the substance, of intrinsic authority. Only when understood in this way can his writing successfully communicate its peculiar brand of soteriologically pragmatic, fictional truth.

Not so long ago I wrote an article in which I intentionally set out to blur the line between talking about what Nāgārjuna does with language and attempting to actually do it. I submitted the essay to the *Journal of Indian Philosophy*. Following protocol, the editors of the journal solicited comments from anonymous reviewers and forwarded those comments on to me. Here is an excerpt from one reviewer’s response:

I have some reservations about the idea that an academic journal article should be written so as to provide a “catalyst for inner transformation” in the reader. Frankly speaking, readers of academic journals are not usually looking for Dharma-preaching in those journals, and I personally think they would be misguided to do so: there exist other more contextually appropriate venues to perform the beneficial function of Dharma-teacher. The article’s present format makes it even difficult for colleagues to engage with its content (how to do so, if rational engagement has been intentionally substituted with literary fiction?) There is a risk for the article to become entirely self-referential and impermeable to academic discussion.

Iris Murdoch once wrote that in order to understand a philosopher we need to look for what it is he or she fears.³⁴ What this anonymous reviewer seems to fear are consequences perceived to follow from the substitution, within an academic context, of religious commitment for fidelity to reason. The article was eventually accepted and published,³⁵ but the reviewer’s comments left me wondering what exactly I was hoping to accomplish in my attempt to carry forward Nāgārjuna’s project within the pages of a scholarly journal.

What my reviewer meant by “Dharma-preaching” is presumably some kind of homily intended primarily for spiritual edification; in essence, a sermon that lacks the requisite critical engagement with original Sanskrit sources. But that is not necessarily a fair characterization of what it means to preach or teach the Dharma. Streng viewed Nāgārjuna’s “Dharma-teaching”—an activity he identified as an explicitly Buddhist genre of “religious philosophy”—as a means “to develop an attitude which frees one from attachment to any single idea or experience without rejecting all ideas or experiences for some projected opposite.”³⁶ To teach or preach the Dharma, then—regardless of the context in which it is done—does not require that one forgo critical engagement with the classical texts, but rather only that one write in such a way as to abandon unquestioning confidence in reason without embracing its opposite, unreason or irrationality. Writing like this is at the core of Nāgārjuna’s soteriological philosophy. Whatever one might think about such a philosophy, it leaves us with plenty to discuss.

NOTES

1. Quoted in Dreyfus, *The Sound of Two Hands Clapping: The Education of a Tibetan Buddhist Monk*, 333.
2. JIP Editorial Board, “Editorial,” 1.
3. *Ibid.*, 2.
4. *Ibid.*
5. *Ibid.*
6. *Ibid.*, 3.
7. Sharf, “Buddhist Modernism and the Rhetoric of Meditative Experience,” 259 & 268.
8. Stepien, “Orienting Reason: A Religious Critique of Philosophizing Nāgārjuna.”
9. Nozick, *The Nature of Rationality*, xi.

10. Nozick, *Philosophical Explanations*, 4.
11. Ibid.
12. Ibid.
13. "Is that a nice way to behave toward someone?" (Ibid., 5).
14. Ibid., 24.
15. Ibid., 19.
16. JIP Editorial Board, "Editorial," 4.
17. Streng, "The Buddhist Doctrine of Two Truths as Religious Philosophy," 263.
18. Ibid.
19. Ibid., 264.
20. Ibid., 270.
21. Ibid., 269.
22. Vaidya *Madhyamakaśāstra of Nāgārjuna with the Commentary Prasannapadā by Candrakīrti*, 108 (*Madhyamakaśāstra* 13.7): *śūnyatā sarvadṛṣṭinām proktā niḥśaranam jinaiḥ/ yeṣām tu śūnyatādrṣṭis tām āśādhyān bahāṣire//*
23. Chiasson, "'2001: A Space Odyssey': What It Means, and How It Was Made."
24. Hamilton, "Review of *The Aesthetics of Argument*, by Martin Warner."
25. Dreyfus, *The Sound of Two Hands Clapping*, 241.
26. If, as Nāgārjuna seems to argue, words do not take their meaning through reference to an extra-linguistic reality, then this applies equally to his own words. Such reflexivity is precisely the sense in which emptiness is itself empty. See, e.g., *Madhyamakāvāṭāra* 6.186 (Huntington, *The Emptiness of Emptiness*, 180), and *Madhyamakaśāstra* 22.11: "One should not definitively assert 'empty', 'not empty', both or neither; [the word 'empty'] is used simply as a heuristic." (Vaidya, *Madhyamakaśāstra of Nāgārjuna with the Commentary Prasannapadā by Candrakīrti*, 192–93: *śūnyamiti na vaktavyam aśūnyamiti va bhavet/ ubhayam nobhayam ceti prajñāpyartham tu kathyate//*)
27. Bruner, *Actual Minds, Possible Worlds*, 39.
28. I've borrowed this phrase from Georges Dreyfus, who eloquently documents the complex strategies that have evolved among Tibetan Buddhists as they attempt to balance the demands of reason and soteriology in the study of Nāgārjuna. As he puts it:

[Tibetan Buddhist] scholastics see their activities not just as a search for knowledge for its own sake but as central to the religious pursuits of the larger tradition of which scholasticism is part. That religious relevance should be understood properly: scholastic studies should be seen not as directly preparing scholars for personal practice but as helping them to construct a cultural universe in which practice makes sense and to develop confidence in the soteriological possibilities of such a universe. (*The Sound of Two Hands Clapping*, 164–65)
29. Vaidya *Madhyamakaśāstra of Nāgārjuna with the Commentary Prasannapadā by Candrakīrti*, 235 (*Madhyamakaśāstra* 25:22–23): *śūnyeṣu sarvadharmeṣu kimanantam kimantavat/ kimanantam antavacca nānantam nāntavacca kim// kim tadeva kimanyat kim śāśvatam kimaśāśvatam/ aśāśvatam śāśvatam ca kim vā nobhayamapyataḥ 'tha//*
30. Vaidya, *Astasāhasrikā Prajñāpāramitā with Haribhadra's Commentary Called Āloka*, 10–11 (*Astasāhasrikā Prajñāpāramitā*): *yathāpi nāma subhūte dakṣo māyākāro vā māyākārāntevāsi vā caturmahāpāthe mahāntam janakāyam abhinirmimite/ abhinirmāya tasyaiva mahato janakāyasyāntardhānam kuryāt/ tatkiṃ manyase subhūte api nu tatra kenacitkaścid dhato vā mṛto vā nāśito vā antarhito vā? [...] evameva subhūte bodhisattvo mahāsattvo 'prameyānasamkhyeyān sattvān parinirvāpayati/ na ca sa kaścitsattvo yaḥ parinirvṛto yena ca parinirvāpito bhavati/*
31. Verchery, "Review of Cho, Francisca, *Seeing Like the Buddha: Enlightenment through Film*"; emphases in original.
32. Garfield, *Engaging Buddhism: Why It Matters to Philosophy*, 320ff.
33. In referring to the "appearance" or "illusion" of rational argumentation I have in mind a central distinction in Madhyamaka texts between "conventional truth" (*saṃvṛti-satya*) and "ultimate truth" (*paramārtha-satya*).
34. Murdoch, *The Sovereignty of Good*, 71.
35. Huntington, "Nāgārjuna's Fictional World."
36. Streng, "The Buddhist Doctrine of Two Truths as Religious Philosophy." Cf. Dreyfus: "To understand emptiness requires that one free oneself from any reification, negative as well as positive" (*The Sound of Two Hands Clapping*, 240).

BIBLIOGRAPHY

Buddhaghosa. *Visuddhimagga or The Path of Purification*. Translated by Bhikkhu Ñānamoli. Onalaska: BPS Pariyatti Editions, 1999.

Bruner, Jerome. *Actual Minds, Possible Worlds*. Cambridge, MA: Harvard University Press, 1996.

Chiasson, Dan. "'2001: A Space Odyssey': What It Means, and How It Was Made." In *New Yorker*, April 23, 2018.

Dreyfus, Georges. *The Sound of Two Hands Clapping: The Education of a Tibetan Buddhist Monk*. Berkeley & Los Angeles: University of California Press, 2003.

Garfield, Jay. *Engaging Buddhism: Why It Matters to Philosophy*. Oxford: Oxford University Press, 2015.

Hamilton, Christopher. "Review of *The Aesthetics of Argument*, by Martin Warner." *Australasian Journal of Philosophy* 95, no. 3. Published online September 2, 2016. <https://doi.org/10.1080/00048402.2016.1225109>.

Huntington, C. W. *The Emptiness of Emptiness*. Honolulu: University of Hawaii Press, 1989.

Huntington, C. W. "Nāgārjuna's Fictional World." *Journal of Indian Philosophy* 46, no. 1 (2018): 153–77. Available at <https://doi.org/10.1007/s10781-017-9340-4>.

JIP Editorial Board. "Editorial." *Journal of Indian Philosophy* 1, no. 1 (1970): 1–3.

Murdoch, Iris. *The Sovereignty of Good*. London: Routledge & Kegan Paul, 1970.

Nozick, Robert. *Philosophical Explanations*. Cambridge: Harvard University Press, 1981.

Nozick, Robert. *The Nature of Rationality*. Princeton: Princeton University Press, 1993.

Rahula, Walpola. *What the Buddha Taught*. New York: Grove Press, 1969.

Sharf, Robert. "Buddhist Modernism and the Rhetoric of Meditative Experience." *Numen* 42, no. 3 (1995): 228–83.

Stepien, Rafal. "Orienting Reason: A Religious Critique of Philosophizing Nāgārjuna." *Journal of the American Academy of Religion* 86, no. 4 (2018): 1072–1106. Available at <https://doi.org/10.1093/jaarel/lfy021>.

Streng, Frederick. "The Buddhist Doctrine of Two Truths as Religious Philosophy." *Journal of Indian Philosophy* 1, no. 3 (1971): 262–71.

Vaidya, P. L. (ed.). *Madhyamakaśāstra of Nāgārjuna with the Commentary Prasannapadā by Candrakīrti*. Buddhist Sanskrit Texts—No. 10. Darbhanga: Mithila Institute of Post-Graduate Studies and Research in Sanskrit Learning, 1960a.

Vaidya, P. L. (ed.). *Astasāhasrikā Prajñāpāramitā with Haribhadra's Commentary Called Āloka*. Buddhist Sanskrit Texts—No. 4. Darbhanga: Mithila Institute of Post-Graduate Studies and Research in Sanskrit Learning, 1960b.

Verchery, Lina. "Review of Cho, Francisca, *Seeing Like the Buddha: Enlightenment through Film*." H-Buddhism, H-Net Reviews. May, 2018. Available at <http://www.h-net.org/reviews/showrev.php?id=51178>.

Decolonizing the Buddhist Mind

Mattia Salvini

MAHIDOL UNIVERSITY, SALAYA CAMPUS
INTERNATIONAL BUDDHIST COLLEGE, SADAO CAMPUS

And so we distort the arts themselves to curry favor for them with the machines. These machines are the inventions of Westerners, and are, as we might expect, well suited to the Western arts. But precisely on this account they put our own arts at a great disadvantage.

– *In Praise of Shadows*, Jun'ichirō Tanizaki¹

In this brief reflection on the present state of research on Buddhist philosophy, I will include observations from my own experience as an educator and administrator, and, primarily, as a student of Buddhist thought. My purpose is to imagine a better direction for the integration of Buddhist philosophy (and philosophers) in a contemporary university setting.

The borrowed title of this essay (Ngūgī wa Thiong'o, 1986) is meant to highlight an intellectual and institutional problem: there is, to the best of my knowledge, no clear institutional space where Buddhist philosophy could be the center and life-force of one's enquiry, instead of being the periphery or object. I do not see institutional spaces which would *explicitly* support someone's work as a Buddhist philosopher: I see how a modern university may support an analytical philosopher writing about Buddhism, or a historian, or a social scientist exploring Asian realities. I also see how someone may remain in one of these contemporary institutional spaces and yet, through some expedient, retain Buddhist philosophizing as one's primary interest (indeed I see this rather well). Yet this is far from an ideal situation, and is after all an expedient. While I very much wish to be proven wrong, my knowledge of the field suggests that Buddhist thought remains a periphery: a "colony," to borrow a metaphor whose historical undertones are here not entirely irrelevant (although without a wish to be too polemical).

We should notice from the start that Buddhist philosophy works within a framework that is somewhat different from contemporary (dominant) Western modes of philosophizing; i.e., it is primarily an exegetical undertaking, meant to unravel the thought of the founder, i.e., the Buddha, and of the most esteemed among the Buddha's followers. This may perhaps bring it closer to theology—in some ways—and yet calling it theology brings in potentially more problems and misunderstandings than it is meant to solve, as Buddhist thought is arguably not theological (pardon the sweeping generalization). Being primarily exegetical is in this context compatible with independent reflection on the nature of reality (*tattva*), with a privileged role for reason/proper reasoning (*yukti*) and with the ample use of logical grounds (*hetu*) to prove or reject specific theoretical positions. How this compatibility works is, however, a topic beyond the scope of the present essay, as it varies, even considerably, in different Buddhist systems and authors.

Whether Buddhist philosophy is a perfect fit for a Western philosophical framework is, however, not so crucial, as rather obviously I am here using the term "philosophy" in a flexible enough way to include what in Sanskrit would be called "Established Conclusion" (*siddhānta*), "Instrument of seeing/View" (*darśana*), "Position" (*sthiti*), etc. Since within Western culture itself "philosophy" has meant many different types of intellectual activities for different people, using the term to broadly reflect the sense of these Sanskrit key-words seems to me acceptable, given some remarkable similarities that have not passed unnoticed through the ages. But if someone may find it more agreeable, I would not object to substituting the word "philosophy" everywhere in this essay (even when it refers to Western thought) with *siddhānta*, and the word "philosopher" with *siddhāntin*. It may in fact be worthwhile to step into the looking glass of using a Buddhist category to parse a Western type of intellectual endeavor, rather than the familiar vice versa.

The absence of a Buddhist voice from the stage of modern academic institutions is not without a price. Some may think that this is not a problem, since traditional Buddhist institutions of learning (such as monasteries) provide the appropriate space for aspirant "Buddhist philosophers" in the sense I delineated above; modern universities offer something different, as they should. Yet this is a simplistic understanding of the situation "on the ground." First of all, traditional institutions are in many ways marginalized; most significantly, the (unfounded) prejudice that traditional scholars are incapable of that magic act called "critical thinking" means that traditional institutions are looked upon with suspicion, as bastions of dogma and intellectual stagnation. There is some truth to this criticism: but the dogma that some "traditional" institutions are defending is that of the superiority of Western modes of learning and research, and some may be quite surprised to find out that a number of Buddhist institutes of learning have discarded any trace of traditional learning in favor of Western so-called "critical thinking." This, more often than not, means that—paradoxically—these very Buddhist institutions favor the marginalization of Buddhist thought, which once again becomes purely an object of research, and not a living intellectual resource. It becomes a colony within one's own mind, whose center is (uncritically) set as the Western mode of "critical thinking."

Hence: modern academic institutions affect Buddhist societies, even their most traditional niches. That Buddhist philosophers, institutionally, do not exist, strongly suggests to Buddhist societies that they are irrelevant. And this in turn affects the intellectual continuity and strength of the Buddhist tradition as a whole, not due to some deep and well-thought-out rational decision, but due to a number of historical contingencies that are either left unaddressed or are addressed in a polarizing and politically overcharged manner.

Some may retort that it is up to individuals to determine the center of their intellectual endeavors. In principle, I agree: I do not wish to explain away individual choice and agency. Yet even individuals who may start their academic careers (as students) with the clear purpose of engaging

with Buddhist thought as a living intellectual resource, and as the center of their lifetime project, receive significant institutional pressure to shift priorities and modes of thinking. They are expected to trust modern academic writing over and above the claims of the primary sources, assuming that modern professors are wise, detached, knowledgeable critical thinkers capable of disentangling valuable intuitions from the net of “pre-modern” and/or “native” dogma, superstition, and lack of critical/historical reflection. We should not be insensitive to the fact that it is difficult enough to study Buddhist philosophy at a high level even while receiving support, encouragement, and a clear indication of its value. Studying it while receiving a constant indication of its marginality is, however, disheartening, and may turn even the most enthusiastic young student into a jaded cynic feigning eloquence well beyond one’s understanding of the subject-matter—just so as to “fit in.”

Monasteries and Buddhist universities, even those which allow for and nourish one’s growth as a Buddhist thinker, are not always easily accessible to all; nor are they public spaces broadly recognized as offering legitimate (and relevant) forms of contemporary enquiry. Just as there are public spaces for students to grow as anthropologists or analytical philosophers, there should be similar spaces for someone to grow as a Buddhist thinker—even if the person cannot or does not wish to enter a monastic institution.

Institutional pressure on students is sometimes, perhaps even often, a brutal matter of economics—such as when a student has to apply for a grant and is expected to advertise novelty at all costs. Can one genuinely keep writing such grants for a lifetime without one’s initial sense of direction being lost? The problem is arguably affecting the humanities as a whole, since, it seems to me, the “novelty/discovery/original contribution” model might primarily come from the hard sciences, and affects the existence of the humanities as an independent form of enquiry even capable of reflecting on the value and proper context of science (the paradoxical assumption seems to be that science is the best form of empirical observation and everyone should simply accept this without any reflection).

Even more significantly, the pressure to adopt a “critical” stance is enormous, where “thinking critically” often means criticizing what one has not yet understood or mastered, with the same gusto and sneering of a sadist vivisectioning a frog. Adopting a detached stance, without genuine intellectual conviction, is an obvious rhetorical shortcut to make one’s writing appear “academic,” and sadly my own experience shows that this shortcut is wholeheartedly adopted even by a number of students who will then quickly disown it in private. Is it desirable, though, to encourage this form of intellectual schizophrenia? Not that “critical thinking” is itself undesirable, but there is great difference between understanding what it means (i.e., a type of thoughtful and open-ended reflection) and posing as an academic by gratuitously criticizing the subject matter of one’s research.² I would argue that someone who understands what “critical thinking” means would not work under the false assumption that traditional Buddhist scholars are or have been incapable of thinking critically.

Precisely because I suspect institutional pressure may adversely affect the minds of prospective Buddhist thinkers, I think it is the greatest threat to the study of Buddhist philosophy: it threatens it from the inside in the most vital sense of the term. Perhaps even unwittingly, modern institutions compel researchers to work under the premise that Buddhist philosophy is “wrong”; at least “wrong” enough that all the tools to be employed for its study must be of a Western heritage.³

This paper, however, is primarily meant as constructive criticism. There does exist a long and remarkable tradition of scholarship on Buddhist thought (primarily thanks to philologists), and some contemporary developments could offer valuable opportunities to create more serene spaces and thoughtful engagement with Buddhist philosophy.

INSTITUTIONAL LOCALES FOR BUDDHIST PHILOSOPHY: AN ARGUMENT IN FAVOR OF TEXTUAL STUDIES

A positive development in the study of Buddhist thought is the general move away from what could appear as institutional segregation of sorts, where departments of philosophy and “professional” philosophers would have little to do with anything apart from so-called Western (i.e., European-American) philosophies. No doubt the increased integration of Asian philosophy within the horizon of, say, analytical philosophy, is commendable. More than that—it is perhaps a matter of common sense that there is no possible justification for excluding Asian philosophies from such institutional spaces.

That said, one may also wish to consider what this integration means, and how it happens; and whether something could be improved or profitably expanded upon. One relatively straightforward problem that may arise when integrating Buddhist philosophy in the framework of modern philosophy departments is: how to also ensure that students go through the (to my mind very much necessary) training in the primary source languages? This is, indeed, a purely administrative/institutional problem. Of course, it is possible to acquire some generic notions about Buddhist philosophy even by relying on Buddhist literature translated in English. Yet I do not think that an in-depth study of any area of Buddhist philosophy is (at least at this point in time) possible without becoming proficient in at least one Buddhist source language. The primary reason is rather simple—translations of Buddhist philosophical texts are after all rather scarce (even a foundational text like the *Abhidharmakośabhāṣya* does not exist in a complete translation from the Sanskrit). Furthermore, the level of the available translations varies considerably, as does the terminology employed to translate recurrent technical terms; often, the terminological choices are philosophically very specified and could reasonably become occasions for philosophical and exegetical debates in their own right, and we should, I think, avoid this tendency towards meta-philosophical complications. Understanding key philosophical terms is unquestionably affected by the quality of their translation.

A simple solution would be to have some amount of collaboration with departments that offer training in the primary source languages and in reading Buddhist philosophical texts in those languages. This should be easily achievable. For those wishing to specialize in Buddhist philosophy, I am convinced that language training should be the first and foremost prerequisite.

Setting up a sensible curriculum for prospective students also poses some challenges. Students who wish to focus on Indian Buddhist philosophy should receive sufficient Indological training, i.e., become acquainted with some of the background taken for granted by Buddhist philosophical texts (*vyākaraṇa*, *abhidharma*, etc.). They should also become sufficiently perceptive of the nuances of the language by studying literary genres other than the purely philosophical, such as *sūtras*, poetry, etc.

Programs offering textual studies are crucial for the survival of reliable research in Buddhist thought. The task of creative philosophy may (to an extent) do away with plausible exegesis, dismissing it as being of purely historical/antiquarian value. Yet this should be made clear: if those who do not mind using Buddhist philosophical texts merely as boxes of new exciting philosophical tools are not interested in whether those “tools” were at all intended by the texts they draw them from, they cannot claim to be producing plausible exegesis. The reply that the difference between “creative philosophy” and plausible exegesis is not so clear-cut is fine, but only in the abstract. There are cases where some lines of exegesis can be very safely dismissed as implausible on philological grounds; this view does not entail being fastidiously nitpicky, but simply reading the texts with a modicum of respect and attention (some interpretations can be discarded as little more than syntactical misreadings, which are sadly more frequent than what one may expect).

This brings to mind another consideration: Buddhist philosophical texts studied in a modern (i.e., “Western”) setting tend to become ornaments of an already established project. Rather than being strong voices of their own, having significant continuities with the concerns that still form the backbone of the Buddhist tradition, they are transplanted into an (I would say, inhospitable) atmosphere of no-rebirth, materialist scientism, neuroscience, and other concerns that are as alien to the texts as they are alienating to much of the tradition within which the texts have been transmitted. They are small ornamental plants, quickly forgotten in the background of some enthusiastic discussion about something else entirely.

AN ALTERNATIVE NETWORK OF LIKE-MINDED INSTITUTIONS

Luckily, Asian universities are on the rise; and among them, several universities located in countries wherein Buddhism was or still is a strong cultural and intellectual force. This presents a great opportunity for the Buddhist tradition; not only could it be integrated within indigenous new models somewhat independent of (or at least not subservient to) Western approaches, but even more significantly, it could help rethink and redefine what a university is in the first

place and what university education is meant to provide. Perhaps this could even help save the contemporary university system from the impending self-annihilation it seems headed towards.

Unfortunately, in Asia too institutions tend to disastrously follow an Anglo-American model of administration, evaluation, ranking, and so forth. The disaster is not in the model itself, but in the conformity and subservience; in the assumption that local cultures and traditions are incapable of generating alternative models more suited to local conditions and needs (such as favoring long-term assimilation of the subject matter over quick results in terms of publication output). Changing this situation is possible, by creating a network of collaboration between like-minded institutions, so that Buddhist textual studies may thrive independently of the dubious game of rankings and eloquent yet vacuous publish-or-perish models that so tragically affect higher education through the medium of English. After at least a century of academic study of Buddhism there are thousands of publications with catchy titles, yet very few translations of foundational Buddhist texts; despite the increase of electronic resources and accessibility of texts, achievements like De La Vallée Poussin’s work on the *Abhidharmakośabhāṣya* seem unthinkable for most present-day university employees. Working “horizontally,” though, Buddhist Studies programs may to some extent even escape or bypass the higher administrative levels that form their present-day straitjackets, or at least gain greater intellectual freedom.

OLD HABITS

Long-standing Buddhist approaches to education can and perhaps should be integrated into the didactics of Buddhist philosophy in contemporary universities (why should *everything* come from Europe and from the Christian educational traditions?). For example, chanting Buddhist philosophical texts in Sanskrit,⁴ Pāli, Tibetan, or Classical Chinese. This simple act of repeatedly chanting aloud brings the student of Buddhist philosophy in tune with the purpose of the text, with its intended usage and manner of assimilation, and with its structure and performative elements. Even a few minutes of daily chanting can have a drastic impact on one’s learning experience, and give at least some of the flavor of what these texts are (they are not modern treatises of analytical philosophy, which are not written so as to be suitable for chanting; Buddhist root treatises are, one could say, philosophical “songs,” or at least “chants,” and reading a song in one’s mind without ever hearing the melody is not the same as learning how to sing it).

There is no reason to forget or marginalize Buddhist traditional didactics in favor of Western approaches. Those who agree that Buddhist philosophical texts should be integrated into the curriculum of modern philosophy departments may have no problem in agreeing that some features of non-Western didactics may also be integrated within a university setting. If the content need not be only Western, the same applies to the didactic approaches, and one may decide on a case-by-case basis rather than starting from a priori assumptions of cultural superiority.

AGAINST EARLY INTERDISCIPLINARITY

Interdisciplinarity may be very fruitful when the researcher is familiar with both the disciplines that he/she is employing: in other words—later. It is not reasonable to expect that a student will effectively learn Buddhist philosophy, with its necessary background and source languages, while being forced to absorb gender studies, comparative religions, and whichever fashionable technicality may encourage eloquence and “professionalism” at the expense of substance and understanding. I think it is very important to ensure that programs that offer basic training in language and textual studies still exist and are supported, without bullying them into the endless labyrinth of “Buddhism and . . .” where one ends up studying anything else but Buddhism and yet ends up passing as an expert in the field.

COLLABORATION

Another positive development in the past few decades has been the increased number of collaborations between scholars with significantly different backgrounds. For Jay Garfield, Tibetan traditional scholars have been “colleagues,” not “informants”; this collegiality has resulted in philosophically well-informed vistas of what contemporary Tibetan thinkers may have to say on, for example, key points of Madhyamaka thought. Another type of collaboration is that between those coming primarily from a contemporary philosophical training and philosophically minded philologists: thus we now have an excellent guide to Nāgārjuna’s thought in the form of Siderits and Katsura’s translation and commentary on the *Mūlamadhyamakārikā*.⁵

The increased frequency of these collaborations allows for a level of optimism, and may result in better opportunities for research in Buddhist philosophy that is both well-grounded in the source texts and philosophically engaged, taking traditional exegesis seriously in an atmosphere of non-competitive intellectual dialogue. It is, however, yet early to see the extent to which these positive results may have an impact on institutional thinking and on the organization of modern universities.

BUDDHISM IN THE UNIVERSITY AND OUTSIDE: CREATING A MEANINGFUL INTERFACE

Buddhism is still important, in different ways, to a large section of human society. A sensible way to ensure support and visibility for Buddhist Studies, and for the study of Buddhist philosophy therein, is to create workable interfaces with the society outside of the university. By “workable” I mean that such interfaces should be fruitful for both parties involved.

For example, opening up classes for auditors is a simple way to allow interaction with larger segments of society. Personally, I have found this to be extremely valuable. Auditors participate without the pressure of a clear-cut career program, and occasionally with a clearer and keener interest in the subject matter than what we find in regular students (alas). From this perspective too, their presence helps create a healthier learning environment.

Going back to chanting: chanting sessions can be didactically relevant for both specialized researchers in Buddhist studies, as well as for Buddhist practitioners from outside of the university. This added interest should not be brushed aside: when we compare this with the integration of, say, sanitized/secularized “mindfulness” sessions, one could argue that sessions of chanting achieve almost the same purpose, i.e., integrating a key traditional practice and improving the student’s ability to focus, while being at the same time truer to the tradition and even less religiously charged (for those participants who do not want them to be so, they function as “mere” linguistic training).

At this point, I consider the simple act of including chanting sessions in the didactics of Buddhist philosophy as a symbol of intellectual freedom: “decolonizing the mind” from the clutches of modernist distortions, from cynicism towards the texts, and from the uniformly Western flavor of contemporary university education. Most importantly, it is a symbol, and a performative enactment, of the relationship between memory and understanding, scholastic training and *Buddhist* philosophical reflection.

ACKNOWLEDGMENTS

I thank Giuliano Giustarini, Stephen Jenkins, Aleix Ruiz-Falqués, Rafal Stepień and Pagorn Singsuriya for very useful feedback on earlier drafts of this paper.

BIBLIOGRAPHY

Chakrabarty, Dipesh. *Provincializing Europe. Post-Colonial Thought and Historical Difference – New Edition*. Princeton: Princeton University Press, 2007.

Ngūgĩ wa Thiong’o. *Decolonising the Mind: the Politics of Language in African Literature*. Heinemann Educational, 1986.

Salvini, Mattia. “Chanting the Abhidharma: A Didactic Tool for Modern Universities.” *Proceeding of the International Conference on Tipitaka Studies: Traditional and Contemporary*. Taunggyi: The Organizing Committee of the International Conference on Tipitaka Studies, Shan State Buddhist University (2017): 16–19.

Siderits, Mark, and Shōryū Katsura. *Nāgārjuna’s Middle Way: Mūlamadhyamakārikā*. Boston, MA: Wisdom Publications, 2013.

Stepień, Rafal. “Do Good Philosophers Argue? A Buddhist Approach to Philosophy and Philosophy Prizes.” *APA Newsletter on Asian and Asian-American Philosophers and Philosophies* 18, no. 1 (Fall 2018): 13–15.

Tanizaki, Jun’ichirō. *In Praise of Shadows*. Sedgwick, Maine: Leete’s Island Books, 1977.

NOTES

1. Tanizaki, *In Praise of Shadows*.
2. Or, by arguing for argument’s sake. On this, and on a possible Buddhist stance towards academic argumentation, see Stepień, “Do Good Philosophers Argue? A Buddhist Approach to Philosophy and Philosophy Prizes.”
3. A similar reflection has been applied to the social sciences by Dipesh Chakrabarty:

This engagement with European thought is also called forth by the fact that today the so-called European intellectual tradition is the only one alive in the social science departments of most, if not all, modern universities. I use the word “alive” in a particular sense. It is only within some very particular traditions of thinking that we treat fundamental thinkers who are long dead and gone not only as people belonging to their own times but also as though they were our own contemporaries. In the social sciences, these are invariably thinkers one encounters within the tradition that has come to call itself “European” or “Western.” I am aware that an

entity called “the European intellectual tradition” stretching back to the ancient Greeks is a fabrication of relatively recent European history. Martin Bernal, Samir Amin, and others have justly criticized the claim of European thinkers that such an unbroken tradition ever existed or that it could even properly be called “European.” The point, however, is that, fabrication or not, this is the genealogy of thought in which social scientists find themselves inserted. Faced with the task of analyzing developments or social practices in modern India, few if any Indian social scientists or social scientists of India would argue seriously with, say, the thirteenth-century logician Gangesa or with the grammarian and linguistic philosopher Bartrihari (fifth to sixth centuries), or with the tenth- or eleventh-century aesthetician Abhinavagupta. Sad though it is, one result of European colonial rule in South Asia is that the intellectual traditions once unbroken and alive in Sanskrit or Persian or Arabic are now only matters of historical research for most—perhaps all—modern social scientists in the region. They treat these traditions as truly dead, as history. Although categories that were once subject to detailed theoretical contemplation and inquiry now exist as practical concepts, bereft of any theoretical lineage, embedded in quotidian practices in South Asia, contemporary social scientists of South Asia seldom have the training that would enable them to make these concepts into resources for critical thought for the present. And yet past European thinkers and their categories are never quite dead for us in the same way. South Asian(ist) social scientists would argue passionately with a Marx or a Weber without feeling any need to historicize them or to place them in their European intellectual contexts. Sometimes—though this is rather rare—they would even argue with the ancient or medieval or early-modern predecessors of these European theorists. (Chakrabarty, *Provincializing Europe. Post-Colonial Thought and Historical Difference – New Edition*, 5–6).

The implicit rationale for Chakrabarty’s qualified usage of terms such as “Western” may be similar to my own; whether one is convinced of the historical accuracy of a subdivision of world culture into “Western” and “Eastern,” this dichotomy informs the type of institutional thinking that I analyse in this paper.

4. See also Salvini, “Chanting the Abhidharma: A Didactic Tool for Modern Universities,” for a more elaborate discussion on chanting Sanskrit as a viable didactic tool.
5. Siderits and Katsura, *Nāgārjuna’s Middle Way: Mūlamadhyamakakārikā*.

Reflecting on Buddhist Philosophy with Pierre Hadot

Matthew T. Kapstein

ÉCOLE PRATIQUE DES HAUTES ÉTUDES, PSL RESEARCH UNIVERSITY, PARIS & THE UNIVERSITY OF CHICAGO

In *The Paris Lectures*, phenomenologist Edmund Husserl tells us something startling. He writes, “First, anyone who seriously considers becoming a philosopher must once in his life withdraw into himself and then, from within attempt to destroy and rebuild all previous learning. Philosophy is the supremely personal affair of the one who philosophizes.”¹ In presenting the Cartesian meditation not as a historical anecdote explaining the origin of a particular path of argument, but as a vital practice that we are called upon to repeat within ourselves, Husserl here seems to broach a key concern of French philosopher Pierre Hadot,² to find in the history of philosophical practice, beyond the catalog of doctrines and arguments, a fully lived awakening of reason, or what Hadot, borrowing from the terminology of Loyola, calls “spiritual exercise.”

Hadot’s redeployment of Loyola’s conception within his own field, the study of ancient philosophy in the West, has had wide repercussions in recent years, eliciting new readings of ancients and moderns, European and Asian thinkers.³ The study of Buddhist philosophy has not been immune to Hadot’s broad influence, for, indeed, he appears to suggest a way for us to overcome the incoherence of treating “Buddhist philosophy” as something alien to the Buddha’s teaching, defined by its therapeutic goal of achieving freedom from suffering. As a number of publications have already taken up varied aspects of Hadot’s thought in its possible relations to Buddhism,⁴ my brief remarks here will be limited to a consideration of some of the themes of Hadot’s “Inaugural Discourse” pronounced at the Collège de France in 1983.⁵

In its brief span, the “Inaugural Discourse” surveys many of Hadot’s chief concerns. It is impossible for a specialist in Buddhist philosophy to read it without a strong sense of recognition on almost every page. This is so whether the topic be the allegorical nature of ancient autobiography, or philosophical wisdom as a “stranger” to the world, or, in respect to methodological issues, the challenges posed for the reconstruction of Hellenistic philosophy by the necessity of referring to Latin, as well as Arabic, sources—for Hadot’s remarks about this clearly resonate with difficulties we face in the study of Buddhist philosophy in respect to the relationship between Indian origins and later Chinese and Tibetan works. Above all, Hadot’s substantive reflections on the constitution of philosophical schools, and the figure of the sage, meditation, and the refinement of argument within them, offer distinct parallels to the shape of philosophy in India overall.⁶ With this in mind, it will be useful to consider just one issue in connection with which several of these topics may be engaged, so as to illustrate something of the pertinence of Hadot’s approach for the study of Buddhist philosophy. We may ask, in respect to the traditions that concern us, just what counts as philosophical progress?

The question, of course, may appear paradoxical, for, as the historian of philosophy Martial Gueroult observed, philosophy, insofar as it stakes its claims on the abiding reality of things, involves a radical denial of history. To insist upon a “history of philosophy” becomes a negation of philosophy itself.⁷ Referring to the Hellenistic schools, Hadot expresses a similar puzzle:

In each school, the doctrines and methodological principles were undisputed. To practice philosophy, in that period, was to choose a school, convert to its way of life and accept its doctrines. That is why the fundamental doctrines and rules governing lifestyle for Platonism, Aristotelianism, Stoicism, and Epicureanism did not change throughout antiquity. . . . That is not to say that theoretical reflection and elaboration were absent from philosophical life. However, such activity did not pertain to the doctrines themselves or to their methodological principles, but to the mode of demonstration and doctrinal systematization, and to secondary points of doctrine that were entailed but without unanimity within the school.⁸

That the conditions of development within Buddhist philosophy were in some respects similar may be illustrated with reference to the great figure in the history of Buddhist epistemology and logic, Dharmakīrti (ca. 600 CE).⁹ Often considered a revolutionary thinker, after whom not only Buddhist philosophy, but Indian philosophy more broadly, came to be characterized by a rigorous adherence to and ongoing refinement of the “modes of demonstration” that he set forth, Dharmakīrti was nevertheless a *Buddhist* thinker, whose commitments to his school were fully explicit. This may be seen most clearly in his major work, the *Verse Commentary on Epistemology (Pramāṇavarttika)*, above all, in the chapter on the “Proof of Epistemology” (“Pramāṇasiddhi”), or, more exactly, the proof of the Buddha’s epistemologically authoritative status.

For Dharmakīrti, the sage to whom he adheres, the Buddha, is the veritable embodiment of genuine knowledge, and this may be established by putting his key doctrines to the test. More than this, the Buddha is forever *compassionate* and has not hoarded his liberating knowledge for himself alone, but has freely dispensed it to ameliorate the world. Much of the “Proof of Epistemology” chapter is therefore concerned to argue that the claims made for the Buddha’s compassion and knowledge are indeed warranted, in short that the “Four Noble Truths” that he proclaimed—suffering, its cause, its cessation, and the path whereby this may be achieved—are supported by reason and not by revelation alone.

At the outset, however, some doubts about the nature of the Buddha’s compassion must be settled. For if we take compassion to be limited to the impulse to alleviate the physical pain of others during this lifetime alone, then we might well affirm the Buddha’s compassion without committing ourselves to much of the rest of his teaching, which asserts that suffering in fact belongs to minds and that these minds are processes continuing through innumerable lifetimes, thereby repeatedly suffering through rebirth and redeath. The Buddha’s compassion, it follows, derives its specific character from its embrace of the sufferings of limitless beings throughout those incalculable lives. It is, in fact, an expression of *omnibenevolence*, a quality not of ordinary mortals, but of divinity.¹⁰ (One may bear in mind here Hadot’s remarks on the divinization of the sage in Western antiquity.¹¹)

Given this characterization, the work of reason faces a hurdle, for, if the Buddhist project is to succeed, it must be warranted to hold that there are minds, not only bodies, and that these are continuous over the course of many bodily lives. Dharmakīrti is a philosopher, no mere dogmatist, precisely because he accepts this as a challenge to reason, to which he seeks to respond through reasoned arguments. It would take us beyond the requirements of this short essay to enter into the details of the arguments he presents, not to speak of their assessment. Our sole concern, rather, is to suggest that the pattern observed by Hadot in the case of the Hellenistic schools—that is to say, their adherence to the doctrine of a sage and their progressive cultivation of philosophical advancement within the ambit thereby defined—is precisely what we find here as well. Dharmakīrti emerges as a revolutionary conservative in the history

of philosophy, whose innovations impelled substantive progress in subsequent Indian and Buddhist philosophical practice, while simultaneously upholding the primacy of the Buddha and his cardinal teachings.

If, however, the trajectory of Dharmakīrti’s thinking is in this sense fundamentally conservative, then just what is the interest of his philosophical project to us? Would we not do better, if we wish to orient ourselves to Buddhist thought, to return to the sources of the tradition? Would this not offer a clearer path than the reasoned subtleties of Dharmakīrtian philosophy, where, as has been sometimes suggested, a major motivation is the polemical defense of Buddhism against its detractors?

To respond to such questions, I believe it worthwhile once again to attend to a suggestion from Hadot, in particular, his manner of defining “meditation” in ancient philosophy:

Græco-Roman philosophical meditation is an exercise that is purely rational, imaginative, or intuitive. . . . It is first and foremost the memorization and assimilation of the school’s fundamental doctrines and rules governing lifestyle. Thanks to this exercise, the worldview of one who endeavors to progress spiritually is totally transformed.¹²

Dharmakīrtian spiritual exercise, I would argue, came to play a similar role in later Indian Buddhism, a role that it continued to enjoy in Tibet. By presenting the essentials of the Buddha’s teaching as a rational and coherent whole, it served not only to prepare its adepts to defend their doctrine, but above all to put their own thought-worlds in order accordingly. As such, philosophical activity became the bridge joining doctrinal study to the practical exigencies of the Buddhist path.¹³

We must recall, as well, that Dharmakīrti’s project was just one of several major efforts within Indian Buddhism to elaborate a reasoned account of the Buddha’s thought. An understanding of the historical movement of Buddhist philosophy emerges when we take account of Nāgārjuna, Vasubandhu, Śāntarakṣita, and others not only in respect to their individual contributions, but above all in relation to the dynamic tensions that arose through the effort to reconcile these and others thinkers with one another. Hadot’s remarks on the disaccord, within ancient Stoicism, between Chryssipus and Zeno and Cleanthes point to a similar phenomenon of creative opposition within the schools.¹⁴

This much will suffice, I believe, to suggest that Pierre Hadot’s thinking on Hellenistic philosophy may serve as a stimulus in our investigations of Buddhist philosophy and, indeed, Indian philosophy more broadly. Some points of caution, however, must also be borne in mind. It would be an error to see in the history of Buddhist thought too close an image of its possible Hellenistic parallels, for, although fruitful comparisons have been proposed between specific aspects of Buddhist philosophy and the doctrines and arguments of particular Hellenistic schools, the precise significance of these is often elusive.¹⁵ In the case of Skepticism, above all, striking resemblances between

certain of the arguments we find in Sextus (ca. 160–210 CE) and those of his near-contemporary Nāgārjuna (second century CE) have been sometimes taken to demonstrate the likelihood of transmission between Hellenistic and Indic worlds, which were, after all, in communication with one another during the period concerned.¹⁶ Whether transmission or parallel development is at stake, however, the peculiar character of each school remains irreducible. Whatever the analogies Buddhism may present in relation to Stoicism, Skepticism, or other Hellenistic philosophies, Hadot's emphasis on the integral constitution of the various schools, with their particular sages, lifestyles, and forms of wisdom, must, I think, be kept always in view. To insist that Siddhārtha and Socrates remain distinct figures, however, should not preclude the exploration of possible connections—whether historical or purely conceptual—among the diverse philosophical traditions their teachings spawned.

NOTES

1. Edmund Husserl, *The Paris Lectures*, trans. Peter Koestenbaum (The Hague: Martinus Nijhoff, 1970), 3–4.
2. For Hadot's own appreciation of Husserl's reading of Descartes, in its relation to his thinking on classical philosophy, refer to Pierre Hadot, *La Philosophie comme manière de vivre: Entretiens avec Jeannie Carlier et Arnold I. Davidson* (Paris: Albin Michel, 2001), 124.
3. This is not the place for a detailed bibliography of Hadot-inspired scholarship. For a range of examples, see Stephen Clark, Michael McGhee, and Michael Chase, eds., *Philosophy as a Way of Life, Antients and Moderns: Essays in Honor of Pierre Hadot* (Oxford: Blackwell, 2013).
4. See, in particular, David V. Fiordalis, ed., *Buddhist Spiritual Practices: Thinking with Pierre Hadot on Buddhism, Philosophy, and the Path* (Berkeley: Mangalam Press, 2018), which includes a thorough bibliography.
5. Pierre Hadot, *Leçon inaugurale* (Paris: Collège de France, 1983) and reprinted many times since under various titles. The edition to which I refer here is: Pierre Hadot, *Éloge de la philosophie antique* (Paris: Éditions Allia, 2012). At this writing, I do not have at my disposal the English translation by Arnold Davidson and Paula Wissing, "Forms of Life and Forms of Discourse in Ancient Philosophy," in Pierre Hadot, *Philosophy as a Way of Life* (Oxford: Blackwell, 1995). The passages that I have translated in the present article may therefore differ to some extent from the published English text.
6. Thus, for instance, the prolific Jain philosopher Haribhadrasūri (eighth or ninth century), in his *Compendium of the Six Worldviews (Śaḍdarśanasamuccaya)*, classifies the schools of thought he treats according to their divine promulgators (*deva*) and characteristic doctrines (*tattva*), the latter embracing both topics of argument and themes of meditation.
7. Martial Gueroult, *Histoire de l'Histoire de la Philosophie: I. En Occident, des origines jusqu'à Condillac* (Paris, Aubier, 1984), 13: "On se demande alors comment la philosophie peut être à la fois une vérité, par définition intemporelle, et une suite de doctrines se succédant dans le temps et s'engloutissant dans un passé révolu."
8. Hadot, *Éloge*, 38–39.
9. Scholarship on Dharmakīrti has assumed considerable proportions in recent years. For a survey with copious references, refer to Vincent Eltschinger, "Dharmakīrti," *Revue internationale de philosophie* 3 (2010): 397–440. On the specific subject-matter briefly summarized below, see Eli Franco, *Dharmakīrti on Compassion and Rebirth*, Wiener Studien zur Tibetologie und Buddhismuskunde 38 (Vienna: Arbeitskreis für Tibetische und Buddhistische Studien Universität Wien, 1997).
10. On the Buddha's compassion as a "maximal" property, refer to Paul J. Griffiths, *On Being Buddha: The Classical Doctrine of*

Buddhahood (Albany: State University of New York Press, 1994), ch. 3.

11. Pierre Hadot, "La figure du sage dans l'Antiquité gréco-latine," in Pierre Hadot, *Discours et mode de vie philosophique* (Paris: Les Belles Lettres, 2014), 177–98.
12. Hadot, *Éloge*, 35.
13. For a much fuller discussion of "spiritual exercise" in relation to the Dharmakīrtian tradition, see my "Spiritual Exercise' and Buddhist Epistemologists in India and Tibet," in Steven Emmanuel, ed., *The Blackwell Companion to Buddhist Philosophy* (Oxford: Blackwell, 2013), 270–89.
14. Hadot, *Éloge*, 39.
15. See, for instance, Christopher W. Gowans, *Philosophy of the Buddha* (London/New York: Routledge, 2003), and my "Stoics and Bodhisattvas," in Clark, McGhee and Chase, eds., *Philosophy as a Way of Life*, 99–115.
16. The possibility of connections between the ancient Sceptics and India has been the subject of much speculation. See, for instance, Thomas McEveilly, *The Shape of Ancient Thought: Comparative Studies in Greek and Indian Philosophies* (New York: Allworth Press, 2002); Adrian Kuzminski, *Pyrrhonism: How the Ancient Greeks Reinvented Buddhism* (Lanham MD: Lexington, 2008). A very thorough account of parallels in Buddhist Madhyamaka and classical Skepticism is set forth in Matthew James Neale, "Madhyamaka and Pyrrhonism: doctrinal, linguistic and historical parallels and interactions between Madhyamaka Buddhism & Hellenic Pyrrhonism," Ph.D. Thesis, The University of Oxford, 2014. The publication of this meticulous work has been impeded following the author's tragic, premature passing in 2017.

Some Suggestions for Future Directions of the Study of Buddhist Philosophy

Jan Westerhoff
UNIVERSITY OF OXFORD

There are many ways in which Buddhist philosophy can be studied, and my following remarks should be understood as suggestions for pursuing one direction that seems to me particularly fruitful. I believe there are three areas where further development would be particularly helpful. The first of these is concerned with texts, the second with the relation between texts and the concepts they express, and the final one primarily with concepts.

1. EDITIONS AND TRANSLATIONS

Even though a number of high-quality translations of central Buddhist philosophical texts have come out over the last couple of decades, it is evident that we are still a long way away not just from comprehensive coverage, but even from coverage of much of the basics. To use my own main area of research, Madhyamaka, as an example, none of the canonical Indian commentaries on Nāgārjuna's *Mūlamadhyakārikā* by Buddhapālita, Bhāviveka, and Candrakīrti are currently available in complete English translations. (An English translation of Buddhapālita's commentary by Ian Coghan is forthcoming from Wisdom Publication). Several chapters of Bhāviveka's commentary have been translated, and the whole of Candrakīrti's commentary is available in partial translations in French, German, and English. Matters do not look much better if we look beyond commentaries on the *Mūlamadhyamakārikā*. We do not have complete translations of other key works such as Candrakīrti's commentaries on

Nāgārjuna's *Śūnyatāsaptati* and Āryadeva's *Catuḥśataka*, or of Avalokitavrata's subcommentary on Bhāviveka's commentary on Nāgārjuna.

Such translations are necessary not simply to aid those researchers without or with only limited command of Sanskrit and Tibetan. A high-quality translation is much more than simply a word-for-word rendering of the Sanskrit or Tibetan into English. The translator's task is more complex; amongst many other things, they have to decide between different possible readings of the text, identify quotations, determine when the author and when an opponent is speaking, analyze the dialectical progression of the text, and describe parts of the ancient Indian philosophical or factual background necessary for understanding the argument. As such, any useful translation will already fulfill some functions of a commentary, even though it will not usually provide a full-scale analysis of the text's argument and an assessment of its philosophical plausibility.

Translations thus occupy a middle position between critical editions of texts and high-level philosophical commentaries on them. Without a reliable philological basis, the translator's work is hardly feasible, and a philosophical commentary requires the presence of the basic commentarial functions a translation provides.

It is unclear how we should ever be able to arrive at a reliable understanding even of Indian Madhyamaka (to say nothing about other Madhyamaka traditions based on the Indian one) in the absence of reliable editions and high-quality translations of the majority of key Indian Madhyamaka texts still extant. As such, one requirement for the progress of the study of Madhyamaka in the West (and, given that similar considerations apply to other Buddhist traditions, a requirement for the progress of the study of Buddhist philosophy more generally) is to increase the amount of key texts currently covered by reliable editions and translations.

Given the complexity of the texts involved, it may be that tackling some of these translations as committees (as has been done, for example, in the translation of Tsong kha pa's *Lam rim chen mo*) might be the best way forward. Apart from speeding up the translation process, the result also benefits from the joint expertise of a community of scholars for resolving difficult points.

2. INTEGRATED TEXTUAL AND CONCEPTUAL PRESENTATION

Another set of resources that would greatly facilitate the further study of Buddhist philosophy is an integrated framework for the presentation of texts and concepts. What I mean by this can be most easily described by a simple example. Consider a verse from Nāgārjuna's *Mūlamadhyamakakārikā*, such as 24:18 (*yaḥ prañīyasamutpādaḥ śūnyatām tām pracakṣmahe | sā prajñaptir upādāya pratipat saiva madhyamā* ||, "Dependent origination we declare to be emptiness. It (emptiness) is a dependent concept; just that is the middle path."¹ There is a plethora of material on how to make sense of this verse both in the traditional Buddhist commentarial

literature as well as in the works of Buddhist studies from the last century or so. Reading texts from either of these two groups, the reader will be made familiar with some of the various ways of understanding Nāgārjuna's connecting dependent origination, emptiness, and conceptual dependence. What would be extremely useful, however, is to compile as far as possible a complete collection of the different ways in which this verse has been understood by interpreters ancient and modern, together with a collection of the relevant bibliographic references, and to do this not just for this very famous verse, but for all the 450 verses in the *Mūlamadhyamakakārikā*. This would allow researchers to move from pieces of the text to its conceptual content, together with references to the authorities ancient and modern that underpin the reading of this particular piece of text in this particular way. If this could be carried out for more than one text, we could then envisage going in the opposite direction, from a specific piece of conceptual content to a piece of text that (as interpreted by a specific authority) expresses this content. One such piece of conceptual content expressed by *Mūlamadhyamakakārikā* 24:18, for example, is the identification of dependent origination and emptiness. We could then imagine a kind of conceptual index, one entry of which would be "identity of dependent origination and emptiness," and which would direct the reader to all the passages in the texts covered by the index that stated (or were interpreted by someone to have stated) that dependent origination and emptiness are identical, including *Mūlamadhyamakakārikā* 24:18.

This opens up the possibility of moving back and forth between texts and concepts, beginning with a piece of text, moving to the concepts it discusses, and moving back from there into different texts; or beginning with a specific concept, moving from there to supporting textual sources, and from these to other concepts also supported by them.

The best way of producing and presenting this information would be in electronic form, not just for facilitating access, but also because the conceptual and bibliographical part of the resource would require continuous updating, incorporating new information as more research on specific texts and their interpretations is published.

It should be evident that the production of such a resource faces a number of challenges. The level of expertise required to provide the necessary linkage between texts and concepts is high, and the process itself is relatively time-consuming, especially if a larger group of texts dealing with technically difficult material is considered. At the same time, such work falls outside the familiar forms of academic output such as monographs and articles, making it difficult for younger academics who need to build up their list of publications to participate. Fortunately, there are at least some examples of electronic resources where this kind of challenge has been overcome. One example is the *Stanford Encyclopedia of Philosophy*, which has attracted a number of highly qualified collaborators in part by the fact that authorship of an article on a particular subject-matter now often signals recognition as an authority on the topic concerned.

The general usefulness of such a resource for developing a reliable, synoptic account of what Buddhist philosophers

had to say on specific topics that is not hindered by restriction to a single text or set of texts is evident. More specifically, however, one area where the potential of an integrated presentation of the textual and conceptual dimension of Buddhist philosophical works is immediately obvious is for constructing a conceptual history or *Begriffsgeschichte*² of specific concepts in Buddhist philosophy. It is common knowledge that such central concepts as *svabhāva*, *satyadvaya*, *vijñāna*, or *ākāra* (to give just a few random examples) are used in a variety of ways by different Buddhist philosophical authors. In order to describe how this usage changes throughout the development of the Buddhist philosophical tradition, it is essential to be able to present key passages where the relevant term is used, together with a reliable account of what the term means in this context. If the resource sketched above covered a sufficient number of texts spread throughout the history of Buddhist thought, it would provide the perfect starting place for compiling such a history of technical terms in Buddhist thought.

3. LINKAGE WITH CONTEMPORARY PHILOSOPHICAL DISCUSSION

Once we have a reliable set of editions and translations, and a clear idea how the texts relate to the concepts they express, there is one more piece of the puzzle still missing: we need to show how the ideas and concepts we find in Buddhist texts *matter*. I am aware that not every scholar of Buddhist studies sees this as a necessary requirement. If we regard Buddhist philosophy strictly as a part of the history of ideas, what we should do, it seems, is to describe the thoughts and arguments of the ancient Indian thinkers as clearly and with as much historical contextualization as we can. Whether their premises are plausible or whether their arguments are sound are questions the historian does not need to have an opinion on. While the intellectual merit of this approach is beyond dispute, and the insight it has produced into the historical background of Buddhist thought is considerable, it also strikes me as incomplete. I believe we need to take one further step, and move from Buddhist philosophy as *history of ideas* to Buddhist philosophy as *philosophy*.³

The contrast between these two approaches can be illustrated by considering the example of the history of technology. When historians study some ancient technological device there are two different approaches they could take. One is to describe in as much detail as the historical sources permit the structure of the device, what it could do, and how it might have related to other devices that were extant at the time. The other is to do all this, and then ask the question: Would the device have worked, i.e., was it able to do what ancient sources tell us it could do? If the information on the structure of the device is incomplete, the historian might try to build a working model that restricts itself to the kind of technological resources that might have been available at the time. Or, in addition to this kind of historical reconstruction, one might try to focus on what the device was supposed to be able to do, and construct a version of it that possesses the same functionality, but uses resources that only became available at a later time. The application of the example to the study

of Buddhist philosophy is evident. The first kind of historian corresponds to one who is interested in arguments simply as parts of the history of ideas. The second kind's ultimate concern is the question: Does the argument work? In order to answer this question, he might pursue two approaches. The first, the route of logical reconstruction, tries to make the structure of the argument found in historical texts as clear as possible in order to be able to assess its validity and soundness. The second, which we might call the route of contemporary amplification, explores the possibility that although a given argument philosopher *x* presents for conclusion *y* we find in the texts is either unsuccessful or weak, there may be a different way of arguing for *y*, one that draws on resources that would not have been available to philosopher *x*, but that also do not stand in conflict with other assumptions the philosopher made. These are resources where we can be relatively confident that if philosopher *x* had been presented with them, he would have regarded them as supportive of his own approach.

I believe that the overall aim of the provision of reliable editions and translations of texts, the development of a clear account of how the texts and the concepts they express are linked, and the study of the Buddhist philosophical tradition as history of ideas is the logical reconstruction and contemporary amplification of Buddhist philosophical arguments. It also seems to me that this conception of the Buddhist philosophical enterprise is fairly close to what the commentators of the ancient Indian Buddhist scholastic tradition were aiming at. They certainly tried to give a faithful interpretation of the author's intent, but always within the horizon of connecting what they took to be this intent with the philosophical interests and concerns of their contemporary audience. This is, I believe, the right model for the Buddhist philosopher to follow.

NOTES

1. Siderits and Katsura, *Nāgārjuna's Middle Way. The Mūlamadhyamakārikā*.
2. For example, along the lines of Gerhard Oberhammer's *Terminologie der frühen philosophischen Scholastik in Indien*.
3. One clear example of what I have in mind by the latter is Jay Garfield's 2015 *Engaging Buddhism*.

BIBLIOGRAPHY

- Garfield, Jay. *Engaging Buddhism*. Oxford: Oxford University Press, 2015.
- Oberhammer, Gerhard, Ernst Prets, and Joachim Prandstetter. *Terminologie der frühen philosophischen Scholastik in Indien: Ein Begriffswörterbuch zur altindischen Dialektik, Erkenntnislehre und Methodologie*. Wien: Verlag der Österreichischen Akademie der Wissenschaften, 1991-2006.
- Siderits, Mark; and Shōryū Katsura. *Nāgārjuna's Middle Way. The Mūlamadhyamakārikā*. Boston: Wisdom Publications, 2013.

Practicing Buddhist Philosophy as Philosophy

Pierre-Julien Harter

UNIVERSITY OF NEW MEXICO

Scholarship on Buddhist philosophy¹ has made impressive progress in the past fifty years. Despite the still unsurpassed contributions of such giants as Thomas and Caroline Rhys Davids, Stanislaw Schayer, Louis de la Vallée Poussin, Theodore Stcherbatsky, Eugene Obermiller, and Étienne Lamotte, among many others, what Buddhism has achieved since 1970 will likely be recognized in future historiographical research as a major advance in the scientific task of understanding the historical development and philosophical significance of the Buddhist tradition. Today's Buddhist scholars owe a debt of gratitude to those who have not only made this tradition more accessible by providing editions, translations, and studies, but who have also made possible the very production of scholarship by pressuring institutions to create research and teaching positions that guarantee the long-term presence of Buddhist scholarship within the academy—hence the continually growing interest in a still largely neglected voice of world philosophy.

And yet, this sense of gratitude should not prevent Buddhist philosophers from acknowledging that the field is, in many ways, still in its infancy. This awareness is all the more important if they want to consider the next steps to take on behalf of Buddhist philosophy: to evolve from the stage of the recovery of texts and ideas to the stage of participation in the conversation of world philosophy. While scholars of ancient Greek philosophy can look into critical editions, dozens of translations, and multiple interpretative theories and syntheses for a single author, Buddhist scholars do not even have full translations available of fundamental texts of the tradition: to restrict the focus to Indian Buddhism only, think of Bhāviveka's major works, of all the works of Dharmakīrti (despite his being one of the current foci of Buddhist scholarship), or of later authors (Jñānaśrīmitra, Ratnākaraśānti, etc.). When translations do exist, they are often scattered in bits and pieces in journals or book chapters, or they exist in certain languages and not in others. The lamentation of Edward Conze about the lack of a full translation of Candrakīrti's *Prasannapadā* in one language and one volume ("what a higgledy-piggledy way of dealing with a great classic all this is!"),² though more than fifty years old, still rings true to this day. Even when we are blessed enough to find a translation, it is not necessarily in the language in which the scholar is supposed to teach, thus prohibiting her from using it in class and introducing students to the text. And, of course, one cannot be picky about the quality of the translation, as every reader of Buddhist philosophy (in English at least—and English is probably the best served European language in this regard) is well aware that translations of Buddhist texts are sometimes plagued by so many mistakes and stylistic infelicities (the famous "Buddhist Hybrid English") that they are almost impossible to use. When, for instance, will we finally get an updated, elegant, and full English

translation of such a fundamental (and delightful) text as the *Milindapañha*?

Translations are, no doubt, critical in enabling a disciplinary field to evolve from a space reserved for specialists to a field open to other scholars and available for teaching—and thus to a larger educated audience. Philosophy has thrived through translations—from the *translatio studiorum* that saw philosophical works in Greek translated and preserved in Latin, Arabic, and Persian, to the translations of works across European languages and between European languages and Japanese in modern times; would Wittgenstein ever have revolutionized philosophy if his *Tractatus* had not been translated from German into English almost immediately? If Buddhist philosophy wants to be a part of philosophy, its texts need to be available to a larger circle of readers than those who can access Sanskrit, Pāli, Chinese, Tibetan, Japanese, or Korean texts in the original. And for this to be possible, institutional support is essential, which is why it is crucial for philosophy departments to recognize the demanding and tedious work that philosophical translations represent—work that is often as *philosophical* as new and original works, despite not receiving the same prestige or weight in tenure files or promotions. One translation has often done much more for the field than countless monographs, now long forgotten.

But the translation of texts has a necessary condition that is almost never fulfilled in the field of Buddhist philosophy: the availability of critical editions. Again, a simple comparison with ancient Greek philosophy or the massive critical editorial work that specialists of nineteenth- or twentieth-century philosophers accomplished (whether for Hegel, Nietzsche, or Husserl) can only make a Buddhist scholar envious of the existence of so much reliably edited material. It goes without saying that often Buddhist scholars do not even have the luxury of hoping for critical editions, since so many texts of the tradition have forever disappeared in the original language. But we also know that manuscripts exist, whether in India, Tibet, or China, that are awaiting the administrative or political green light, or the development of intellectual interests and efforts on the part of scholars, to make it to the academic showroom.

These considerations are far from being trivial if we want to think seriously about the academic and intellectual future of Buddhist philosophy. They nevertheless belong more to what Sheldon Pollock calls "philology" (with all the positive attributes that Pollock bestows upon that term) than to philosophy proper. In my opinion, there will be no good Buddhist philosophy without good Buddhist philology, but, surely, a necessary condition is not a sufficient one.

Philosophy—at least when it explicitly claims to be rooted in a tradition of texts, thought, and practice (that is, not only of textual transmission—*āgama*—but also of intellectual and spiritual comprehension—*adhigama*—two aspects claimed by many Buddhist authors, which could arguably also be ascribed to Western philosophy)—is not just constituted by philological research. Otherwise philosophy would run the risk of cultivating a relation with its past on the model of an "antiquarian history," as Nietzsche puts it. It would consist in preserving the texts of the tradition

in their proper contexts, shielding them from the present, as traces of a lost past that is worthy of being preserved in its own right. If we want Buddhist philosophers today to be more than curators, they need to invest in another dimension of philosophy as well, a dimension where the concern with the texts of the tradition does not stifle the critical life of thought.

In a sense, there is an aspect of philosophy that is radically ahistorical: a philosophical thought should be able to engage with texts, ideas, arguments from the past as if they were not from the past, but belonged to a sort of eternal present we can always participate in. This present is the present of arguments and claims to truth, in which we can read Nāgārjuna's verses or Vasubandhu's arguments and wonder whether their arguments hold true or not, despite their being from another time and another place. This presupposition of a philosophically eternal present grounds the very possibility of taking these ideas and arguments seriously enough to wonder whether they are true, convincing, believable, or plainly false. To practice philosophy is precisely to keep alive this capacity of conversing with texts and arguments from other times and other places; it is what turns the *āgama*, the textual basis, into an *adhigama*, an engagement with the ideas of those texts that retrieves their claims to truth and their theoretical and practical potentialities. This is why I would argue that practicing Buddhist philosophy as philosophy, and not simply treating it as a relic of the past, remains true to Buddhist philosophy itself, as well as to philosophy more broadly.

Certainly, there is no shortage of views in (academic) philosophy opposed to such an approach. The historicist trend would tend to reduce all ideas and arguments to context—as if that were even possible—and to suspend judgment about the actual worth or significance of the texts and arguments. On the opposite side, there is a (once ubiquitous, but now less widespread) view that considers anything remotely resembling historical work to be outside the scope of philosophy, and philosophy to begin only when we leave the history of philosophy behind.

It seems to me that the future of Buddhist philosophy hinges on avoiding the historicist Scylla and the anti-historicist Charybdis, and on navigating with the compass of *āgama* and *adhigama*. A Buddhist philosophy that would dispense with the meticulous and tedious confrontation and interpretation of the texts of the tradition (*āgama*) along the philological lines mentioned above would be hard pressed to justify its very name. But scholarship that would restrict itself to archival work would hardly be more than intellectual history, missing proper philosophical comprehension (*adhigama*). Guardians of the *āgama* sometimes turn out to be the censors of *adhigama*.

It is this "philosophically eternal present" that, it seems to me, requires a new specific effort on the part of Buddhist philosophers. For too long Buddhist scholarship has considered Buddhist philosophy as an object of study—though, of course, it was necessary for scholars to explore texts that they did not know, remote in time and cultural circumstances, and subject them to study through rigorous

academic methodology and standards. This effort certainly should never stop. As we know from the example of traditional exegetical traditions in Buddhist countries and from Western scholarship on the history of philosophy, it is an infinite task, always renewable and always promising. But it is time as well for those who consider themselves philosophers to shift their attitude towards Buddhist philosophy from an "objective" approach to a "subjective" one; that is, to occupy Buddhist texts, ideas, and arguments as subjects so as to use them in producing philosophical thought. Such an idea was expressed masterfully by M. P. Rege regarding the situation of Indian philosophy, but it would perfectly fit the picture of Buddhist philosophy as well:

In a way [Indian philosophy] is already in the arena but only as a passive object of examination and assessment. What is necessary is for it to assume the role of a subject actively criticizing and evaluating western philosophical theories from a perspective yielded by its own philosophical standpoints.³

This movement has already begun, and I will not make the mistake of mentioning any names, since mentioning is always a matter of selecting and forgetting. Many talented individuals have already changed the face of Buddhist scholarship through their philosophical writings. Nevertheless, it seems difficult to point at contributions that have changed the general philosophical conversation in any of the major specialized fields in philosophy. The fault certainly falls in part on philosophers who neglect, if not frankly and openly despise, non-Western philosophies. But it seems appropriate to be self-critical as well and recognize that Buddhist philosophers might not yet have taken the steps to really produce new, independent works that would alter general philosophical conversations. Inhabiting Buddhist philosophy as creative subjects means that Buddhist philosophers should engage in such different fields as epistemology, metaphysics, ethics, and (likely less obviously) aesthetics and political philosophy, so as not only to add their contributions to the existing conversation, but also (as great philosophers have always done) to change the conversation and criticize in turn the presuppositions of such fields, even the very constitutions and definitions of those fields.

These two approaches can be understood as two strategies that are available for those who would want to practice Buddhist philosophy as philosophy: one we could call "the appeasement strategy," the other the "confrontation strategy." The "philosophically eternal present" can be more or less extended (as we know since Augustine at least, the present is not a point in time, but has a certain extension), and the whole question is how long we want it to be. One strategy is to make every effort to speak to philosophers trained in and reading exclusively Western philosophy: to speak their language, to use their categories, and to make the same references they do. With this approach, the philosophically eternal present remains restricted mainly to the questions and concepts of Western philosophy, which dictates the extension of that present. This has been the dominant strategy for Buddhist philosophers and has

proved successful to the extent that Buddhist philosophy has evolved into a field that is now recognized in philosophy departments and professional philosophical organizations.

But one can always wonder whether the result has been as successful as it could be. If we acknowledge that Buddhist philosophy remains largely marginalized, there is still probably some room for improvement. Another issue with this approach is that it seems to convey a certain lingering fragrance of colonialism. Western philosophy still remains in command with its questions, its concepts, and its frame(s) of reference. Certainly, there is no discussion without those, and so the point is never to dream of a discussion that could do away with them, as hermeneutics has long taught us, and as the Buddha's *suttas* often exemplify. The question is rather—to follow the trend above—how to extend the philosophical present so as to welcome new questions, concepts, and references. The second strategy puts its effort in this endeavor by empowering Buddhist philosophy as a real subject that is not merely, again, the passive object of investigation of Western scholarship, which would maintain Buddhist philosophy in the role of the respondent or, worse, the defendant. It encourages Buddhist philosophy to take on the role of questioning Western philosophy's concepts and frameworks. It is not enough, for instance, to interrogate Buddhist philosophy about the absence of the concept of free will in its ethical developments. It is also necessary to wonder why this concept appears in its full force in the Latin Middle Ages and acquired such a central role in Western ethics and moral psychology—its presence might be as much of an issue as its absence. Conversely, the detailed development of theories of ethical and moral cultivation in Buddhist philosophy finds very little equivalent in Western philosophy, which is more interested in the analysis of moral principles (i.e., Ethical Theory). Is there here a blind spot in Western philosophy that Buddhist philosophy could help point out and fill in? The concept of the path, so central to so many Buddhist philosophical developments, finds few parallels in Western philosophy: Could Buddhist philosophy perhaps help provide new ways to frame philosophical fields?

Such a strategy will be possible only if we are patient enough to listen to Buddhist texts and hear what they have to say from their own perspectives, as opposed to interrogating them with our own questions. Certainly, this approach has been extensively criticized from the standpoint of certain philosophical, literary, and critical theories in the twentieth century that pushed against the notion of the author's intent. Nonetheless, it has a long, established tradition, both in Western philosophy and in Asian Buddhist traditions, and one can say that the interpretative results it has yielded are enough justifications of its heuristic and pragmatic value. This is where, interestingly, the philosophical conjoins with the philosophical: recovering the voices of Buddhist philosophers, in a way that is not necessarily different from the way scholastic Buddhist traditions have done for centuries, can fuel the practice of Buddhist philosophy as philosophy. To dig out concepts that are not usually included in the philosopher's toolbox—*karma*, *nirvana*, *path*, means of valid cognition—requires an effort to bring out their philosophical dimensions, which in turn allows one to develop further their philosophical potentialities.

Buddhist philosophers can thus draw from Buddhist texts and traditional hermeneutical practices to tune their own voices, at once Buddhist and philosophical, so as to participate fully as subjects in the general philosophical conversation.

Buddhist philosophers who want to extend the philosophical present might find unsuspected allies. I can think of two who are not usually even on the radar of Buddhist philosophers, in contrast to more popular candidates such as neuroscience or analytic epistemology. I believe Buddhist philosophy might develop a promising association with ancient Greek and Latin philosophies as well as medieval Islamic, Jewish, and Latin philosophies that could help move the conversation in directions hitherto unexplored. Medieval philosophies show similar trends in scholastic and exegetical cultures, and in the autonomous developments of logic and epistemology. They are often guided by a certain conception of the *summum bonum*, which is also shared by ancient Greek and Latin philosophies. All these aspects hint at a potential fecund dialogue with Buddhist philosophy. One could imagine the emergence of new fields, *Ancient Philosophies* and *Medieval Philosophies*, which would help to de-regionalize and decenter the usual way to compartmentalize Western and non-Western philosophies—thus changing the conversation, introducing new questions and concepts.

There are arguments to be made—in another paper—about the shared commitments of pre-modern philosophies. But would this direction not preclude the emergence of a Buddhist philosophy as philosophy with its own, proper voice, since it would bring Buddhist philosophy back to a more historical-comparativist approach that tries to find equivalents between philosophical traditions? This is definitely a danger. But it will be avoided if Buddhist philosophers can protect themselves against the historical-comparativist perspective by making sure that they remain conversation partners in whichever conversation they have—instead of being used as tokens for comparison. Ancient philosophy might have the advantage, compared to other conversation partners, of the historical and conceptual distance with our own time, which helps in preserving oneself against contemporary prejudices and starting new subjects of conversation. Ancient philosophy, after all, was a source of inspiration for Alasdair MacIntyre and Martha Nussbaum, two of the most influential philosophers in the twentieth and twenty-first centuries. Buddhist philosophers could certainly be inspired by their examples, drawing from their own *āgamas* to develop an *adhigama* that can speak to and be in conversation with philosophers in general. Only then will Buddhist philosophers actually be part, with their distinct and hopefully diverse voices, of the global philosophical conversation.

NOTES

1. By "Buddhist scholarship" I mean the kind of modern scholarship that has developed since the nineteenth century using scientific methods in the study of Buddhist texts and the Buddhist tradition in general. Buddhist scholarship is not exclusively Western, at least in the geographical sense of the category. Much of Japanese and Chinese scholarship on Buddhist texts produced in the past fifty years is included in the category of Buddhist scholarship, and conversely many Western publications on the Buddhist tradition do not really fall under the category of scholarship.

2. "Recent Progress in Buddhist Studies," in *Thirty Years of Buddhist Studies: Selected Essays* (Oxford: Bruno Cassirer, 1967), 20.
3. Daya Krishna, M. P. Rege, R. C. Dwivedi, and Mukund Lath (eds.), *Samvāda. A Dialogue between Two Philosophical Traditions* (Delhi: Indian Council of Philosophical Research, 1991), xxiv.

Emptiness, Multiverses, and the Conception of a Multi-Entry Philosophy

Gereon Kopf

LUTHER COLLEGE & UNIVERSITY OF ICELAND

As a scholar of Buddhist philosophy, I am frequently asked what the purpose of the study of Buddhist philosophy may be for people who are not practitioners of any of the Buddhist schools. Besides the obvious contribution of thinkers and texts within the Buddhist tradition on topics central to the philosophical project for almost 2,500 years, I believe that one can find within the Buddhist tradition, like in any tradition that furthers intellectual engagement with the world and reflection on the place human beings inhabit in it, quite a few gems that could significantly contribute to pressing issues of our times. One of these gems is the blueprint of what I have come to call a "multi-entry philosophy." In short, a multi-entry philosophy challenges the assumption pervasive in many schools and practices of academic philosophy that to every philosophical question the philosophical inquirer can and will find one correct way of answering, and then one correct answer, while alternative options are deemed untenable or inconsistent. The assumption is that since philosophical inquiry will inevitably lead to the truth, there is only one right way of answering any philosophical but especially any so-called perennial question. I believe that some of the most pervasive conceptions developed in the Mahāyāna Buddhist traditions render a different approach, namely, one that suggests that there are multiple equally persuasive and insightful ways of answering the same question. This is what I call "multi-entry philosophy."

Recent decades have seen a proliferation of conferences, positions, publications, and journals in the fields of, first, comparative and more recently, global philosophy. This expansion of what is included in the discourse and discipline of academic philosophy has been accompanied by discussions about the definition of what philosophy is and what tradition can be recognized as "philosophical" and not merely "intellectual." In general, these discussions have essentialized a modernistic definition of philosophy and traditions as entities with impervious boundaries. These definitions, which have been critiqued by postmodern philosophers as well, are clearly culturally conditioned and ignore other, alternative models such as the understanding expressed by the concept of *sanjiao*, literally "three teachings," which challenges the notion of separate traditions. More interestingly, however, these definitions bring to the fore the fundamental predicament of philosophy as, depending on one's view, either culturally conditioned discourses with universal significance or as universal insights and rules expressed in culturally

conditioned terminology. In his "Zen to sekai" 禪と世界 ("Zen and the World"), UEDA Shizuteru distinguishes "Zen philosophy" from other Zen discourses by emphasizing that the former envisions "one world," what NISHIDA Kitarō 西田幾多郎 (1870–1945) called "global world" (J. *sekaiteki sekai* 世界の世界),¹ in a specific vernacular. Subsequently, located at the intersection of the local and global, the philosopher faces the conundrum of how these two realms can be reconciled. How can we make global claims using a vernacular that is regional and contextualized? In my readings of Buddhist texts, I have found a refreshing solution to this conundrum.

My foray into the Buddhist terminology that renders a multi-entry philosophy begins with Mahāyāna Buddhist conception of "emptiness" (S. *śūnyatā*). This concept finds its first systematic formulation by the philosophers of the Madhyamaka school such as Nāgārjuna (second/third century CE) and Jizang 吉藏 (549–629). In its basic stratification, it possesses an ontological, an epistemological, and a soteriological significance. In a general sense the "theory of emptiness" (S. *śūnyatāvāda*) denies causally independent and permanent substances as well as the possibility of absolute truth and proposes a method of systematic detachment from all desires as well as their objects including the desire to attain detachment itself. What is important here is that while many texts in the Madhyamaka tradition contain *bona fide* ontological, epistemological, as well as soteriological observations, it is almost impossible to clearly separate them into discrete categories or discourses. Since there are no independent permanent entities, it is impossible to make final and definite truth claims. Therefore, we need to let go of many of our long-held beliefs about reality.

But how does this denial of the possibility of arriving at ultimate truth claims explain the construction of knowledge and even the relevance of the doctrine of emptiness itself? To develop a cogent and practicable epistemology, Madhyamaka philosophers introduced the theory of the "two truths" (S. *satya dvaya*, J. *nitai* 二諦), namely, the "conventional truth" (S. *saṃvṛti-satya*, J. *zokutai* 俗諦) and the "ultimate truth" (S. *paramārtha-satya*, J. *shintai* 眞諦). The doctrine of the two truths has been frequently interpreted to restrict linguistic formulae as insufficient and to deem ultimate truth claims impossible. Some representatives of the Chan Buddhist tradition have taken this doctrine to privilege silence over language and to advocate non-linguistic means of communication and verification of a student's insight. However, the doctrine of the two truths does not necessarily imply that linguistic formulations are to be eschewed and to be replaced by the practice of silence as articulated as "non-reliance on words and letters" (C. *buli wenzi* 不立文字), the second of the so-called four principles of Chan Buddhism. For example, WATSUJI Tetsurō (1889–1960) argued in his volume *The History of Buddhist Ethical Thought* that the conventional truth is not to be considered inferior to the ultimate truth since it constitutes the "dharma talk of Buddha."² Both, conventional and ultimate truth, are equally necessary—there is no hierarchy among them: While *paramārtha-satya* facilitates the "destruction of attachments to being," *saṃvṛti-satya* implies the "destruction of the attachment to

emptiness."³ Not only are both truths correlated, they are equally necessary in the pursuit to describe reality. In his *Emptiness and Omnipresence*, Brook Ziporyn paraphrases these two truths appropriately as "local coherence" and "global incoherence":⁴ since all truth claims are contextual, linguistic formulations are not able to articulate universal truths. Conceptions and ideologies illuminate specific contexts but the global perspective is doomed to be incoherent.

One conception that coalesces this understanding of the two truths is the concept of "skillful means" (S. *upayā*). Introduced by the *Lotus Sūtra* (S. *Saddharmapuṇḍarīka sūtra*), this conception and the parables that illustrate it are designed to articulate the belief that while no concept refers to one clearly demarcated signified (F. *signifié*) and no truth claims are universalizable without becoming incoherent, every concept and theory communicates the nature of reality and thus helps propel the listener or reader from a state of "ignorance" (S. *avidyā*) to one of "wisdom" (S. *prajñā*). The theory of "skillful means" thus adds a soteriological dimension to the epistemological rhetoric of the two truths. The central question raised by this terminology—a question the answer to which Buddhist thinkers obviously cannot agree on—is whether complete ignorance and absolute wisdom actually exist or whether they constitute two extreme limit functions that can only be asymptotically approached. The rhetoric of "false views" (S. *mithyā dr̥ṣṭi*) and "correct views" (S. *samyag dr̥ṣṭi*) seems to indicate that ignorance and wisdom constitute real existing states, while the "theory of emptiness," as articulated in the doctrine of the two truths, and the notion of the skillful means seem to indicate, on the contrary, that linguistic formulae are locally coherent and globally incoherent. In his essay "The Standpoint of Zen" (J. *Zen no tachiba* 禪の立場),⁵ NISHITANI Keiji 西谷啓治 (1900–1990) argues that philosophical positions tell us less about the nature of reality and more about the "standpoint" (J. *tachiba* 禪の立場) of the proponents holding them. He does, however, suggest that the standpoint of Zen articulates the theory of emptiness, transcends all conventional hypotheses, and embraces what his student ABE Masao (1915–2006) calls the "positionless position."⁶ While this approach constitutes a common if not mainstream reading of *śūnyatāvāda* among the Zen Buddhist thinkers, it nevertheless falls into the "transcendence trap"⁷ and defies our earlier discussion of the two truths. An alternative interpretation of *śūnyatāvāda* that is held by, among others, Ziporyn implies that our understanding and articulation of reality is necessarily *in the grey* between the extremes of the false and the correct views. As implied by the notion of skillful means and the "ranking of the doctrines" (C. *panjiao* 判教) common in East Asian Mahāyāna Buddhism, philosophical concepts, positions, and systems are indicative of the standpoint of the beholder or adherent. This insight adds the ontological dimension of *śūnyatāvāda*, i.e., the relationship of the local to the global, to our discussion. An elucidation of this relationship illustrates how philosophy as an academic project can be re-envisioned in the light of the "theory of emptiness."

The final piece of the puzzle in our inquiry is the relationship between particular standpoints and the totality of the

cosmos. As is well known, the Chinese Buddhist thinkers of the Tang dynasty (618–907) developed a dynamic since non-essentialist conception of the universe. Thus, for example, Chengguang 澄觀 (738–839) in his *Mysterious Mirror of the Huayan Dharma World* (C. *Huayan fajie xuanjing* 華嚴法界玄鏡) used the common Huayan Buddhist slogan "one-and-yet-many" (C. *yijiduo duojiyi* 一即多多即一) as well as neologisms such as "the unhindered penetration of the universal and particulars" (C. *lishiwuai* 理事無礙) and "the unhindered penetration among particulars" (C. *shishiwuai* 事事無礙) to articulate the philosophical foundation underlying the image of Indra's net.⁸ According to Huayan scriptures, in Indra's net a multifaceted jewel is attached to every intersection in the net. Each jewel reflects simultaneously the net as a whole and every other jewel. Each individual jewel symbolizes one awareness event, and the net the totality of the cosmos. Not only does each awareness event reflect all other awareness events, it does so in a unique way. Therefore, there exist infinite manifestations of what we call cosmos. Ziporyn calls this conception of the cosmos not a "universe" but a "multiverse."⁹ This multiverse is constantly changing and, at any given time, is embodied in multiple manifestations. The Japanese Zen master Dōgen 道元禪師 (1200–1253) refers to these manifestations as "expressions" (J. *dōtoku* 道得). Like the Huayan thinkers before him, Dōgen was convinced that particular awareness events express all other awareness events as well as their totality. In his fascicle "Expression" (J. *Dōtoku* 道得), however, he highlights another important dimension of this relationship when he coins the term "expression-and-non-expression" (J. *dōtoku fudōtoku* 道得不得)¹⁰: Expressions, be they discursive or in non-linguistic modalities, comprise full but incomplete embodiments of the buddha-dharma, what we would call the "truth." They constitute unique yet incomplete views on the world. They disclose as much about their own standpoint, which is defined *vis-à-vis* other standpoints, as they reveal about the nature of reality. Consequently, an integration of as many standpoints as possible increases our understanding of world and our place within it.

Thus understood, Dōgen's notion of expression provides the blueprint for a multi-entry philosophy. This particular model of doing philosophy is inspired by the trope of Indra's net developed in the writings of Huayan Buddhism and Dōgen's conception of "expression" but also by the rhizomatic model of Gilles Deleuze (1925–1995) and Félix Guattari (1930–1992)¹¹ as well as Francois Lyotard's (1924–1998) "small narratives" (F. *petit récit*).¹² It is based on Dōgen's insight that "when one aspect is brought to the fore, another aspect is obscured."¹³ A multi-entry approach to philosophy realizes that even the best philosophical systems are only locally coherent insofar as they highlight some aspects while ignoring others. Finally, every philosophical system expresses one standpoint and is constructed *vis-à-vis* other standpoints: e.g., atheism is constructed *vis-à-vis* theism, materialism *vis-à-vis* idealism, and rationalism *vis-à-vis* empiricism. To be fruitful, a philosophical inquiry will invite a multiplicity of approaches, listen to their language, understand their standpoints, and discern their contribution to the specific inquiry as well as to the philosophical project in general. Of course, the format of this essay does not allow for a careful stratification

of such a multi-entry philosophy. I hope, however, that the current sketch of such a philosophy illustrates the kind of systematic thinking Mahāyāna Buddhist texts that propagate the notion of emptiness, the rhetoric of skillful means, the vision of a multiverse, and our understanding of human activity as expression engender. It is a method that embraces epistemic humility and emphasizes listening as a path to self-awareness and to a deeper understanding of the world.

NOTES

1. Abbreviations are as follows: C for Chinese, F for French, J for Japanese, S for Sanskrit.
2. Tetsurō Watsuji, *Watsuji tetsurō zenshū 和辻哲郎全集 [The Complete Works of Watsuji Tetsurō]*, 21 volumes. (Tokyo: Iwanami Shoten, 1961–1963) [hereafter abbreviated as “WTZ”], 19: 297.
3. WTZ 19: 297.
4. Brook A. Ziporyn, *Emptiness and Omnipresence: An Essential Introduction to Tiantai Buddhism* (Bloomington: Indiana University Press, 2016), 151.
5. Keiji Nishitani, *Nishitani keiji chosakushū 西谷啓治著作集 [Collected Works of Keiji Nishitani]*, 26 volumes. (Tokyo: Sōbunsha, 1986–1995), 11: 3–157.
6. Masao Abe, *Zen and the Modern World: A Third Sequel to Zen and Western Thought*, ed. Steven Heine (Honolulu: University of Hawai’i Press, 2003), 10.
7. James Whitehill, “Buddhist Ethics in Western Context: The ‘Virtues’ Approach,” *Journal of Buddhist Ethics* 1 (1994): 6. This essay was republished as James Whitehill, “Buddhism and the Virtues,” in *Contemporary Buddhist Ethics*, ed. Damien Keown, 17–36 (Richmond, Surrey: Curzon Press, 2000), 21. See also Jin Y. Park, “Wisdom, Compassion, and Zen Social Ethics: The Case of Chinul, Sōngch’ōl, and Minjung Buddhism in Korea,” *Journal of Buddhist Ethics* 10 (2006): 10.
8. *Taishō taizōkyō 大正藏經*, ed. Junjirō Takakusu and Kaigyoku Watanabe (Tokyo: Taishō Shinshū Daizōkyō Kankōkai, 1961), 1883.45.672-683.
9. Brook A. Ziporyn, *Emptiness and Omnipresence: An Essential Introduction to Tiantai Buddhism* (Bloomington: Indiana University Press, 2016), 143.
10. *Dōgen zenji zenshū 道元禪師全集 [Complete Works of Zen Master Dōgen]*, ed. Dōshū Ōkubo 城大久保道舟, 2 volumes (Tokyo: Chikuma Shobō, 1969–1970) [hereafter abbreviated as “DZZ”], 1: 305.
11. Gilles Deleuze and Félix Guattari, *A Thousand Plateaus: Capitalism and Schizophrenia*, transl. Brian Massumi (New York: Continuum, 2003), 3–25.
12. Francois Lyotard, *The Postmodern Condition: A Report on Knowledge*, trans. Geoffrey Bennington and Brian Massumi (Minneapolis: University of Minneapolis Press, 1984), xxiv.
13. DZZ 1: 7.

become designated as “contemplative neurosciences,” a field evolving in mutual interaction with the development of mindfulness therapies that are now being applied in a growing number and variety of social contexts. Historically and traditionally, the Buddhist practices targeted in neuroscientific research were embedded in soteriological, ethical, and philosophical frameworks which provided them with their meaning and function, and they continue to be so embedded among traditional Buddhist communities across Asia. Critical observers of mindfulness studies and practices have variously problematized the disassociation of meditation from these frameworks, and in doing so concentrated on ethical and religious contexts.¹ In this essay I want to broaden the discussion by offering some critical reflections on the place of Buddhist philosophy in contemplative neurosciences.

To begin with, it is not entirely unreasonable to expect that philosophical ideas developed in Buddhist traditions would be relevant to the empirical study of contemplative practices that can claim to be grounded in the same traditions. An argument for this could be developed along the following lines. Generally speaking, in Buddhism the philosophical analysis of consciousness and the elaboration of meditative states form part of the same overarching soteriological discourse aimed at determining the causes for human bondage, at making a case that liberation is possible, and at devising methods for attaining it. The relationship between philosophy and meditation is complex, as can be expected from an internally heterogeneous religio-philosophical tradition with a history of over 2,500 years that geographically extends over nearly all of today’s Asia and has in the course of its history been shaped by vastly divergent societies and cultures. But the following appears to apply across many historical, societal, cultural, and doctrinal divides. Buddhist traditions develop fundamental philosophical principles such as the idea that there is no enduring self-as-substance that serves as a core of human individuality and subjectivity (selflessness), the idea that consciousness is ultimately non-dual and devoid of subject and object (a notion frequently taken to entail a denial of the external world), or the idea that all phenomena, material as well as mental, lack inherent existence that would allow them to exist independently of causes and conditions (emptiness). Not all traditions assent to all of these principles, and among those who assent to the same principles there is, again, variation in their interpretation. More important for the present context, however, is the dual function that these principles fulfill: on the one hand they serve as presuppositions in general theories of the mind that also extend to meditative states. On the other hand such philosophical principles are themselves the focus and content of certain forms of meditation. In other words, there is a type of meditative practice that consists in, or includes, the realization of such principles as selflessness, non-duality, or emptiness, which are understood as representing the true nature of reality. This realization is thought to be indispensable for attaining liberation. Following a helpful taxonomy of Buddhist meditation devised by Florin Deleanu, we can describe this kind of meditation as reality-directed and reflection-centered.² The category of “meditation” in Buddhism is by no means exhausted by this type of practice; others—

Buddhist Philosophy and the Neuroscientific Study of Meditation: Critical Reflections

Birgit Kellner

AUSTRIAN ACADEMY OF SCIENCES

Since the early 2000s, the scientific study of Buddhist (or Buddhism-inspired) meditation has expanded and intensified considerably. The study of the effects of contemplative practices on the brain has more recently

again following Deleau's taxonomy—focus on eliminating disturbing emotions, on the suppression or elimination of consciousness, on the cultivation of moral qualities such as loving-kindness, or on controlling or manipulating the esoteric physiology of the body. Most of these practices were historically developed in ancient and medieval India, and a number of them are shared with non-Buddhist religious-philosophical movements.

Although this background would seem to justify some consideration of Buddhist philosophy in the empirical study of Buddhist or, to put it more cautiously, Buddhism-derived contemplative practices, this has scarcely been the case. Indeed, this seems to be one of the issues that this line of research is currently grappling with. But why have philosophical ideas played such a marginal role in scientific studies of meditation? The answer to this question can be sought in the nature of the cross-cultural and cross-disciplinary engagement that initiated, catalyzed, and shaped the development of contemplative neurosciences—a process on which I will take the liberty of commenting in broad brushstrokes for the sake of argument, while being aware that some of my extrapolations will not do justice to the variety of standpoints taken by the various actors involved in the process at different stages.

Instrumental in this process were the activities of the Mind & Life Institute, founded in 1990 as a result of conversations of the fourteenth Dalai Lama Tenzin Gyatso with scientists (in which the neuroscientist Francisco Varela was crucially involved).³ Since 1998 the institute has worked to “create a collaborative research program for investigating the mind with scientists, Buddhist contemplatives, and contemplative scholars,”⁴ resulting, among others, in neuroscientific studies with long-term meditators conducted in Paris and in Wisconsin–Madison. The principle guiding the institute's activities is that a purely scientific investigation of the nature of reality is incomplete, and must be supplemented by contemplative practices (including, but not limited to, Buddhism-derived practices), which are regarded as playing an equal role as instruments of investigation.⁵ Buddhism is in this context seen as a resource for contemplative practices with the potential of contributing to overall human well-being. The view that both partners, science and Buddhism, each have their own unique investigative methods that complement each other and should work together can be seen as a step to move the scientific study of meditation forward, away from a mere objectification of Buddhism to a more dialogical framework.⁶ At the same time, the primacy of science in this dialogue is reflected in a more or less explicit agreement to sidestep issues where Buddhist doctrine might be in conflict with modern scientific worldviews—an agreement which also permits the Buddhist participants in these dialogues to compartmentalize science and its scope of authority while protecting their own interests.⁷

Obviously, this engagement reflects a complicated process of negotiation between actors with different and divergent perspectives on meditation and its study, in pursuit of a common goal. These complexities notwithstanding, the framework as such arguably has the tendency of producing what one might call, for lack of a better term, an “empiricist

filter” which regulates what is and what is not fed into these dialogues from the rich repertoire of Buddhist thought. This filter appears to leave little room, if any, for philosophical ideas that Buddhist thinkers developed as an analysis of reality—unless these ideas can be claimed as having been the product of meditation as a method of investigation. An example may serve to illustrate how this filter operates in practice. In one conversation between the Dalai Lama and scientists, scientists enquired about the intriguing notion of a “storehouse consciousness” (Sanskrit *ālayavijñāna*), a subliminal layer of consciousness that the Buddhist Yogācāra school postulates to run alongside manifest forms of consciousness such as sense-perception. In response, the Dalai Lama criticized and dismissed the storehouse consciousness on the ground that the Yogācāras were compelled to formulate it “because of their rational presuppositions, rather than through empirical investigation or realization.”⁸ Many scholars studying the intellectual history of Buddhism would find themselves in general agreement with this assessment, if it is taken as a claim about how the notion of the storehouse consciousness was historically introduced. One plausible hypothesis for its genesis—in texts datable to the early fourth century—is that it was introduced as an explanatory principle to account for the continuity of consciousness across states of deep meditative absorption where all ordinary mental functions have ceased—and for how it is possible that after such an interruption ordinary mental functions can reoccur.⁹ However, the storehouse consciousness also fulfills other important functions in later discussions of consciousness, after the principle was initially introduced. To occlude Buddhist philosophical ideas from dialogues with science because they were not empirically discovered is therefore problematic because it reduces the relevance of such ideas to the mode of their initial discovery, disregarding how they might nonetheless become relevant in different theorizations at a later stage. On a more fundamental level, claims as to which principles were or were not discovered through meditative practice remain inherently problematic, as evinced by the fact that scholars of Buddhism continue to debate the issue controversially.¹⁰

The even more basic question is whether meditation can claim the status of an empirical investigative method in the first place. At the very least, this conception contradicts some of the historically significant ways in which meditative practice was conceptualized in traditional Buddhist contexts, especially in relationship to philosophical analysis conducted with the help of logical reasoning. According to an influential Indian Buddhist approach that also impacted upon Tibetan traditions—some of whose exponents are mentioned further below—meditation is precisely *not* conceived as an independent investigative method leading to new discoveries about the mind. Rather, this approach is premised on the idea that meditative practice—here referred to with the Sanskrit term *bhāvanā*—serves to turn principles like emptiness that are first proven through scriptural study and rational inquiry into objects of experience. Following Tom Tillemans, we can extrapolate the principle underlying this approach as a “continuity thesis” claiming that meditative understanding is continuous with and dependent upon preceding philosophical analysis; it has no independent epistemic value of its own.¹¹ On this

understanding, meditation internalizes and inculcates principles determined by reasoning and is thought to thereby fundamentally transform consciousness so that it operates in accordance with those principles—principles which are taken to reflect the true nature of reality. While meditation affects the quality of knowledge (it produces non-conceptual awareness of something known before only in propositional terms), it does not generate knowledge of new objects or contents. The eighth century Indian Buddhist scholar-philosopher Kamalaśīla (c. 740–795 CE) famously advocated this position in polemical exchanges with the Chinese Chan master Moheyan in imperial Tibet, and the matter is traditionally taken to have been settled in a public debate presided over by the emperor himself at Samyé monastery. Moheyan’s position, more difficult to reconstruct from a patchy textual record, was likely premised on an “independence thesis” (again, Tillemans’s term) that denies soteriological value to philosophical analysis and conceptual understanding. On Moheyan’s view meditation has a unique role in disclosing the true and primordially awakened nature of the mind.

Without going deeper into the long and complicated further history of these approaches in Tibetan Buddhism—in the course of which models were developed that do not view them as diametrically opposed—let us note that a continuity thesis of the kind advanced by Kamalaśīla would challenge the attribution of an independent investigative role to meditation. Meditative cultivation is here not admitted to be a reliable source for knowledge in its own right. Following ideas prominently advanced by the earlier Buddhist epistemologist Dharmakīrti (mid-sixth to mid-seventh centuries), Kamalaśīla would as a matter of fact insist that meditative cultivation on its own is not reliable; it is no warrant for the reality of its objects. Dharmakīrti and his many followers in India and Tibet consider the basic mechanism of meditative practice, that of transforming abstract, conceptual knowledge into direct experience, to be effective regardless of whether meditation is applied to something real or fictitious. Thus, for example, if one contemplates an object of desire long enough, it will eventually appear vividly as if in front of one’s eyes. Precisely for this reason meditation must be preceded by a process of reasoning that proves its content to represent the nature of reality. The direct experience which meditation produces is thought to be necessary because of its presumed special power to effect a lasting transformation of consciousness; it is not, as recently claimed,¹² thought to corroborate or confirm the validity of any preceding reasoning. The process of reasoning is not regarded as incomplete in terms of the evidence that it produces. Rather, since it is limited to producing a merely conceptual certainty, it is simply not capable of transforming consciousness in the way that this is necessary from a soteriological perspective, and which meditation eminently accomplishes.

The particular engagement of Buddhism with science that fed into the formation of contemplative neurosciences thus conceives of meditation in a way at odds with its function in at least some influential traditional contexts. While this itself does not discredit a conception of meditation that takes it to perform an investigative role, one has to be clear about the extent to which this marks a departure from

traditional conceptions. Moreover, once we look at how Buddhist scholar-philosophers like Kamalaśīla conceived of meditation (in the sense of *bhāvanā*), we may also formulate the question differently and ask precisely how it is that meditation is understood to play an “investigative” role. At the current stage of research, and quite apart from the fact that at least one influential tradition of Buddhist thought on the matter explicitly argues *against* meditation possessing an investigative function, the nature and extent of the investigative function meditation *could* play remain rather unspecified in the realm of contemporary contemplative neurosciences.

One could question the relevance of my reflections inspired by Kamalaśīla by pointing out that the form of meditation he discusses—in Deleau’s terms: reality-directed and reflection-centered meditation—is just one of many, and that others are far more pertinent to the dialogue between Buddhism and science. Emotion-regulating practices, for instance, might be considered far more relevant to a dialogue conducted in pursuit of human well-being. This, however, leads to yet another problem: Deleau’s taxonomy notwithstanding, philosophical notions such as emptiness are also relevant to practices that in themselves are not directly directed at an understanding of reality. In the 1980s, a Tibetan meditator called Lobsang Tenzin participated in studies on the biomedical effects of meditation conducted by Herbert Benson, a pioneer of biofeedback research.¹³ Tenzin was asked to participate in such studies on account of his expertise in the Tibetan *tummo* practice by which a meditator becomes capable to regulate their body temperature; in Deleau’s taxonomy, *tummo* would belong to the class of practices serving to control or manipulate the esoteric physiology of the body. Tenzin’s report offers insights into a variety of problems involved in experimental studies from the perspective of participants—problems which also raise serious ethical questions.¹⁴ Among others, he expresses fears that his practice might be adversely affected by the scientists since the drawing of blood by doctors intervened with the subtle physiology of channels and winds that form the basis of the practice. The environmental conditions of the exams he underwent weakened his practice, he writes, but the scientists were very pleased.

Particularly relevant for our context is that Tenzin astutely draws attention to epistemic limitations of the scientific apparatus. The scientists, he writes, could see certain things with their instruments, but there were also things that they could not see: the non-dual insight into “great bliss” and emptiness, the root of all gnosis. But as he was able to confirm from his own experience, without this subjectively experienced wisdom, the control over wind and channels that *tummo* practice required would not be possible at all. To reformulate the problem that Tenzin points out in other words, there are certain factors which render a given practice effective according to those who are engaged in it, and precisely these factors are not being taken into account when scientists study these practices. These factors notably include elements of Buddhist philosophical theories such as non-dual insight and emptiness, even in a practice that in Deleau’s terminology would not in itself be reality-directed.

It is now widely recognized in scientific studies of meditation that the subjective reports by meditators must be included in the research process. Ambitious proposals on how to combine first-person phenomenological reports with third-person scientific methods have been put forward.¹⁵ But at more or less the same time, scientific research paradigms seem to have started to exercise methodological constraints that push in the exact opposite direction: one where empirical studies become even more divorced from the universe of meaning, including philosophical meaning, that renders Buddhist meditation effective according to those who practice it. A study conducting EEG measurements on eight experienced meditators practicing unconditional loving-kindness and objectless compassion illustrates this dynamic well. In this practice, benevolent attitudes are radiated outward without being confined to any one person or group. Studies indicated regular abnormalities in high-frequency gamma waves.¹⁶ The practice appeared capable of synchronizing neurons in great numbers and to change neuronal activity in a sustainable fashion. This suggests long-term meditation practice can effect lasting changes in the brain, but according to the scientists who conducted the study one cannot exclude that these EEG measures derive from preexisting peculiarities of individual participants. Such peculiarities might have motivated the participants to take up serious practice in the first place; they need not be the results of long-term practice itself. Meditation practice, moreover, is over the long term a highly individual affair, as meditators may after all choose, upon the advice of their teachers, to combine different practice styles.

One conclusion that may be drawn from this is that if more precise causal relationships were to be established between meditation practice and measurable brain activation patterns, the training process would have to become part of the experiment so as to become subject to methodological control. The “attentional blink” study subsequently conducted in Wisconsin–Madison illustrates this approach.¹⁷ It combined a group of subjects trained in *vipassanā* for three months as part of the study, and a control group that received only a one-hour crash-course. When these two groups, as well as totally untrained subjects, were asked to identify two visual objects shown to them in a 500 milliseconds interval, their responses differed significantly. Only the retreat participants were able to consistently identify the second object since the practice had trained them to suspend cognitive and emotional reactions that untrained and lesser-trained persons exhibit immediately after registering a visual object.

My concern here is not with the merits of this experiment or the reliability of its results, but with the sociocultural consequences entailed by a larger-scale adoption of its underlying approach: if only a strictly controlled environment in which scientists cooperate with therapists and *vipassanā*-teachers were admitted to provide results with scientific authority, the scientific study of meditation would end up producing the very context that it is to study. A self-reproducing environment would emerge that takes the disassociation of meditative practice from its soteriological, ethical, and philosophical frameworks to an entirely new level.

Scientists and scholars of Buddhism who collaborate in the neuroscientific study of meditation occasionally express the hope that future studies should more strongly engage with questions of context,¹⁸ which may signal that the production of closed and controlled training environments is not viewed as a direction the field should take as a whole. What those who conduct such studies have in mind when they speak of context are primarily ethical principles, traditional background, culture as well as environment, including such factors as altruistic motivations and teacher-student relationships. At this stage, it seems to be recognized as a challenge how neuroscientific research might be able to incorporate such aspects to explore their function in modulating the production of particular mental states, but attempts to address this challenge appear to be still in their initial stages. The authors of one research paper consider it to be extremely difficult to take subjective reports by participants into account because it is hard to distinguish clearly defined descriptions of meditative states from a first-person perspective from other statements that at first sight look like descriptions, but might in the end better be treated as certain “cultural and religious exigencies that are not strictly rooted in scientifically tractable observations.”¹⁹ This distinction seems opaque and might even be viewed as yet another expression of the “empiricist filter.” If a meditator were to report on their experiences by saying “and then I entered a state of clear light,” perhaps the scientists might consider this close enough to a “descriptive” statement because ways could be found to translate such statements into phenomenological descriptions, or into the language of empirical psychology. But one can imagine that the claim “and then I entered non-dual awareness of emptiness and experienced the Great Bliss” would pose greater challenges and end up being dismissed as not “properly” descriptive.

If Buddhist philosophical frameworks are admitted as shaping the kinds of experiences contemplative practitioners from such traditions make, then a research process conducted in dialogue with such contemplatives will have to somehow attend to these frameworks. The research contexts discussed in this essay do not yet seem to have reached a stage where this particular challenge can be met. Whatever direction they may take in the future, it stands to reason that the more room that is given to the methodological constraints of natural sciences—and the investigative primacy of science—the more difficult it will become to incorporate contextual factors that are outside the sphere of control of scientists. One conclusion to draw from this is that the research framework for the empirical study of contemplative practices might have to be reconsidered, or that new research frameworks might have to be devised in order to facilitate a more open-ended dialogue that does full justice to the richness and variety of Buddhist philosophical discourse.

NOTES

1. Harrington and Dunne (“When Mindfulness Is Therapy”) review ethical critiques and place them in a historical context; Lopez (*Buddhism & Science: A Guide for the Perplexed*, chapter 5) gives a general and critical account of how the mindfulness movement and neuroscientific research disassociated meditation from religious frameworks.
2. Deleanu, “Agnostic Meditations on Buddhist Meditation.”

3. A brief history of the Mind & Life Institute is given as an Appendix in Wallace *Buddhism & Science, Breaking New Ground*, 417–21.
 4. Thompson, *Waking, Dreaming, Being. Self and Consciousness in Neuroscience, Meditation, and Philosophy*, Kindle version, position 1795.
 5. See, among others, the mission statement of the institute at <https://www.mindandlife.org/mission/>, last accessed January 4, 2019.
 6. Cabezón, "Buddhism and Science: On the Nature of the Dialogue," 36–39, for a critical discussion of "objectification" as a dominant mode of interaction between Buddhism and science.
 7. Dunne, "What Is Inner Science?," 323.
 8. The Dalai Lama and Varela, *Sleeping, Dreaming, and Dying: An Exploration of Consciousness with the Dalai Lama*, 88.
 9. Schmithausen, *Ālayavijñāna: On the Origin and the Early Development of a Central Concept of Yogācāra Philosophy*.
 10. The view that central philosophical doctrines of Buddhism derive from spiritual practice was advanced in Schmithausen ("On the Problem of the Relation of Spiritual Practice and Philosophical Theory in Buddhism"), subjected to critique in Franco ("Meditation and Metaphysics: On their Mutual Relationship in South Asian Buddhism"), to which Schmithausen responded in Schmithausen (*The Genesis of Yogācāra-Vijñānavāda: Responses and Reflections*).
 11. Tillemans, "Yogic Perception, Meditation, and Enlightenment: The Epistemological Issues in a Key Debate." Tillemans's discussion is confined to the eighth-century thinker Kamalāśīla whom I discuss further below, but the "continuity thesis" as he outlines it is also applicable to a great number of other Indian Buddhist scholar-philosophers.
 12. Adam, "Philosophy, Meditation, and Experience in the Great Debate at Bsam yas," 364.
 13. Tenzin, "Biography of a Contemporary Yogi," is a first-personal account of his experiences; for background discussion see Harrington, *The Cure Within: A History of Mind-Body Medicine*, epub version location 475, 5–481, 7.
 14. Cabezón, "Buddhism and Science: On the Nature of the Dialogue," 37f.
 15. Lutz and Thompson, "Neurophenomenology: Integrating Subjective Experience and Brain Dynamics in the Neuroscience of Consciousness."
 16. Lutz et al., "Long-Term Meditators Self-Induce High-Amplitude Gamma Synchrony During Mental Practice."
 17. Slagter et al., "Mental Training Affects Distribution of Limited Brain Resources."
 18. See Lutz et al., "Attention Regulation and Monitoring in Meditation," 168; Lutz et al., "Investigating the Phenomenological Matrix of Mindfulness-Related Practices from a Neurocognitive Perspective."
 19. Lutz, Dunne, and Davidson, "Meditation and the Neuroscience of Consciousness: An Introduction," 544.
- Deleanu, Florin. "Agnostic Meditations on Buddhist Meditation." *Zygon* 45 (2010): 605–26.
- Dunne, John D. "What Is Inner Science?" In *In Vimalakīrti's House: A Festschrift in Honor of Robert A. F. Thurman on the Occasion of his 70th Birthday*, edited by Christian Wedemeyer, 319–44. New York: AIBS/ Columbia University Press, 2014.
- Franco, Eli. "Meditation and Metaphysics: On their Mutual Relationship in South Asian Buddhism." In *Yogic Perception, Meditation and Altered States of Consciousness*, edited by Eli Franco, in collaboration with Dagmar Eigner, 1–51. Vienna: Verlag der Österreichischen Akademie der Wissenschaften, 2009.
- Harrington, Anne. *The Cure Within: A History of Mind-Body Medicine*. Cambridge: Harvard University Press, 2008.
- Harrington, Anne, and John D. Dunne. "When Mindfulness Is Therapy." *American Psychologist* 70, no. 7 (2015): 621–31.
- Lopez, Donald S. *Buddhism & Science: A Guide for the Perplexed*. Chicago and London: University of Chicago Press, 2008.
- Lutz, Antoine, John D. Dunne, and Richard J. Davidson. "Meditation and the Neuroscience of Consciousness: An Introduction." In *Cambridge Handbook of Consciousness*, edited by P. D. Zelazo, M. Moscovitch, and E. Thompson, 499–551. Cambridge: Cambridge University Press, 2007.
- Lutz, Antoine, Lawrence L. Greischar, Nancy B. Rawlings, Matthieu Ricard, and Richard J. Davidson. "Long-Term Meditators Self-Induce High-Amplitude Gamma Synchrony During Mental Practice." *Proceedings of the National Academy of Sciences of the United States of America* 101, no. 46 (2004): 16369–73.
- Lutz, Antoine, A. P. Jha, J. D. Dunne, and C. D. Saron. "Investigating the Phenomenological Matrix of Mindfulness-Related Practices from a Neurocognitive Perspective." *American Psychologist* 70, no. 7 (2015): 632–58.
- Lutz, Antoine, Heleen Slagter, Richard J. Davidson, and John Dunne. "Attention Regulation and Monitoring in Meditation." *Trends in Cognitive Science* 12 (2008): 163–69.
- Lutz, Antoine, and Evan Thompson. "Neurophenomenology: Integrating Subjective Experience and Brain Dynamics in the Neuroscience of Consciousness." *Journal of Consciousness Studies* 10 (2003): 31–52.
- Schmithausen, Lambert. "On the Problem of the Relation of Spiritual Practice and Philosophical Theory in Buddhism." In *German Scholars on India. Contributions to Indian Studies, Vol. II*, 235–50. Bombay: Nachiketa Publications Ltd., 1976.
- Schmithausen, Lambert. *Ālayavijñāna: On the Origin and the Early Development of a Central Concept of Yogācāra Philosophy. Parts I & II*. Tokyo: The International Institute for Buddhist Studies, 1987.
- Schmithausen, Lambert. *The Genesis of Yogācāra-Vijñānavāda: Responses and Reflections*. Tokyo: International Institute for Buddhist Studies of the International College for Postgraduate Buddhist Studies, 2014.
- Slagter, Heleen A., Antoine Lutz, Lawrence L. Greischar, Andrew D. Francis, Sander Nieuwenhuis, James M. Davis, and Richard J. Davidson. "Mental Training Affects Distribution of Limited Brain Resources." *PLoS Biology* 5 (2007): e138.
- Tenzin, Lobsang. "Biography of a Contemporary Yogi." *Cho yang (chos dbyangs)* 3 (1990): 102–11.
- Thompson, Evan. *Waking, Dreaming, Being. Self and Consciousness in Neuroscience, Meditation, and Philosophy*. New York: Columbia University Press, 2015.
- Tillemans, Tom J. F. "Yogic Perception, Meditation, and Enlightenment: The Epistemological Issues in a Key Debate." In *A Companion to Buddhist Philosophy*, edited by Steven M. Emmanuel, 290–306. Chichester: Wiley-Blackwell, 2013.
- Wallace, B. Alan (ed.). *Buddhism & Science. Breaking New Ground*. New York: Columbia University Press, 2003.

BIBLIOGRAPHY

Adam, Martin. "Philosophy, Meditation, and Experience in the Great Debate at Bsam yas." *Journal of the International Association for Buddhist Studies* 39 (2016): 351–74.

Cabezón, José Ignacio. "Buddhism and Science: On the Nature of the Dialogue." In *Buddhism & Science*, edited by B. Alan Wallace, 35–68. New York: Columbia University Press, 2003.

The Dalai Lama, His Holiness; and Francisco Varela. *Sleeping, Dreaming, and Dying: An Exploration of Consciousness with the Dalai Lama*. Somerville: Wisdom Publications, 2004.



INTRODUCTION

#MeToo and Philosophy

Lauren Freeman
UNIVERSITY OF LOUISVILLE

[Content warning for discussion of sexual violence.]

Continuing an important legacy of Black women's advocacy for sexual violence survivors, in 2006 Tarana Burke, a Black feminist, social activist, and community organizer, founded the Me Too movement. Her founding of this movement was a culmination of sorts, after having worked with Black female survivors of sexual violence for over a decade. Beginning in the late 1990s, Burke was working in Selma, Alabama, with youth, many of whom had been victims of sexual violence. Burke herself had also been a victim of repeated sexual assault, beginning from the age of six. One day, Burke was talking to Heaven, a thirteen-year-old girl who she could tell wanted to confide in her about abuse she'd experienced. Knowing where the conversation was going, and that Burke herself was not in a place where she could properly respond to someone else's trauma, she cut the conversation short, before it went there. Later, she regretted that decision immensely. "I cut her off," Burke recounted. "She was so hurt and devastated. That moment taught me the power of empathy. All she needed was to hear the words, 'Me, too'."¹

Burke knew that these girls needed help. She also knew that, at the time, she was not equipped to offer the kind of help they needed, so she sought advice from and possible collaboration with a local rape-crisis center. This proved to be a dead end, as she was told that the center would only work with girls who had gotten a referral from the local police department after they'd filed an official report. Burke knew how sexual assault worked and how police are often the last people survivors go to, let alone trust, after having experienced sexual assault. So, in 2003, she started Just Be, Inc., a non-profit organization that aims to help women heal from sexual assault. In 2006, she named the movement Me Too. The focus of this movement was to respond to the experiences of abuse suffered by Black and brown girls and women who are and remain disproportionately vulnerable, specifically, by connecting survivors of sexual assault to the resources they need in order to heal.

More than a decade later, in the wake of the growing accusations of sexual harassment and assault by film mogul Harvey Weinstein, the term "me too" resurfaced in a different

way and to a different end. At around noon on October 15, 2017, Actress Alyssa Milano encouraged spreading the hashtag #MeToo on social media, in order to draw attention to the prevalence and pervasiveness of sexual harassment and assault by showing how many people have personally experienced such events. By six o'clock p.m. on the same day, the hashtag had been shared over 200,000 times; by noon the next day, over 500,000 times. On Facebook, the hashtag was used by over 4.7 million people in 12 million posts in the first 24 hours. Facebook has estimated that 45 percent of US users had at least one friend who had posted the statement "me too."² As of October 2018, #MeToo has been tweeted more than 19 million times.³

As Tarana Burke has recently said, "What #MeToo allowed people to do was create community with these shared experiences. You have a built-in group of people who automatically gets you, who automatically believes you, who automatically wants to hear you. That's the wildfire of it."⁴ As Miranda Pilipchuk discusses in her contribution to this issue, "This emphasis on marginalized survivors is particularly important given that many of the white-led anti-violence movements that have gained national uptake have largely failed to address the concerns of marginalized survivors." For Burke, the aim of the movement is to provide resources to people to help them heal and to advocate for changes to laws and policies. Specifically, she's emphasized goals like processing all untested rape kits, re-thinking local school policies, improving the ways in which teachers are vetted, and updating sexual harassment policies.⁵ For Burke, the aim of the movement ought to be focused on survivors and not on perpetrators. For Milano, a priority of #MeToo is to challenge the laws surrounding sexual harassment and assault, for example, implementing legislation that makes it difficult for publicly traded companies to hide cover-up payments from their stockholders and to make it illegal for employers to require new employees to sign non-disclosure agreements as a condition of employment.⁶

#MeToo continues to loom large in the national and international consciousness as increasingly more men are accused and charged of sexual harassment and sexual assault. And yet the number of cases that go unreported and the number of women who remain silent is even larger, pointing to the systemic problems of injustice for victims of abuse, assault, and harassment and the systematic failures of our institutions to bring about justice. All of these problems are complicated by the class, race, nationality, immigration status, sexuality, gender identity, and disability of victims.

#MeToo is a crucial form of resistance and the #MeToo movement, in all of its complexity, is ripe for philosophical engagement and analysis. In October 2018, in collaboration with The CUNY Graduate Center Advanced Research Collaborative, the Center for the Humanities, the Philosophy Program, and Hunter College, Linda Martín Alcoff and Charles Mills organized and hosted an interdisciplinary conference on “#MeToo and Epistemic Injustice.” Aside from that, there has been no sustained philosophical treatment of the #MeToo movement. The present issue of the *Newsletter on Feminism and Philosophy* aims both to fill this lacuna and also to start more of a profession-wide discussion about what philosophy and philosophers have to contribute to the movement. Because this is the first sustained treatment of #MeToo by philosophers, it does not claim to be comprehensive by any means. Yet, the papers that you will find here begin some important conversations.

Miranda Pilipchuk’s article, “**Good Survivor, Bad Survivor: #MeToo and the Moralization of Survivorship**,” examines how #MeToo’s reliance on disclosure has resulted in moral dilemmas for survivors and, more generally, in a moralization of survivorship itself. She begins by situating Burke’s original version of Me Too within the historical legacy of Black women’s advocacy for sexual violence survivors and then highlights the central features that differentiate Burke’s original version of Me Too from the viral social media version of the movement. Pilipchuk then goes on to discuss the nature of disclosure within the context of #MeToo and argues that the impetus to disclose has created a moral hierarchy of survivors, where survivors who publicly disclose receive greater moral standing than survivors who remain silent. In response to this problem, she recommends that #MeToo realign itself with the priorities of Burke’s original Me Too movement, namely, to focus on survivors (rather than perpetrators) and to provide resources for survivor healing—especially survivors from marginalized communities. Ultimately, her claim is that in order to be truly inclusive of all survivors, #MeToo must expand beyond mere disclosure.

Sarah Clark Miller’s paper grapples with the question of whether to disclose at all and the kinds of moral conundrums that disclosing poses to survivors. She opens “**Beyond Silence, Towards Refusal: The Epistemic Possibilities of #MeToo**” by admitting that for decades, she was silent about her own sexual assault. However, after watching Dr. Christine Blasey Ford’s harrowing testimony before Congress on September 27, 2018, something in her changed, something she’s come to understand as “a moment of tremendous epistemic refusal galvanized by the similar gestures of refusal [she] saw all around [her].” That moment, within the larger context of #MeToo, she began to see the complicated and myriad ways in which victims are pressured to remain silent and, as a result, how our confidence in our knowledge about the violence we endured is undermined. As Miller goes on to show, such mechanisms of silencing and undermining are insidious, “accomplished through widespread practices of credibility erosion as well as ostracization and shaming.” Miller uses her reflection on her own decision to come forward as a victim of sexual assault as a point of departure for considering the nature of such

an epistemic transformation, and, more broadly, as a lens to better understand some of the epistemic dimensions of what she saw happening around her. Of import, Miller coins, unpacks, and develops the term “epistemic refusal,” which refers to the strategy of breaking one’s silence, in one’s own way, on one’s own terms, thereby refusing the dominant epistemic structures that have kept us tightly in check.

Miller’s paper shows how the mass informal disclosure of survivor status of #MeToo that took place in conjunction with other hashtags like #WhyIDidntReport and #BelieveHer has created space for epistemic, ethical, and political community between survivors of sexual violence, specifically by denying the hegemonic epistemic discourses of contemporary rape culture. A key insight of Miller’s paper is her discussion of the ways that mass informal disclosure of sexual violation can “shift the focus from the credibility of the survivor to the wrongful actions of the perpetrator—moving feelings of shame and responsibility away from victim and back onto perpetrator.” The implications of this point have great moral and political promise as we consider how to move forward in building communities of solidarity and epistemic subcultures for and between survivors in the #MeToo era and beyond. This idea might also help us to reconcile some of the goals of #MeToo with the original guiding initiatives of the Me Too movement.

Though it seems as though #MeToo has launched a paradigmatic shift in the way that we collectively understand sexual assault—just look at the number of times the hashtag has been shared on social media—in “**The Speech Acts of #MeToo**,” Cassie Herbert asks whether the movement has really been all that successful in bringing about justice. She notes that as we look around us, there does not seem to be a corresponding increase in holding perpetrators accountable for the sexual violations they’ve committed. Insofar as that’s the case, she writes, it seems as though #MeToo has not been the great success that one might initially assume. Yet Herbert contends that this is the wrong way of measuring the success of the #MeToo movement for at least two reasons: First, by using such a metric, its emphasis is misplaced by focusing on perpetrators rather than on the support given to survivors (recall, Burke’s concerns about the way that sexual assault is understood and discussed); second, it misunderstands the precise kinds of actions that comprise the movement. According to Herbert, speech act theory can help to provide a better sense of what #MeToo was and is, and, subsequently, it can help us better evaluate the success of the movement. In order to do this, Herbert explores two ways of understanding the social media posts that comprised #MeToo: namely, as accusations (speech acts that call for holding the perpetrator to account) and as reports (truth-claims about the world rooted in the speaker’s first-personal experience). She goes on to argue that #MeToo ought to be understood as composed primarily of reports. This move, she contends, better enables us to evaluate the successes of the movement; moreover, it also allows us to recognize how the movement signifies an important shift in the dominant frameworks by which survivors’ speech is understood.

Varying successes of #MeToo notwithstanding, in her paper “#MeToo?,” Lori Watson points out, importantly, that “[i]ntersectional inequalities, including those on the basis of race, class, sexuality, disability status, as well as gender expression and identity mean that the power of #MeToo is unequally distributed.” This means that credibility attached to the testimonies of victims of sexual assault is also not equally assigned and “[t]he more socially subordinated, the more inequality on the basis of group membership, the less credibility or support is extended.” Consequently, she contends that “[t]he potential power of #MeToo depends on equality within it; hierarchies of standing that underwrite social inequality such that only some women finally get their moment of justice, whether legal or social, must be dismantled for the reality of the #MeToo moment to unfold.” Thus, we must articulate and underscore the experiences of the most socially marginalized across all intersecting inequalities in order for #MeToo to bring about the kind of all-encompassing social change that many call for.

Watson paves one of the many roads necessary in order to do this. Specifically, she notes that for the most part, #MeToo has focused on heterosexual encounters of sexual harassment and violence. Same-sex sexual harassment and violence has played far less a role in the movement. And even when same-sex violations have been revealed (most famously, Terry Crews and James Van Der Beek), they have primarily been by men calling out abuse they’ve experienced by other men. Same-sex sexual harassment and violence between women has not been a part of the narrative of #MeToo. Watson’s contribution to this issue begins to fill this lacuna by shedding light on same-sex sexual harassment between women. Specifically, she explores the gendered dynamics that shape and define such harassment and, in so doing, reveals the layered complexities of power dynamics at work. The broader aim of her paper is to begin to help make the #MeToo movement more inclusive, particularly for those who fall outside of the heterosexual female mold to which it has mostly catered up until now. As she indicates, clearly, there is still much more work to be done.

In her essay, “#MeToo vs. *Mea Culpa*: On the Risks of Public Apologies,” Alice MacLachlan analyzes how we should understand and what we should do with all of the apologies—and, in particular, the plethora of non-apologies and quasi-apologies—we’ve seen in the wake of the accusations and allegations of assaults and violence of #MeToo. MacLachlan notes that “[b]ad apologies can be frustrating, infuriating, and profoundly painful—in some ways, they are more hurtful than outright denials of wrongdoing.” In response, MacLachlan asks, what if they’d been better, namely, “what if the #MeToo movement had produced, along with an upsurge of truth-telling and solidarity among survivors and allies, an equivalent rise in genuinely sincere statements of responsibility and remorse by perpetrators and collaborators?” Her answer to this question is not what we might intuitively expect. Instead of going on to argue that our situation would have been much improved had we received the kind of ideal apologies many might have hoped for and wanted, in a more nuanced manner, MacLachlan suggests that this

would not have been a straightforwardly happy ending. Her view is that “the risks of public apologies are not limited to the faults and flaws of bad apologies; in many ways, good ones are more insidious.” Consequently, her essay focuses not on how to assess and compare individual apologies; nor on the necessary and sufficient conditions for the ideal apology, or even a good apology in the age of #MeToo. Rather, MacLachlan identifies three risks to both good and bad public apologies and shows how these risks increase when it comes to the gender politics of #MeToo. Ultimately, she argues that the aims of #MeToo may be in tension with and may even be undermined by the practice of public *mea culpae*.

In the final paper in this volume, “**Women, Work, and Power: Envisaging the Radical Potential of #MeToo**,” Robin Zheng considers #MeToo as a form of political struggle aimed at social change. To begin, she situates the movement between two poles of structured organizing and mass protest: the Women’s March of 2017 and the recent wave of teacher’s strikes that we’ve seen across the United States in the last few years. This comparison is important as it casts into relief the need for movements to translate *symbolic* solidary power, which shifts discourses and norms, into exercises of power that can alter *structural* conditions in a more fundamental way. She goes on to argue that while #MeToo has been highly successful in disrupting sexist mores and patriarchal norms on the cultural front, there is still a good deal of work to be done structurally. That is, we must match this success with a larger commitment to transforming the fundamental material conditions that enable men’s dominance over women. If we are going to eliminate sexual harassment, Zheng argues—particularly of non-elite women—then there are many related issues that must also be simultaneously addressed, since sexual harassment never happens in a bubble and is a part of the larger issues of systemic and systematic discrimination against women. Specifically, Zheng urges the inclusion in our discussion of sex-based discrimination the related issues of underlying job insecurity, poor working conditions, and economic vulnerability that threatens almost all workers. In short, her claim is that #MeToo must go radical.

The narrative essay in this issue, “**Field Notes on Conference Climate: A Decade with the Philosophy of Science Association Women’s Caucus**,” by Julia Bursten, chronicles many of the changes that occurred at the conference between 2008 and 2018. Bursten herself served as the early-career co-chair of the PSA Women’s Caucus from 2014 to 2018. In her narrative essay, she reflects on their aim “to make a more humane conference environment for everyone involved, and in so doing, to make it easier for women to exist in the conference space—and to exist as philosophers, rather than as *women* philosophers.” This essay will be helpful for all conference and symposium organizers to read. Of particular note is the initiative to assemble and disseminate in all conference registration packages a flyer on bystander intervention that reviews basic information on what constitutes sexual harassment and discriminatory behavior and offers a number of quick, in-the-moment strategies to reroute an instance of discrimination and support the targeted person. Also of note is the dependent care offered by the conference,

details of which can be found here: <https://psa2018.philsci.org/en/83-information-for-attendees/108-dependent-care>. Though it seems like there is always a taller mountain on the horizon to climb when climate issues in philosophy are on the table, Bursten is optimistic that we are on the right track.

Book reviews in this issue include Caleb Ward's discussion of Hilke Charlotte Hänel's *What Is Rape? Social Theory and Conceptual Analysis*; Mari Mikkola's discussion of Andrew Altman and Lori Watson's *Debating Pornography*; and Catherine Clune-Taylor's discussion of Shelley L. Tremain's *Foucault and Feminist Philosophy of Disability*.

In closing, I'm so grateful to all of the wonderful contributors to this volume for giving the newsletter the opportunity to profile their superb work, as well as to Ann Cahill, Rebecca Fraser, Rebecca Harrison, Grayson Hunt, Debra Jackson, and Heather Stewart for their helpful comments on papers and for their service to the profession. I look forward to the discussions that result from this issue. I hope that you find the contributions to be as illuminating as I do.

As a final note, I'd like to welcome Kate Norlock to her new post as chair of the APA Committee on the Status of Women. Kate is taking over from Charlotte Witt, who has served in this position for the past three years. On behalf of the entire committee, we offer a heartfelt thanks to Charlotte for her hard work and service, and a warm welcome to Kate!

NOTES

1. E. Wellington, "Tarana Burke: Me Too Movement Can't End with a Hashtag."
2. CBS News, "More Than 12M 'Me Too' Facebook Posts, Comments, Reactions in 24 Hours."
3. M. Anderson and S. Toor, "How Social Media Users Have Discussed Sexual Harassment Since #MeToo Went Viral."
4. U. Dabiero, "Tarana Burke, Creator of #MeToo, Says the Movement 'Has Lost It's Way' for This Reason."
5. C. Snyder and L. Lopez, "Tarana Burke On Why She Created the #MeToo Movement—And Where It's Headed."
6. For a far more detailed discussion of the key differences between the two movements, see section I of Pilipchuk's paper in this issue.

REFERENCES

- Anderson, Monica, and Skye Toor. "How Social Media Users Have Discussed Sexual Harassment Since #MeToo Went Viral." *Pew Research Center*. 2018. Accessed July 2019. <https://www.pewresearch.org/fact-tank/2018/10/11/how-social-media-users-have-discussed-sexual-harassment-since-metoo-went-viral/>.
- CBS News. "More Than 12M 'Me Too' Facebook Posts, Comments, Reactions in 24 Hours." *CBS News*. October 17, 2017. Accessed June 2019. <https://www.cbsnews.com/news/metoo-more-than-12-million-facebook-posts-comments-reactions-24-hours/>.
- Dabiero, Una. "Tarana Burke, Creator of #MeToo, Says the Movement 'Has Lost It's Way' for This Reason." *FairyGodBoss*. 2018. Accessed July 2019. <https://fairygodboss.com/articles/tarana-burke-creator-of-metoo-says-the-movement-has-lost-its-way-for-this-reason>.
- Snyder, Chris, and Linette Lopez. "Tarana Burke On Why She Created the #MeToo Movement—And Where It's Headed." *Business Insider*. December 13, 2017.
- Wellington, Elizabeth. "Tarana Burke: Me Too Movement Can't End with a Hashtag." *The Philadelphia Inquirer*. October 23, 2017. Accessed July 2019. https://www.inquirer.com/philly/columnists/elizabeth_wellington/philly-me-too-movement-founder-tarana-burke-20171023.html.

ABOUT THE NEWSLETTER ON FEMINISM AND PHILOSOPHY

The *Newsletter on Feminism and Philosophy* is sponsored by the APA Committee on the Status of Women (CSW). The newsletter is designed to provide an introduction to recent philosophical work that addresses issues of gender. None of the varied philosophical views presented by authors of newsletter articles necessarily reflect the views of any or all of the members of the Committee on the Status of Women, including the editor(s) of the newsletter, nor does the committee advocate any particular type of feminist philosophy. We advocate only that serious philosophical attention be given to issues of gender and that claims of gender bias in philosophy receive full and fair consideration.

SUBMISSION GUIDELINES AND INFORMATION

1. Purpose: The purpose of the newsletter is to publish information about the status of women in philosophy and to make the resources of feminist philosophy more widely available. The newsletter contains discussions of recent developments in feminist philosophy and related work in other disciplines, literature overviews and book reviews, suggestions for eliminating gender bias in the traditional philosophy curriculum, and reflections on feminist pedagogy. It also informs the profession about the work of the APA Committee on the Status of Women. Articles submitted to the newsletter should be around ten double-spaced pages and must follow the APA guidelines for gender-neutral language. Please submit essays electronically to the editor or send four copies of essays via regular mail. All manuscripts should be prepared for anonymous review. References should follow *The Chicago Manual of Style*.

2. Book Reviews and Reviewers: If you have published a book that is appropriate for review in the newsletter, please have your publisher send us a copy of your book. We are always seeking new book reviewers. To volunteer to review books (or some particular book), please send the editor, Lauren Freeman (lauren.freeman@louisville.edu), a CV and letter of interest, including mention of your areas of research and teaching.

3. Where to Send Things: Please send all articles, comments, suggestions, books, and other communications to the editor: Dr. Lauren Freeman, University of Louisville, lauren.freeman@louisville.edu.

4. Submission Deadlines: Submissions for spring issues are due by the preceding November 1; submissions for fall issues are due by the preceding February 1.

ARTICLES

Good Survivor, Bad Survivor: #MeToo and the Moralization of Survivorship

Miranda Pilipchuk
VILLANOVA UNIVERSITY

This article analyzes the issue at the heart of the viral social media “Me Too” campaign (#MeToo) that began in the fall of 2017: disclosure. Survivor disclosure is a key defining feature of the campaign that distinguishes it from other anti-violence campaigns—including Tarana Burke’s original version of Me Too.¹ Unlike previous anti-violence campaigns, #MeToo relies exclusively on survivor disclosure to educate the general public about issues of sexual violence. This article examines how #MeToo’s reliance on disclosure has led to moral dilemmas for survivors, and a moralization of survivorship itself. Section I outlines the central features of both versions of the Me Too movement—Burke’s original version and the viral #MeToo version—paying careful attention to how #MeToo has departed from Burke’s original movement. Section II focuses specifically on the issue of disclosure in #MeToo. I argue that the impetus to disclose has created a moral hierarchy of survivors, with survivors who publicly disclose being given greater moral standing than survivors who remain silent. Finally, in section III I recommend #MeToo realign itself with the priorities of Burke’s original Me Too movement, which is to provide resources for survivor healing—especially survivors from marginalized communities. Ultimately, I argue that in order to be truly inclusive of all survivors, #MeToo must expand beyond mere disclosure.

I. TWO ME TOOS

Burke’s original version of Me Too and #MeToo share two prominent similarities: a name and a theme. Both movements are titled “Me Too,” and both movements explicitly address issues of sexual violence. Beyond these similarities, the two movements differ substantially in their history, their target audience, and their desired outcomes. This section traces the genesis of each movement and explicates the key features that distinguish the movements. I argue that although the movements share a name, they are in many ways engaged in different projects.

Burke founded the original version of Me Too after a personal encounter with a young survivor of sexual violence named Heaven, which took place more than a decade before #MeToo went viral. When Heaven disclosed her experience with sexual violence to Burke, Burke “listened until I literally could not take it anymore . . . cut her off, and immediately directed her to another female counselor who could ‘help her better.’”² Burke expresses regret about the exchange, recounting that “I couldn’t even bring myself to whisper . . . me too.”³ Burke founded Just Be, Inc., a nonprofit organization dedicated to providing resources to sexual violence survivors, in 2003. In 2007, she named the movement Me Too.⁴

Burke’s original version of Me Too continues an important legacy of black women’s advocacy for sexual violence survivors.⁵ Black women were among the first to raise sustained, comprehensive critiques about sexual violence in the United States. In 1861, Harriet Jacobs published one of the first accounts of sexual violence written from a survivor’s perspective. Jacobs discloses her abuse by Dr. James Norcom, documenting that “He threatened me with death, and worse than death, should I make any complaint.”⁶ Though Jacobs’s work contains a brilliant critique of the relationship between power, oppression, and sexual violence that foreshadows analyses from contemporary scholarship, she is not widely credited as a sexual violence theorist.

In 1892, Ida B. Wells published the first “political understanding of sexual assault.”⁷ Wells is perhaps best known for her analyses of how rape was used to justify the lynching of black men, but an important part of her anti-lynching work involved calling attention to how frequently black women and girls were raped by white men, and how white America refused to acknowledge or address this problem. Wells’s work emphasizes the double standard in white America’s concern with rape, highlighting how race led to extreme discrepancies in both legal and extra-legal responses to rape. For example, Wells reports that “A white man in Guthrie, Oklahoma Territory, two months ago inflicted such injuries upon another Afro-American child that she died. He was not punished, but an attempt was made in the same town in the month of June to lynch an Afro-American who visited a white woman.”⁸

Black women were also the first to stage organized protests against sexual violence. Angela Davis reports that “Before the end of the nineteenth century pioneering Black clubwomen conducted one of the very first organized public protests against sexual abuse.”⁹ Darlene Clark Hine argues that from the start, the National Association of Colored Women’s Clubs was thoroughly invested in addressing sexual violence against black women. Mary Church Terrell, the founding president of the Clubs, launched a campaign designed to undermine the negative tropes that have traditionally been used to justify sexual violence against black women.¹⁰ And Frances Ellen Watkins Harper, a leading member of the Women’s Christian Temperance Union, started a campaign to urge white women to speak out about sexual violence against black domestic workers.¹¹

Burke’s original version of Me Too continues this legacy of black women’s advocacy by explicitly calling attention to the disproportionate impact of sexual violence on marginalized communities. One of the central purposes of Burke’s original version of Me Too is to provide resources for survivors who might be marginalized from mainstream sources of assistance.¹² The official Me Too website states that “The ‘me too’ movement was founded in 2006 to help survivors of sexual violence, particularly Black women and girls, and other young women of color from low wealth communities, find pathways to healing.”¹³ This emphasis on marginalized survivors is particularly important given that many of the white-led anti-violence movements that have gained national uptake have largely failed to address the concerns of marginalized survivors. Despite the fact

that black women were the earliest, strongest advocates against sexual violence, popular movements led by white activists have historically neglected issues of racism as well as heterosexism, transphobia, and ableism.¹⁴ In the mid-late nineteenth century, white suffragettes who advocated on behalf of sexual violence survivors focused exclusively on white women survivors.¹⁵ In the 1970s, some prominent white anti-violence activists explicitly endorsed racialized myths surrounding sexual violence—including the myth that black men were more likely to be rapists.¹⁶ And many current anti-violence campaigns have called for solutions to sexual violence that could have a disproportionately negative impact on marginalized communities.¹⁷ By being attentive to survivors from marginalized communities, Burke has explicitly addressed the shortcomings of other anti-violence movements and reoriented activism against sexual violence to include survivors that have been alienated by other movements.

#MeToo has a somewhat different origin story. In October 2017, Alyssa Milano tweeted, “If all the women who have been sexually harassed or assaulted wrote ‘Me too’ as a status, we might give people a sense of the magnitude of the problem.”¹⁸ Milano later reported that she was inspired to send out the tweet by Rose McGowan, who had been struggling with issues of sexual violence.¹⁹ Both Milano and McGowan are famous white women. In less than a day, the hashtag went viral. Facebook reports that 4.7 million people engaged with “me too” in the first 24 hours.²⁰ The Me Too hashtag was tweeted 1.7 million times in 85 countries in the first week alone. Within 6 months it had been used over 12 million times globally.²¹

Although Milano was originally credited with starting Me Too, she has since publicly acknowledged Burke as the Me Too founder. This acknowledgment is important. Given that white-led anti-violence movements have tended to overlook or obscure the importance of black women’s activism, it is crucial that all anti-violence movements publicly recognize the contributions of black women. Burke is the founder of Me Too and should be credited as such. Yet even while Burke should indeed be credited for founding both versions of the Me Too movement, the two versions themselves are actually quite different. This difference becomes most evident when answering one key question: Who is the target audience of the movement(s)?

The target audience of the original version of the Me Too movement is survivors. Burke designed the movement to be a way for survivors to connect with and support one another. In Burke’s words, “The work is really about survivors talking to each other and saying, ‘I see you. I support you. I get it.’”²² The original version of Me Too is an intra-survivor movement, meaning that it is a movement made by survivors, for survivors, and involves survivors working with each other. The original version of Me Too is also survivor-centric, meaning that the primary goal of the movement is meeting the needs of survivors themselves, and making sure they have the resources they need to recover. Burke reports that Me Too “is actually about . . . centering victims or survivors of sexual violence, making sure the most marginalized survivors have access to resources, and making sure all survivors have resources to craft their own

healing journey.”²³ Note that Burke’s description of Me Too encourages survivors to support each other, but it does not require that survivors educate non-survivors about sexual violence.

In contrast, the target audience of #MeToo explicitly includes non-survivors. Milano’s original tweet specifies that the purpose of #MeToo is to “give people a sense of the magnitude of the problem.”²⁴ This phrasing indicates that #MeToo is about reaching people who are unaware of issues of sexual violence, which substantially expands the range of the movement beyond survivors. The difference in audience between the two versions of Me Too can appear to be a relatively minor detail, but it fundamentally alters the purpose of Me Too. As a survivor-centric movement, the purpose of Burke’s original version of Me Too is to build resources and support for people who have experienced sexual violence. Milano’s stated purpose was public education through survivor disclosure. As I outline below, disclosure can itself be a source of healing, empowerment, and bonding between survivors, which indicates that #MeToo has in many ways expanded beyond simple public education. Although public education may have been one of the original motivations for #MeToo, the movement has come to mean much more for many survivors. In this article, I want to acknowledge the importance #MeToo has had for many survivors, while also exploring the problematic elements of basing public education about issues of sexual violence on survivor disclosure. I argue that this project of public education through survivor disclosure may unduly burden survivors who do disclose, and creates a moral dilemma for survivors who do not wish to disclose.

II. GOOD SURVIVOR, BAD SURVIVOR

The difference in audience between the two versions of Me Too shifts the roles of survivors in each version of the movement. In Burke’s original version of Me Too, survivors have the opportunity to fill two roles: (1) they can use the movement to access support and resources for their healing journey; and (2) they can themselves become a resource and support to other survivors embarking on their own healing journeys. #MeToo, by contrast, turns survivors who participate in the movement into public educators. This section focuses on the consequences of asking survivors to become public educators. Before beginning, I want to emphasize that my critique in this section applies only to #MeToo, not to Burke’s original version of the movement. Burke’s original version of Me Too does not necessarily ask survivors to become public educators, and thus falls well outside the scope of my critique.

#MeToo centers entirely around survivors disclosing their experiences with sexual violence. The movement does not specify what form the disclosure should take, and survivors participating in #MeToo have taken different approaches to disclosure. But in each case, one central feature remains the same: survivors must identify themselves as survivors. I use the word “must” deliberately. Even though #MeToo has come to mean different things to different survivors, all survivors participating in #MeToo have disclosed that they are survivors. The central feature of the movement is, precisely, disclosure, and most often disclosure through social media platforms. These acts of disclosure have

been crucial to the public education element of #MeToo. The movement has engaged in public education primarily by having survivors identify themselves as survivors in order to demonstrate the vast number of people who have experienced sexual violence, and the different forms their experiences take. Milano's original #MeToo tweet suggests that if all survivors identify themselves, then people will have a better understanding of the prevalence of sexual violence.

In order for this public education project to work, and for the general public to get an accurate picture of just how extensive sexual violence is, two conditions must be met. First, survivors must disclose. #MeToo relies virtually exclusively on survivor disclosure to educate the public about sexual violence. There are few alternative sources of information. If survivors choose not to disclose, then the public education campaign will have relatively limited success. The second condition is that a critical mass of survivors has to disclose. If #MeToo is to convey not just the existence of sexual violence but also the extent of sexual violence, then a substantial number of survivors have to disclose. Without a critical mass of survivors disclosing, the movement cannot convey the widespread nature of sexual violence. The success of the movement is in many ways reliant on the number of survivors who are willing to disclose.

Disclosure can be a source of healing and empowerment for some survivors. Sexual violence has long been closely associated with practices of silencing and denial. Jacobs wrote *Incidents in the Life of a Slave Girl* for an audience of white women,²⁵ at least in part to encourage them to finally acknowledge a history of sexual violence that white women knew about but spoke little of. In 1987, Sandra Harding recounted "the struggle we have had to get women's testimony about rape, wife battering, sexual harassment, and incest experiences accepted as reliable by police, psychiatrists, other men and women, etc."²⁶ More recently, Kate Manne has noted the structural similarities between intimate partner violence and silencing.²⁷ For some survivors, publicly disclosing their experience with sexual violence serves as a powerful way of undermining the testimonial injustices that often accompany acts of sexual violence, and dismantling the very systems that perpetuate sexual violence.

Disclosure can also be a way for survivors to bond with each other, and find sources of comfort, support, and understanding that non-survivors cannot provide. Susan Brison reports that hearing the testimony of other survivors can help survivors to heal from their own trauma: "[survivors] come to have greater compassion for their earlier selves by empathizing with others who experienced similar traumas. They stop blaming themselves by realizing that others who acted or reacted similarly are not blameworthy."²⁸ Some of the more highly publicized survivors who came forward during #MeToo do report that disclosing their experiences with sexual violence has helped them to find unexpected communities with other survivors. For example, Rachel Renock—who disclosed being harassed by an investor and has since testified before Congress on issues of workplace harassment—says that "I had a very special bond with the

other women who have come forward . . . I don't know that I quite expected or understood what kind of bond would come from that. That's been really an incredible force, having a support system."²⁹ In this aspect, public disclosure can sometimes be beneficial for some survivors.

However, acts of disclosure can also jeopardize the well-being of survivors. There are (at least) four problems that make any act of disclosure potentially dangerous to survivors. First, it is not always safe for survivors to disclose. A striking example of the danger of disclosure is sexual violence in US correctional facilities. As I will argue below, the dangers of disclosure can make #MeToo inaccessible for members of marginalized communities, such as incarcerated persons. According to the US Department of Justice, in 2011 alone there were almost 9,000 reported incidents of sexual harassment and assault in adult correctional facilities.³⁰ Given that incidents of sexual harassment and assault in correctional facilities are severely underreported, this number is likely inaccurate. An independent survey conducted with incarcerated persons leaving correctional facilities estimates that 80,600 incarcerated persons experienced sexual violence in 2011–2012.³¹ The difference between these two numbers is telling. If they are both correct, it would imply that over 90 percent of acts of sexual violence in prisons and jails remain unreported.

There are many reasons incarcerated survivors choose to remain silent, but one of the most commonly cited reasons is the very real threat of retaliation. Survivors who choose to report abuse by a staff member must file an administrative report known as a "602" in order to request an investigation of their case, which means they must report their abuse to the administration that is responsible for their abuse. As Diana Block notes, "you are filing the 602 basically with, if not the actual people, the friends of the people, the coworkers of the people, who have abused you."³² And as Jerry Metcalf reports, prison guards often retaliate against survivors who file complaints: "The guards make it a point to label you a 'rat' and destroy your peace of mind and what little you may own in a thousand different ways if you tell on them or one of their coworkers."³³ Even when survivors do choose to report, only 1 percent of correctional facility staff accused of sexual misconduct are ever convicted.³⁴ The US prison system is structured in a way that makes publicly disclosing experiences of sexual violence dangerous for incarcerated survivors. For many incarcerated survivors, disclosure is simply not an option, and this will be of crucial importance for my normative claims in what follows.

The second problem with disclosure, specifically as it relates to #MeToo, is that survivors are being asked to disclose to communities that likely include the people who assaulted them. RAINN reports that 80 percent of survivors know their perpetrators.³⁵ This statistic makes it highly likely that when survivors participating in #MeToo disclose their status using social media, their social media audience will include either the perpetrator themselves or friends and/or acquaintances of the perpetrator. Disclosing to the perpetrator or the perpetrator's community is an entirely different phenomena than disclosing to someone with no connection to the perpetrator, and may affect survivors'

experiences of disclosure. Psychological research on trauma suggests that when confronted about their actions, abuse perpetrators tend to “deny or minimize the abuse, attack the victim’s credibility, and assume a victimized role.”³⁶ Sarah Harsey, Eileen Zurbriggen, and Jennifer Freyd report that individuals who confront perpetrators that used these kinds of evasive tactics experienced an increase in self-blame, which could result in further psychological distress.³⁷ Survivors who publicly disclose to a community that supports or includes their perpetrator could experience evasive or minimizing behaviors that may increase their self-blame. These evasive or minimizing behaviors could become more pronounced in a high-stakes social media disclosure if the perpetrator feels called out in front of their community and feels the need to defend themselves in order to maintain their standing in that community. In these cases, public disclosure could result in increased psychological and emotional harm for survivors.

The third problem with disclosure is that stereotypes about the sexual dangerousness of men of color—especially black men—can make disclosure especially weighted for survivors of color. Survivors of color who are harassed or assaulted by men from their own community may be reluctant to disclose out of concern that their disclosure could be used as fuel for negative stereotypes about men of color. Nellie McKay describes this dilemma as a struggle between sides of the self: “In all of their lives in America . . . black women have felt torn between the loyalties that bind them to race on the one hand, and sex on the other. Choosing one or the other, of course, means taking sides against the self.”³⁸ For survivors of color, disclosure can aggravate racial injustices, harming both themselves and their communities.

Finally, the fourth problem with disclosure is that disclosing itself can be an additional psychological burden on survivors. Burke has repeatedly acknowledged this problem, and the need to address it, recounting that “There’s a cycle that people go through that requires support, even in just saying ‘Me, Too.’”³⁹ Disclosing a history of sexual violence means confronting and addressing highly personal acts that may have had a profound impact on the survivor’s life. Working through the impact of sexual violence in private, or among trusted allies, can be an incredibly difficult experience for many survivors. Publicly addressing this impact to a broad audience on highly visible social media platforms may increase the burden of the already difficult process for some survivors.

The increased burden on survivors is especially true if survivors are asked to educate uninformed non-survivors in addition to disclosing their survivor status. Non-survivors who are shocked or confused by survivors’ testimonies, or simply want to become better informed about issues of sexual violence, may turn to survivors as a source of explanation, information, and advice. Since survivor testimony is the only form of public education #MeToo offers, it makes sense that non-survivors would turn to survivors for help in understanding sexual violence. The problem is that in doing so, non-survivors substantially increase the amount of emotional and other types of labor survivors have in addressing issues of sexual violence. Not

only must survivors find healing from the impact of sexual violence, they must also convince non-survivors to care about issues of sexual violence, and teach them how to be good allies to survivors. Nora Berenstain refers to this additional burden as “epistemic exploitation,” namely, “when privileged persons compel marginalized persons to produce an education or explanation about the nature of the oppression they face.”⁴⁰ Some survivors might be willing to undertake this additional work, but others could find it overwhelming.

There are two key problems with epistemic exploitation in the context of #MeToo. First, as Berenstain correctly argues, epistemic exploitation “maintains structures of oppression by centering the needs and desires of dominant groups and exploiting the emotional and cognitive labor of members of marginalized groups.”⁴¹ Asking survivors to educate non-survivors could thus maintain traditional structures of privilege and marginalization instead of dismantling them. Secondly, as Burke notes, if #MeToo is going to ask survivors to disclose, then it should also be prepared to help survivors manage the burdens of disclosure: “If you make something [like this] viral, you have to be prepared to help people deal. You have to give people something else besides the disclosure.”⁴² The official Me Too website—which is run by Burke and her organization—does provide information about resources survivors can turn to for support. However, #MeToo itself has provided no additional resources for survivors. It has merely asked them to disclose without helping them manage the consequences of disclosure. This lack of support is especially problematic given the disproportionate burden disclosure can have on survivors from marginalized communities.

The tension between the incitement to disclose and the potential harms of disclosure creates a moral dilemma for survivors. #MeToo is structured in a way that provides survivors with two responses to the movement: (1) either they participate in the movement by disclosing their status, or (2) they remain silent and do not directly participate in the movement. Any survivor who wishes to participate in #MeToo must disclose their status as survivors. I do want to note that #MeToo does not explicitly demand that all survivors disclose their status. Survivors can choose not to disclose their status, but doing so means effectively opting out of the movement. A survivor’s silence could even be interpreted as positive assertion that they have *not* experienced sexual harassment or assault. Sanford Goldberg argues that, especially in conditions of oppression, silence is frequently misinterpreted as assent or agreement.⁴³ In the context of a massive outpouring of survivor testimonies, the public might assume that people who do not testify are not survivors.

Even though #MeToo does not explicitly demand disclosure, it has created substantial implicit pressure to disclose. Part of this implicit pressure comes from the unique nature of social media and the way it influences relationships between users. Kathryn Norlock argues that social media platforms serve as a means for users to become part of influential social groups, and increase their standing in those groups. She notes that users have become invested in being perceived in a positive light, and this concern

directly informs how users interact both with each other and with viral social media campaigns, writing that “we cannot help thinking about and caring about the strong and positive reactions to our words online.”⁴⁴ The very fact that #MeToo went viral increases the pressure on survivors to disclose. As the movement gained momentum, it started to achieve a social prominence, and some survivors who disclosed ended up increasing their social standing as a result of their disclosure. For example, Wendy Walsh—the first woman to publicly accuse Bill O’Reilly of sexual misconduct—reports that public disclosure has made her more successful than she was before disclosing. Walsh notes: “Although I was worried this would hurt my business . . . somehow this increased my credibility. Suddenly all kinds of ethical companies wanted me to represent their brands.”⁴⁵ Because of #MeToo’s status as a social media movement that could alter a survivor’s standing within their community, survivors may experience implicit pressure to disclose in order to join such an influential movement. If they want to participate in #MeToo, and potentially gain the social benefits that could come from being identified as a participant in the movement, then survivors must disclose.

The implicit pressure to disclose creates a moral dilemma for survivors who are unwilling or unable to disclose. As detailed above, disclosure can be harmful for some survivors. But the only way any survivor can participate in #MeToo is through public disclosure. The very structure of #MeToo requires survivors to make a morally weighted choice: either they accept the harms that may come from disclosure, or they forfeit their place in the movement. This structure creates an in-group/out-group distinction in #MeToo. Survivors who publicly disclose automatically become part of #MeToo. Survivors who choose not to publicly disclose remain essentially external to the movement.⁴⁶

My concern is that because the movement itself deals with such a high-stakes moral issue, in-group survivors may receive an elevated moral standing that is not available to out-group survivors. Debra Jackson argues that survivor testimony is often evaluated within the context of social scripts: “The rhetorical spaces in which a victim’s testimony is expressed are shaped by social scripts, attitudes, stereotypes, and discourses which are culturally and historically situated.”⁴⁷ #MeToo has—perhaps inadvertently—created a social script for how survivors should behave in relation to their own survivorship: survivors should use their survivorship to raise awareness about sexual harassment and assault, despite the potential risks to their own well-being. Survivors might now be publicly judged based on whether or not they perform this script. Those who perform the script may receive public praise and credit for behaving as good survivors should. By participating in the movement, in-group survivors publicly identify themselves not only as survivors of sexual violence, but also as the right kind of survivors—as advocates against sexual violence, and as part of the solution to changing cultures of sexual violence. By contrast, the out-group survivors who do not perform the script may remain hidden from the public view, and will not necessarily be publicly identified as advocates against sexual violence or as part of the solution to end sexual violence. Out-group survivors

may thus experience little to no public credit for behaving as good survivors, decreasing their moral standing in relation to in-group survivors.

The in-group/out-group structure of #MeToo may thus have created a moral hierarchy that grants survivors moral standing based on public disclosure. The moral standing accorded to survivors increases when they disclose and decreases when they do not, turning survivors who disclose into “better” survivors than survivors who do not. The survivors who disclose become “good survivors”; the survivors who choose not to disclose become comparatively “bad” survivors. Survivorship itself then becomes moralized to the extent that how survivors manage their own survivorship is morally ranked based on whether or not they perform the script of the “good” survivor, and it becomes a way for survivors to gain or lose moral standing in their communities. Good survivors are the ones who say “me too.” Bad survivors are the ones who remain silent.

This moral hierarchy is especially concerning given that #MeToo may be the most inaccessible to survivors from marginalized communities—such as incarcerated survivors. Survivors from marginalized communities are the least likely to be able to disclose without significant backlash, to experience significant uptake of their stories, and to have access to the additional resources they may need to support them if they do disclose. For example, Josephine Yurcaba argues that incarcerated survivors do not fit the stereotypical definition of a good survivor, and thus are less likely to receive sympathy or support from the general public. Yurcaba notes that “those women who have become the face of the [Me Too] movement usually haven’t been charged with crimes. There’s an extra battle that prison abuse survivors face: The public thinks prisoners deserve what they get, and there’s a persistent belief that inmates do not have rights.”⁴⁸ For survivors from marginalized communities, the choice between public disclosure and personal safety may be more extreme, and could lead to more severe consequences than for famous white survivors. Survivors from marginalized communities are thus the most likely to end up at the bottom of the moral hierarchy created by #MeToo. #MeToo may indeed be very beneficial to some survivors. However, the benefits of #MeToo are not equally accessible to all survivors, and I worry that in creating a platform for some survivors to speak, the movement has also created an additional moral burden for survivors who choose to remain silent (often, in not fully autonomous ways).

III. A RETURN TO ORIGINS

In this last section, I suggest that in order to make good on this problem, #MeToo should realign itself with the priorities of Burke’s original version of the movement. I want to emphasize here that I am not arguing that #MeToo should be abandoned. The movement has been a source of support, encouragement, and empowerment for many survivors. As Catharine MacKinnon argues, #MeToo has also brought an unprecedented level of public awareness of and concern for issues of sexual harassment and assault.⁴⁹ Public awareness of and support for sexual harassment and assault may itself be an important precursor to funding increases for organizations that provide resources and

support for dealing with issues of sexual harassment and assault, as well as important recruiting tools for people willing to engage with this kind of work.⁵⁰ Realigning with Burke's original Me Too movement does not mean giving up on the beneficial elements of #MeToo. It does, however, mean shifting the priorities and strategies of #MeToo to be more genuinely inclusive of all survivors.

Realigning #MeToo with Burke's original movement should take two forms. First, #MeToo should prioritize the needs of survivors themselves, including developing resources to help survivors heal. #MeToo currently relies on survivor labor, but it does not center the needs of survivors themselves. Unlike Burke's original version of Me Too, #MeToo largely prioritizes the educational needs of the public instead of the healing needs of survivors. As noted above, while public disclosure can be part of the healing process for some survivors, it alone is not a sufficient healing process for all survivors. Public disclosure should be a viable option for any survivor who wishes to take it, but it cannot be the only option available to survivors. There must be additional resources available to both survivors who disclose and survivors who choose to remain silent. Without these additional resources, #MeToo will be at best incomplete, and at worst potentially harmful to the well-being of survivors.

The second form of realigning that I am suggesting is specifically in regard to marginalized communities. #MeToo should remain profoundly attentive to how sexual violence impacts members of marginalized communities, and should work to ensure that all communities have resources to deal with sexual violence. A core feature of Burke's original version of the movement was this attentiveness to marginalized communities, but this feature has been largely lost in #MeToo. It is important to stress that survivors from marginalized communities have indeed been speaking out,⁵¹ but their stories have not received the same uptake in the mainstream media as those of famous white survivors.⁵² And while some famous white survivors have acknowledged marginalized survivors and survivors who choose not to disclose, this acknowledgment has primarily taken the form of speaking for silent survivors. In a *New York Times* op-ed, Diana Nyad states that "I will continue to tell my story until all girls and women find their own voice."⁵³ In a *Rolling Stone* article, Milano states that "[w]e will succeed if we are the voice of the voiceless," and that "I have a platform that I will continue to use to amplify those who don't have a voice."⁵⁴ Nyad's statement indicates that she intends her voice to stand in for the voice of other survivors. Milano's statement explicitly indicates that she considers herself to be the spokesperson of all survivors, but most especially survivors who have not themselves spoken out—including survivors from marginalized communities.

This kind of speaking for risks reestablishing racist, classist, and nationalist tendencies that have often been problematic in white-led anti-violence movements. As Linda Alcoff argues, speaking for others can have the effect of asserting the power of the speaker over the power of those who are spoken for: "The effect of the practice of speaking for others is often, though not always, erasure and a reinscription of

sexual, national, and other kinds of hierarchies."⁵⁵ Speaking for survivors from marginalized communities overwrites the speaking those survivors themselves have done, and recenters the spotlight on the famous white women who speak for others. Instead of addressing the needs of survivors from marginalized communities, this speaking for can become a way of ensuring that famous white survivors and their needs remain the central focus of #MeToo.

I am not suggesting that famous white survivors should refrain from speaking to issues that impact survivors from marginalized communities since, indeed, they have a lot of social capital that can be leveraged in productive ways. These issues deserve far more attention than they have received, including from more privileged survivors. However, I do recommend that famous white survivors who speak to these issues do so in ways that resist the traditional power dynamic between privileged and marginalized communities instead of reinscribing it. Alcoff suggests that "anyone who speaks for others should only do so out of a concrete analysis of the particular power relations and discursive effects involved."⁵⁶ In the context of #MeToo, this kind of remaining attentive may mean having difficult conversations about how and why privileged survivors continue to receive so much more public attention and support than marginalized survivors, and about how privilege and marginalization can affect the experience of survivorship itself. It may also mean focusing more on engaging with women-of-color-led organizations such as *Sista Il Sista*⁵⁷ and Burke's own Just Be, Inc., which were actively addressing the needs of survivors from marginalized communities long before #MeToo became a trending hashtag.

Realigning #MeToo with Burke's original version of Me Too will have two benefits. First, this realigning will provide the best way forward for survivors themselves. Providing survivors with additional resources will better enable them to move forward on their own healing journeys, instead of focusing most of their energy on educating the public. Second, this realigning will make the movement more accessible to all survivors. Explicitly addressing the needs of survivors from marginalized communities will ensure that the needs and perspectives of all survivors have a place in the movement, not just the needs and perspectives of famous white survivors. If #MeToo is to be more than just a platform for famous white survivors, then it must return to the spirit of Burke's original Me Too movement.

ACKNOWLEDGMENTS

The idea for this article began at the American Philosophical Association Eastern 2019 symposium "Gender Violence as a Problem of Alterity." I would like to thank the attendees of that panel, and especially my co-panelists—Al-Yasha Williams and Rónké Oké—for inspiring this project. Many thanks also to three anonymous reviewers and editor Lauren Freeman for their thoughtful engagement with this article.

NOTES

1. In order to distinguish between the two versions of the Me Too movement, I will refer to the viral social media campaign as #MeToo, and reserve Me Too (without hashtag) for Burke's original version of the movement.
2. Tarana Burke, "The Inception," Just Be Inc., <https://justbeinc.wixsite.com/justbeinc/the-me-too-movement-cmmi>.

3. Ibid.
4. Elizabeth Wellington, "Tarana Burke: Me Too Movement Can't End with a Hashtag," *The Inquirer*, October 23, 2017, https://www.philly.com/philly/columnists/elizabeth_wellington/philly-me-too-movement-founder-tarana-burke-20171023.html.
5. See Danielle L. McGuire, "Recy Taylor, Rosa Parks and the Long History of Black Women's Resistance to Sexual Violence," presented at the #MeToo and Epistemic Injustice Conference, CUNY University, October 5, 2018. Available online at <https://livestream.com/roosevelthouse/metoo-epistemic-injustice/videos/181288887>.
6. Harriet Jacobs, *Incidents in the Life of a Slave Girl* (Oxford: Oxford University Press, 2015), 33.
7. Maria Bevacqua, *Rape on the Public Agenda* (Boston: Northeastern University Press, 2000), 21.
8. Ida B. Wells, *On Lynchings* (New York: Arno Press and the New York Times, 1969), 12.
9. Angela Y. Davis, *Women, Race and Class* (New York: Vintage Books, 1983), 175.
10. Darlene Clark Hine, *Black Women in America*, 2nd ed., online version (Oxford: Oxford University Press, 2005).
11. Ibid.
12. Wellington, "Tarana Burke: Me Too Movement Can't End with a Hashtag." Burke "decided to seek advice from a local rape-crisis center. These girls needed help. But when she got there, she was told rather coldly that the center would only work with girls who were referred from the local police department after they filed a police report. She thought to herself: 'Who is going to do that?'" In 2003, she started Just Be Inc., a nonprofit that helps women heal from sexual assault. She named the movement Me Too in 2007."
13. "About me too," Me Too Movement, <https://metoomvmt.org/about/#history>.
14. See, for example, Kim Q. Hall, "'The Epidemic No One Talks About,'" presented at the #MeToo and Epistemic Injustice Conference, CUNY University, October 5, 2018.
15. Bevacqua, *Rape on the Public Agenda*, 19–20.
16. See, for example, Susan Brownmiller, *Against Our Will* (New York: Simon and Schuster, 1975), 247.
17. See, for example, Beth E. Richie, *Arrested Justice* (New York: New York University Press, 2012); Critical Resistance and INCITE! Women of Color Against Violence, "Gender Violence and the Prison-Industrial Complex," in *Color of Violence*, ed. INCITE! (Durham: Duke University Press, 2006), 223–26.
18. Christie D'Zurilla, "In saying 'MeToo,' Alyssa Milano Pushes Awareness Campaign about Sexual Assault and Harassment," *The Los Angeles Times*, October 16, 2017, <http://www.latimes.com/entertainment/la-et-entertainment-news-updates-metoo-campaign-me-too-alyssa-milano-1508173882-htmistory.html>.
19. "MeToo Creator Tarana Burke and Alyssa Milano: 'It's Not Just a Moment, It's a Movement,'" *Today*, December 6, 2017, https://www.youtube.com/watch?v=MxuZB-K_Q1A.
20. Cassandra Santiago and Doug Criss, "An Activist, a Little Girl and the Heartbreaking Origin of 'Me Too,'" CNN, October 17, 2017, <https://www-m.cnn.com/2017/10/17/us/me-too-tarana-burke-origin-trnd/index.html>.
21. J'na Jefferson, "A Long Road Ahead: #MeToo Founder Tarana Burke on Sexual Assault, Stigmas and Society," *Vibe*, April 3, 2018, <https://www.vibe.com/featured/tarana-burke-me-too-feature>.
22. Wellington, "Tarana Burke: Me Too Movement Can't End with a Hashtag."
23. J'na Jefferson, "A Long Road Ahead."
24. D'Zurilla, "In Saying 'MeToo,' Alyssa Milano Pushes Awareness Campaign about Sexual Assault and Harassment."
25. See C. Riley Snorton, *Black on Both Sides* (Minneapolis: University of Minnesota Press, 2017), 67.
26. Sandra Harding, "Ascetic Intellectual Opportunities" in *Science, Morality, and Feminist Theory*, ed. Marsha P. Hanen and Kai Nielsen (Calgary: University of Calgary Press, 1987), 77.
27. Kate Manne, *Down Girl* (Oxford: Oxford University Press, 2018), 3–4.
28. Susan J. Brison, *Aftermath* (Princeton: Princeton University Press, 2002), 63.
29. "How Saying #MeToo Changed Their Lives," *The New York Times*, June 28, 2018, <https://www.nytimes.com/interactive/2018/06/28/arts/metoo-movement-stories.html>.
30. Allen J. Beck, Ramona R. Rantala, and Jessica Rexroat, "Sexual Victimization Reported by Adult Correctional Authorities, 2009–11," US Department of Justice Bureau of Justice Statistics, January 2014, <https://www.bjs.gov/content/pub/pdf/svraca0911.pdf>.
31. Valerie Bauman, "The Forgotten Voices of #MeToo" Just Detention International, December 10, 2018, <https://justdetention.org/the-forgotten-voices-of-metoo-prison-rape-in-america-perpetrated-by-both-inmates-and-staff-remains-pervasive-and-under-reported-15-years-after-the-passage-of-a-law-intended-to-end-crisis/>.
32. See Marisa Endicott, "'No Longer Human': Women's Prisons Are a Breedings Ground for Sexual Harassment, Abuse," *Think Progress*, August 29, 2018, <https://thinkprogress.org/sexual-harassment-abuse-womens-prisons-me-too-5231b62c1785/>.
33. Jerry Metcalf, "When Prisoners Say #MeToo," *The Marshall Project*, September 20, 2018, <https://www.themarshallproject.org/2018/09/20/when-prisoners-say-metoo>.
34. Josephine Yurcaba, "For Survivors of Prison Rape, Saying 'Me Too' Isn't an Option," *Rewire News*, January 8, 2018, <https://rewire.news/article/2018/01/08/survivors-prison-rape-saying-isnt-option/>.
35. "Perpetrators of Sexual Violence: Statistics," *RAINN*, <https://www.rainn.org/statistics/perpetrators-sexual-violence>. The notable exception to this statistic is Indigenous women. Indigenous women are more likely to experience sexual violence than women of any other race or ethnicity, and they are most often assaulted by perpetrators from outside their own communities (see "Victims of Sexual Violence: Statistics," *RAINN*, <https://www.rainn.org/statistics/victims-sexual-violence>).
36. Sarah J. Harsey, Eileen L. Zurbriggen, and Jennifer J. Freyd, "Perpetrator Responses to Victim Confrontation," *Journal of Aggression, Maltreatment, & Trauma* 26, no. 6 (2017): 645.
37. Ibid., 656 and 658.
38. Nellie McKay, "Remembering Anita Hill and Clarence Thomas," in *Race-ing Justice, En-gendering Power: Essays on Anita Hill, Clarence Thomas, and the Construction of Social Reality*, ed. Toni Morrison (New York: Pantheon Books, 1992), 277–78.
39. Abby Ohlheiser, "The Woman Behind 'Me Too' Knew the Power of the Phrase When She Created It—10 Years Ago," *The Washington Post*, October 19, 2017, <https://www.washingtonpost.com/news/the-intersect/wp/2017/10/19/the-woman-behind-me-too-knew-the-power-of-the-phrase-when-she-created-it-10-years-ago/>.
40. Nora Berenstain, "Epistemic Exploitation," *Ergo* 3, no. 22 (2016): 570.
41. Ibid.
42. Wellington, "Tarana Burke: Me Too Movement Can't End with a Hashtag."
43. Sanford C. Goldberg, "Arrogance, Silence, and Silencing," *Aristotelian Society Supplementary Volume* 90, no. 1 (June 2016): 96.
44. Kathryn J. Norlock, "Online Shaming," *Social Philosophy Today* 33 (2017): 193.
45. "How Saying #MeToo Changed Their Lives." Note that not all survivors who publicly disclose experience this same increase in social standing. Gabrielle Union has importantly pointed out that while many famous white women have received a lot of uptake through #MeToo, the stories of women of color have received considerably less attention. See Bethonie Butler, "Gabrielle Union on #MeToo: 'The Floodgates Have Opened for White Women,'" *The Washington Post*, December 8, 2017, <https://www.washingtonpost.com/news/arts-and-entertainment/>

wp/2017/12/08/gabrielle-union-on-metoo-the-floodgates-have-opened-for-white-women/.

46. Some scholars argue that in at least some contexts, survivors may have a duty to disclose in spite of the potential harms of disclosure (see, for example, Ashwini Vasanthakumar, "Epistemic Privilege and Victims' Duties to Resist their Oppression," *Journal of Applied Philosophy* 35, no. 3 [2018]: 465–80). Survivors' duties to disclose is an important subject of conversation that is well beyond the scope of this article. For the purposes of this article, I'm most concerned with exploring the impact #MeToo has on survivors, instead of whether or not survivors have a duty to disclose. In other words, I'm more interested in evaluating #MeToo than in evaluating the duties of survivors themselves.
47. Debra L. Jackson, "'Me Too': Epistemic Injustice and the Struggle for Recognition," *Feminist Philosophy Quarterly* 4, no. 4 (2018): 5.
48. Yurcaba, "For Survivors of Prison Rape, Saying 'Me Too' Isn't an Option."
49. Catharine A. MacKinnon, "#MeToo Has Done What the Law Could Not," *The New York Times*, February 4, 2018, <https://www.nytimes.com/2018/02/04/opinion/metoo-law-legal-system.html>.
50. Thanks to an anonymous reviewer for this point.
51. See, for example, the documentary *Surviving R. Kelly*.
52. In some ways, #MeToo has counteracted its own goal of public education. Since the movement has tended to be most accessible and receptive to the stories of famous white survivors, it may have perpetuated the false notion that the way famous white survivors experience sexual harassment and assault is universal to all survivors. Thanks to an anonymous reviewer for this point.
53. Diana Nyad, "My Life After Sexual Assault," *The New York Times*, November 9, 2017, <https://www.nytimes.com/2017/11/09/opinion/diana-nyad-sexual-assault.html>.
54. Alyssa Milano, "Alyssa Milano on Joining Time's Up," *Rolling Stone*, January 4, 2018, <https://www.rollingstone.com/culture/culture-news/alyssa-milano-on-joining-times-up-women-are-scared-women-are-angry-204035/>.
55. Linda Alcoff, "The Problem of Speaking for Others," *Cultural Critique* 20 (Winter 1991–1992): 29.
56. *Ibid.*, 24.
57. See Sista II Sista, "Sista's Makin' Moves" in *Color of Violence*, ed. INCITE! Women of Color Against Violence (Durham: Duke University Press, 2016), 196–207.

Beyond Silence, Towards Refusal: The Epistemic Possibilities of #MeToo

Sarah Clark Miller
PENNSYLVANIA STATE UNIVERSITY

On September 27, 2018, I sat watching the painfully careful and exceedingly brave way in which Dr. Christine Blasey Ford navigated testifying in front of the Senate Judiciary Committee regarding now Supreme Court Justice Brett Kavanaugh's alleged assault of her. As this took place, something happened: survivors of sexual violence—friends and acquaintances, primarily women—began to post their own stories of sexual violence on Facebook and Twitter. The strength that Dr. Blasey Ford demonstrated inspired many survivors to disclose the details of their stories, many for the first time. As they did so, the sense of solidarity among them grew. It was not entirely unlike and was undoubtedly inspired by the #MeToo moment that preceded it nearly a year before in October 2017 when the hashtag exploded on social media. What was originally called the #MeToo moment had grown into a movement.

Dr. Blasey Ford's testimony was also personally significant. In that moment of watching her and seeing the many posts catalyzed by her testimony, something in me shifted. It is what I have come to understand as a moment of tremendous epistemic refusal galvanized by the similar gestures of refusal I saw all around me. For decades, I had remained largely quiet about my own sexual assault. I had done so for the reasons that many victims of sexual violence do: because of the threats, obstacles, and harms that sexual violence survivors commonly face in rape culture. These include how we are pressured to remain silent and how our confidence in our knowledge about the violence we endured is undermined. This silencing and undermining are accomplished through widespread practices of credibility erosion as well as ostracization and shaming. Sexual violence survivors can be harmed both with regard to their ability to know the truth of their own experiences, as well as the ability to share their knowledge with others. Spurred on by the bravery of other survivors and a desire to exhibit the same form of bravery as a gesture of solidarity, I decided to break my silence.

So, I disclosed to 1,020 or so of my closest friends on Facebook that I, too, was a rape survivor. This is what I wrote:

Dr. Blasey Ford's story is much like my own, except I was unable to escape from my perpetrator, who raped me when I was 16. It was a violent event that shattered my life and shaped much of who I am today.

I have spent over two decades largely silent and ashamed. Not any longer. Not after today. After what I just witnessed, I stand firmly in solidarity and in strength with all survivors of sexual violence and will no longer be ashamed about what he did to me. That shame was never mine. It was always his.

My story was one in a deluge of stories that poured forth publicly on that day. In coming forward, I added my voice to a chorus of millions who had already spoken on social media as part of the wider #MeToo movement. What became clear to me in that moment was the power of the #MeToo movement to affect not only cultural, ethical, and political change, but equally importantly, *epistemic* transformation. It is that epistemic transformation that I want to consider in this article. For me, the drive to understand the possibilities of such transformation is rooted in a desire to understand my own epistemic shift, as well as the wider meaning of what I saw happening all around me on that day.

There are many ways to understand the meanings of the #MeToo movement. Analyses of its significance have proliferated in popular media; some academic analyses have also recently appeared.¹ Commentary on the philosophical and epistemic significance of the #MeToo movement has been less plentiful.² The specific moment of the #MeToo movement in which Dr. Blasey Ford's testimony garnered a widespread social media response from sexual violence survivors highlighted the power of a particular form of epistemic response, what I call

“epistemic refusal.” In breaking our silence, those of us who are sexual violence survivors have used this strategy to refuse the dominant epistemic structures that have kept us tightly in check. Mass informal disclosure of survivor status represented in conjunction with hashtags such as #MeToo, #WhyIDidntReport, and #BelieveHer creates space for epistemic, ethical, and political community between survivors of sexual violence by denying hegemonic epistemic discourses of contemporary rape culture. Regarding Dr. Blasey Ford’s testimony and the social media response it garnered, analysis of three main elements proves particularly illuminating: the nature of mass informal disclosure of sexual violence, what the hashtag #WhyIDidntReport reveals about refusing silence, and what the hashtags #BelieveHer and #BelieveSurvivors can show us about what it takes to begin to overcome epistemic gaslighting.

MASS INFORMAL DISCLOSURE

The advent of the #MeToo Movement in October 2017 and the occasion of Dr. Blasey Ford’s testimony represented the power of disclosure and, more specifically and technically, mass informal disclosure as a mechanism of epistemic refusal. Mass informal disclosure of sexual violence takes place when a survivor divulges information regarding their assault to a large group of people, often in a public context, absent the intention of that disclosure having some form of official implication or effect. Undoubtedly, the moment of #MeToo and #WhyIDidntReport gave rise to other forms of disclosure, too—countless private moments between friends, spouses, child and parent in which those who had been violated shared what had happened to them. These private, informal modes of disclosure represent a form of disclosure with a more limited scope than is found with mass informal disclosure. #MeToo and #BelieveHer presumably also gave rise to some cases of formal disclosure too—the realization that one was one of many survivors may have provided some impetus for the lodging of formal complaints—for example, with the police—regarding the sexual violence one had suffered. Formal modes of disclosure differ from mass informal disclosure in terms of the institutional or bureaucratic power and intervention they activate, sometimes unwantedly.

There is arguably always vulnerability in the face of the disclosure of sexual violence, no matter what its form, though the nature of that vulnerability will differ based on the kind and circumstances of disclosure.³ Similarly, choosing to disclose one’s status as a survivor is nearly always risky, though the extent and forms of associated risk can also differ. In its peculiar form of being simultaneously a disclosure to no one and everyone, mass informal disclosure is not without its own specific kinds of vulnerability and risk. Should the intention behind the disclosure be one of hoping for a public embracing of one’s status as a survivor of sexual violence (an approach I do not recommend), then one can be vulnerable to everything from shaming to recrimination to indifference from those to whom one blasts one’s status. The risks are myriad, as well, and include a possible reorienting of how others see you—layered now with pity, discomfort, or even fear. Other possible risks include all of the forces that kept you from saying anything in the first place—threats

of violence, being called a liar or a whore, retraumatizing oneself through the process of disclosure, internalizing all of these consequences, and so on.

The predominant mode of disclosure catalyzed through the hashtags #WhyIDidntReport and #BelieveHer or #BelieveSurvivors was mass informal disclosure. And it is the epistemic space and possibilities that mass informal disclosure can open that I explore here. The mass informal disclosure that survivors engaged in during and after Dr. Blasey Ford’s testimony embodied epistemic refusal in a couple of respects. They served as examples not so much of *resisting* our collective rape culture, in the sense of pushing back against it, as *refusing* its logic and outcomes entirely. Rather than pushing back against prejudicial epistemic standards that harm survivors, such as testimonial injustice and credibility deficit, those who engage in mass informal disclosure assert their experiences of violation in a way that denies the importance of epistemic uptake from others, hence beginning to reclaim the autonomy and power that was taken from them. Mass informal disclosure of this nature involves an implicit shift in the terms of epistemic exchange—it is an articulation of personal knowledge absent the requirement of epistemic uptake of that knowledge by the broader public—perhaps most especially those likely to doubt the credibility of rape survivors. Instead of further engaging the various ways in which rape victims are made to hustle for their worth through prejudicial epistemic standards, those who disclose begin to create a different knowledge economy by generating epistemic spaces of their own.

Of equal importance is how forms of mass informal disclosure of sexual violation can shift the focus from the credibility of the survivor to the wrongful actions of the perpetrator—moving feelings of shame and responsibility away from victim and back onto perpetrator. This was true in my own case: the form of mass informal disclosure in which I engaged served to take a festering bucket of shame and firmly shove it where it always should have been—with the person who raped me. And mass informal disclosure can also lay the groundwork for the creation of communities of solidarity among survivors of sexual violence, which is an aspect of what Tarana Burke has called for and represents one intention behind the important work she has done. This is one way of understanding the kind of epistemic spaces the #MeToo movement can create. These communities afford the creation of a different sort of epistemic subculture for and between survivors—a subculture that can be governed not by dominant epistemic structures that are prejudicial against survivors but rather by the forms of mutual understanding of both the harms of sexual violence and the epistemic injustices that keep survivors silent about those harms.

A CONTRAST IN #HASHTAGS

What I am designating as the second major moment of the #MeToo Movement, which involved Dr. Blasey Ford’s testimony and the mass informal disclosures that it engendered, differed in several key regards from the initial #MeToo moment. One way to characterize that difference is through the associated hashtags as well as the meanings those hashtags were meant to carry. #MeToo was and is a

way of saying, this also happened to me. It was a way of raising one's hand, as Tarana Burke has noted.⁴ One aspect of the impactful force of #MeToo was how it shattered some of the layers of denial around the sheer magnitude of sexual violation, bringing to light what many of us have long known: sexual violence is pervasive, systemic, and not at all rare. #MeToo garnered more than 12 million posts, comments, and reactions on Facebook alone in less than 24 hours, hence showing the world—or at least those who cared to take notice—that sexual violence is a shockingly widespread phenomenon. It was also a pivotal moment for the ways in which it helped survivors feel much less isolated, as they could concretely see that they were very much not alone. Thus, the advent of the #MeToo hashtag was perhaps more about quantity, that is, the overwhelming numbers of women who signaled in tweets and Facebook posts that they, too, had been victims of sexual violence, often without providing contextualizing details. In fact, this was one part of the power of it—the ability to disclose without needing to offer extensive details designed to justify the disclosure, to forestall challenges to one's credibility, or to seek affirmation.

In the second main moment of the movement, survivors who shared their own experiences during and after Dr. Blasey Ford's testimony often appended the hashtag #WhyIDidntReport as they detailed the myriad challenges survivors experience regarding formal disclosure in the wake of a sexual violation. And supporters of both Dr. Blasey Ford and the many others who told their stories on that day affixed #BelieveHer or #BelieveSurvivors to their posts, to emphasize the importance of giving proper epistemic weight to the claims of violation sexual violence survivors make. We can therefore see that Dr. Blasey Ford called forth a response that wasn't primarily about sheer quantity—some victims had, after all, already disclosed during the earlier #MeToo moment. Instead, it was about quality and circumstance in the sense that many focused on the contextual details of the obstacles to reporting. Concurrently, #BelieveHer and #BelieveSurvivors represented a way to signal support for all of those who came forth on that day and to emphasize the importance of confronting patterns of credibility deficit that survivors so often experience.

#WHYIDIDNTREPORT

The hashtag associated with Dr. Blasey Ford's testimony, #WhyIDidntReport, highlights specific epistemic angles of the difficulties of disclosure in a society steeped in rape culture. There are many reasons why survivors of sexual assault don't report. What epistemic analysis of the underreporting of sexual violence highlights is at least twofold: both the widespread epistemic silencing and the epistemic gaslighting victims endure. One overarching way to understand social media posts that feature #WhyIDidntReport is as a strategy of epistemic refusal in response to epistemic silencing. (Gaslighting plays an interesting role for both #WhyIDidntReport and #BelieveHer/#BelieveSurvivors. Gaslighting in the context of the latter will be my focus in the next section.) Survivors deploy mass informal disclosure to lay bare the mechanisms of silencing and gaslighting they and others encounter. Rather than continuing to remain silent and doubtful about

their own experiences or to rail against dominant epistemic systems, those who use #WhyIDidntReport do something else: they reveal the very mechanisms of epistemic oppression that so profoundly harm them for all to see.

While #WhyIDidntReport was designed to explore the barricades to formal reporting, it also represents an opportunity to consider the broader patterns of silencing survivors experience. There are different ways we don't tell and a multitude of reasons why we remain silent. In a sense, then, #WhyIDidntReport explores the vast and varied temporal landscape prior to disclosure, which is not to assume that disclosure is an inevitability—far from it. Disclosure also isn't a toggle switch. While moments of mass informal disclosure tend to have a certain sense of loudness or force about them, there can be many quieter, private moments of partial or attempted (and often inevitably thwarted) disclosure that survivors face. Thus, #WhyIDidntReport can tell us a lot about the risks and vulnerabilities that survivors know are there when they contemplate telling others they were violated.

As for my own story, the truth of what happened to me attempted to bubble up multiple times in my late teens, only to be forcibly stuffed back down. In addition to being outright threatened for my attempted informal disclosure, the indignities I encountered included being ostracized for being supposedly promiscuous and being told that I was lying in an attempt to stoke drama in my otherwise apparently very boring life. On the one hand, this reads as the ordinary drama of middle-class suburban girlhood. On the other hand, it is absolutely appalling that such an experience reads as ordinary drama at all. I was subjected to multiple instances of outright aggression that threatened my well-being and social standing in my small-town Pennsylvanian community. I was threatened in ways that attempted to shatter my very sense of self. After experiences like that, formal reporting seemed an absolute impossibility.

My experience shows how the forces of violence visited upon survivors represent a complicated intermingling of the formal and informal. In my particular case, the informal mechanisms of social policing designed to maintain the impunity of boys and men functioned to ensure that the costs of formal reporting would be so high as to appear unfathomable. This is not to say that once survivors feel they can report, they will then be safe in doing so. The power that institutions such as law enforcement and the medical establishment possess can be, and all too frequently is, wielded against survivors themselves. While, in theory, institutional power may be designed to protect survivors, it can and often does, in fact, exacerbate their vulnerability. This can happen either by the devastating forms of resistance to the truth of the prevalence of sexual violence that institutions can enact or by a similarly damaging violation of survivor autonomy when institutions insist on the details of disclosure being shared and pursued in case they are formally actionable—as we find with contemporary mandatory reporting requirements on today's college and university campuses.

We can bring some of the difficulties of #WhyIDidntReport into focus by peering through lenses of epistemic analysis. The first lens is that of epistemic silencing. Among other epistemic feats rape culture accomplishes, silencing is one of the most pernicious. #WhyIDidntReport points to specific practices of epistemic silencing, which Kristie Dotson, following Gayatri Spivak understands as “a type of violence that attempts to eliminate knowledge possessed by marginal subjects...”⁵ There is more than one way to silence a survivor of sexual violence. Dotson identifies two primary ways of silencing: testimonial quieting and testimonial smothering.⁶ Testimonial quieting takes place “when an audience fails to identify a speaker as a knower. A speaker needs an audience to identify, or at least recognize, her as a knower in order to offer testimony.”⁷ If you don’t believe a particular sexual violence survivor has the epistemic authority to be a knower in the first place, and are therefore incapable of identifying her as someone who could have reliable, valuable knowledge to share, why would you bother to listen to what she has to say? It is important to note that survivors of sexual violence may experience differing degrees of testimonial quieting based on the particular social position they occupy, as well as how their speaking might be perceived to support or hinder patriarchal aims. The degree of testimonial injustice visited on a survivor through forms of silencing often varies based on the particular race, sex, gender, and/or class of that survivor.

The second variety of silencing that helps to shed light on the epistemic mechanisms that cause survivors not to report is testimonial smothering. Testimonial smothering “occurs because the speaker perceives one’s immediate audience as unwilling or unable to gain the appropriate uptake of proffered testimony. . . . Testimonial smothering, ultimately, is the truncating of one’s own testimony in order to insure that the testimony contains only content for which one’s audience demonstrates testimonial competence.”⁸ When people are raised in rape culture, they are epistemically conditioned to a kind of ignorance that renders them incapable of receiving some forms of testimony from sexual violence survivors. Many survivors already know this and therefore deliver only part of their experience—the parts that have some chance of being heard. This is a second way in which epistemological silencing makes plain why survivors often don’t disclose and report. Or, if they do disclose, it makes clear why they sometimes choose only to disclose slivers of their experience.

When survivors deploy the hashtag #WhyIDidntReport, they reveal the mechanisms of testimonial quieting and smothering, thereby beginning to refuse the epistemic terms of engagement to which they have long been subjected. They refuse testimonial quieting by asserting their status as knowers. And they refuse testimonial smothering by asserting what they know without caring whether the broader audience has testimonial competence. In that moment, they are far more interested in connecting and building epistemic community with other survivors who believe what they know and who possess the testimonial competence to hear what they are saying.

BEYOND #BELIEVEHER AND #BELIEVESURVIVORS

Epistemic refusal is also present in how sexual violence survivors come to believe the truth of their own experiences, a process that can be thwarted mightily by another mechanism of epistemic oppression: epistemic gaslighting. Kate Abramson characterizes gaslighting as “a form of emotional manipulation in which the gaslighter tries (consciously or not) to induce in someone the sense that her reactions, perceptions, memories and/or beliefs are not just mistaken, but utterly without grounds—paradigmatically, so unfounded as to qualify as crazy. Gaslighting is . . . quite unlike dismissing someone, for dismissal simply fails to take another seriously as an interlocutor, whereas gaslighting is aimed at getting another not to take herself seriously as an interlocutor.”⁹ While Abramson characterizes gaslighting in the quote above primarily as a form of emotional manipulation, I take it to be in the spirit of what she is saying to assert that it is also very much a form of epistemic manipulation. Sexual violence survivors are frequently subjected to gaslighting of both an emotional and epistemic nature. They are told they are overreacting. It wasn’t that bad. They are generating false memories. They were too drunk to really recall. They were too emotional to see the situation clearly, etc.

It is in these ways and more that sexual violence survivors’ view of their own epistemic authority is undermined and sometimes outright obliterated. When one is told over and over again that what they thought had happened did not, in fact, actually happen, and that their belief that it did arises through their inability to properly perceive the true nature of experience because they are crazy, too emotional, inherently deceitful, etc., they all too often start to believe that they did not experience what they, in fact, actually did. Even more perniciously, they will begin to internalize the very mechanism that destabilizes and can obliterate their own sense of epistemic credibility and authority. That is to say, they will do it to themselves.

When sexual violence survivors engage in mass informal disclosure of their violation, they refuse the terms of engagement necessary to get gaslighting off the ground. Mass informal disclosure amounts to a broadcasting of their epistemic confidence in themselves and of their self-conception as agents with epistemic authority. They thereby refuse to buy into the undermining of the knowledge they have of their own experiences. They refuse to let others undermine their self-trust. They, in short, refuse to be gaslighted.

It is in view of these very common and utterly destructive experiences of epistemic gaslighting that I want to respond to the prevalence of #BelieveHer and #BelieveSurvivors. Both hashtags offer an interesting window into this particular problem. The intentions behind #BelieveHer and #BelieveSurvivors are admirable ones—those who use such hashtags want to signal that they support survivors in trusting themselves. Use of both hashtags also functions as an implicit recognition of the pervasive epistemic gaslighting survivors encounter—a form of recognition that is very much needed. And it offers the beginnings of a shift away from a dominant framework of credibility in which survivors tend to come out on the losing end.

While appreciating the goodness of these intentions, I want to assert that in light of the pervasive nature of the gaslighting of sexual violence survivors, what is equally, if not more important is that survivors believe themselves and that they come to believe one another by stepping into a shared epistemic space of their own creation. This is to say that while believing her and believing survivors are both important things to do, there is perhaps something even more significant to accomplish. It is something that I take to be a paradigmatic move of epistemic refusal: for survivors to center on their own knowledge, build their self-trust and trust in one another, and in doing so, create epistemic communities through which they can further support one another.

Dr. Blasey Ford serves as a potent example in this regard: it was in and through her characterization of herself as “100 percent” sure that Kavanaugh was her attacker—even though some details of the evening in question were fuzzy—that other survivors could step into the truth of their own experiences, the certainty of their own testimony, and the power of their own epistemic authority. It is in and through such self-trust, truth, and power that we begin to refuse the epistemic deck that for so long has been so carefully and relentlessly stacked against us.

ACKNOWLEDGMENTS

I would like to thank Lauren Freeman and Lori Watson for providing feedback on this paper.

NOTES

1. See, for example, Michelle Rodino-Colocino, “Me Too, #MeToo: Countering Cruelty with Empathy,” *Communication and Critical Cultural Studies* 15, no. 1 (2018): 96–100; Reshma Jagsi, “Sexual Harassment in Medicine—#MeToo,” *The New England Journal of Medicine* 378, 3 (2018): 209–11; Kaitlynn Mendes, Jessica Ringrose, and Jessalyn Keller, “#MeToo and the Promise and Pitfalls of Challenging Rape Culture through Digital Feminist Activism,” *European Journal of Women’s Studies* 25, no. 2 (2018): 236–46; Ritty Lukose, “Decolonizing Feminism in the #MeToo Era,” *The Cambridge Journal of Anthropology* 36 (2018): 34–52; and Martha Chamallas, “Will Tort Law Have Its #MeToo Moment?” *Journal of Tort Law* 11, no. 1 (2018): 39–70.
2. There are some examples from within philosophy of analysis of the #MeToo Movement, most notably “#MeToo and Epistemic Injustice,” a conference organized in fall 2018 by Linda Martín Alcoff and Charles W. Mills. See also Linda Martín Alcoff, *Rape and Resistance* (Medford, MA: Polity, 2018); and Debra Jackson, “Me Too’: Epistemic Injustice and the Struggle for Recognition,” *Feminist Philosophical Quarterly* 4, no. 4 (2018): article 7.
3. See Ann J. Cahill, “Disclosing and Experience of Sexual Assault: Ethics and the Role of the Confidant,” in *Vulnerability and the Politics of Care* (Oxford University Press, forthcoming).
4. Liz Rowley, “The Architect of #MeToo Says the Movement Has Lost Its Way,” *New York Magazine*, October 23, 2018, <https://www.thecut.com/2018/10/tarana-burke-me-too-founder-movement-has-lost-its-way.html>, accessed 5/1/19.
5. Kristie Dotson, “Tracking Epistemic Violence, Tracking Practices of Silencing,” *Hypatia* 26, no. 2 (Spring 2011): 236.
6. *Ibid.*, 242–51.
7. *Ibid.*, 242.
8. *Ibid.*, 244.
9. Kate Abramson, “Turning Up the Lights on Gaslighting,” *Philosophical Perspectives* 28, no. 1 (2014): 2.

The Speech Acts of #MeToo

Cassie Herbert

ILLINOIS STATE UNIVERSITY

In the fall of 2017 I, along with many others, watched and participated as #MeToo unfolded across social media. Women, nonbinary folks, and some men shared their experiences of being sexually harassed or assaulted. Some went into gut-wrenching detail. Some simply wrote the hashtag #MeToo. Some named their assailants, though most did not. As more and more people chimed in, I remember being struck by the realization that *I did not know a single woman who hadn’t been sexually violated in some way*. This thought was likely too sweeping, but it speaks to how overwhelming the moment was: it felt like everyone had a horror story. One month after it gained prominence, a Facebook estimate held that 45 percent of US users had at least one friend who had posted the statement “me too.”¹ One year later, the hashtag had been used at least 19 million times across Twitter.² In the time since, the movement has continued to maintain a presence in the global social landscape.

#MeToo took work and energy; it took courage to speak up, and attending to the flood of stories day after day could be emotionally draining. It required emotional and epistemic labor to process the stories coming out, to sort through the various responses to those stories, and to figure out what to do next. For many it was retraumatizing, as survivors³ relived their own violations by sharing their experiences publicly or by taking in the stories told by others. The movement made stark the extent of sexual violations occurring around us. Survivors and others performed this difficult labor with the hope that it would make some sort of difference.

One of the repeated questions to come up in the wake of the movement is, “Has #MeToo helped?” Or, put slightly differently, “Has #MeToo been successful?” Despite all the experiences of sexual harassment and abuse shared on social media, there hasn’t been a corresponding significant uptick in holding perpetrators accountable for the sexual violations they commit. Some high profile figures lost their jobs (though often, the severance package they received would hardly be termed a hardship), some people were publicly censured (though think pieces are already heralding their “comebacks”), and a very few people have had legal charges filed against them (though it’s not clear if anyone in the US, to date, has actually been prosecuted or found guilty as a result of #MeToo). If these are the measures of success, then it seems like #MeToo hasn’t succeeded at much.

I hold that these metrics are the wrong way to evaluate #MeToo. Not simply because they focus on the wrong things (centering perpetrators rather than, say, the support given to survivors), but also because they misunderstand what kind of actions made up the movement. We need to have a better sense of what the movement was in order to evaluate the success of the movement. Speech act theory can help us do this. In what follows, I explore two ways of understanding the social media posts that comprised

#MeToo: as accusations and as reports. *Accusations* are speech acts that call for holding the perpetrator to account, while *reports* are truth-claims about the world rooted in the speaker's first-personal experience. #MeToo, I argue, ought to be understood as primarily composed of reports. This better enables us to evaluate the successes of the movement and, in fact, allows us to recognize the ways in which the movement signifies a noteworthy shift in the dominant frameworks by which survivors' speech is understood.

I. SPEECH ACTS, REPORTS, AND ACCUSATIONS

Before discussing #MeToo, I first need to lay out the core features of speech act theory. As J. L. Austin and others have pointed out, we do more with speech than simply put forward information. Speech is an action, and through this action we make real normative changes to the world. (Note that "speech" here is maximally inclusive, and involves spoken interactions along with Tweets, texts, social media posts, and any other form of communication.) We can command, invite, entreat, persuade, call to order, or any number of things via speech. Each of these actions has a particular performative structure. To start with a familiar example, a speaker must have the appropriate standing to issue a command and must issue it to the right people, in the right context, for it to succeed. A professor can command a student to turn in their assigned paper, but that student cannot order the professor to return their exam. However, the professor cannot command a student who is not in the course to turn in a paper and they cannot arbitrarily order their student to wear purple, eat bananas for breakfast, or any number of things the professor simply has no authority over. Drawing on the framework developed by Rebecca Kukla and Mark Lance, while commands are agent-relative and require certain authority conditions to be felicitous, other speech acts are agent-neutral in their inputs—anyone can properly issue them, regardless of whether or not they have authority in that context.⁴ A declarative, for example, fits this mold: declaratives are publicly accessible truth-claims about the world ("The sky is blue," or "Trump is the 45th president.") Anyone may issue these speech acts, it doesn't take a particular kind of standing or authority to do so.

Not only are speech acts distinct from one another based on their entry conditions and to whom they can be directed, but they differ by their function and what they do. Declaratives call on the audience to incorporate the truth-claim into their belief structure.⁵ Going back to commands, a command, issued with the proper authority, in the right context, and to the right person, creates a requirement for that person to do something. A boss commanding their underling to attend a meeting properly expects the underling to comply; if they don't, they have responded inappropriately. (Of course, sometimes life happens and we miss meetings—but when that happens we offer excuses and apologies in recognition of having failed to do something we ought to have done.) On the other hand, an invitation creates an option which one can take up or not, and both are appropriate responses. An invitation to coffee between friends does not automatically require attendance—the whole purpose of an invitation is that it is optional, and one would be dismayed to find out a friend

came because they felt they had to. Yet an invitation is not a neutral laying out of the options either: someone has offered us something by extending the invitation and they are properly owed gratitude, however minimal, for doing so ("Thanks for the invite, but I just grabbed coffee a bit ago!").⁶ Invitations are agent-relative in their outputs; only the person to whom it is issued is authorized to take up the invitation (a stranger who overhears me invite my friend to coffee has not, themselves, been invited to join me for coffee.) An invitation is not simply a weak command or one issued by someone with only marginal authority. Invitations and commands have different performative structures and pragmatic outputs, even when they both aim at the same result of getting someone to show up for a particular event. Even when the effects are similar, the particular speech act performed makes a difference in *why* the person is there, whether they ought to thank the other for the invitation, the tenor of their ongoing interactions, etc.

Crucially, speech acts can be indistinguishable at the level of surface-grammar. "Gee, I sure am cold" might be a mere statement said to the room at large or, if said while making pointed eye contact with the person near the open window, might be a request for them to close it. We rely on non-linguistic features of the situations, such as context and body language, as well as established discursive and social conventions, to govern how to properly decipher and take up speech acts.

Here, I want to explore two speech acts that have received little direct philosophical attention: reports and accusations. Accusations seek to hold a person accountable for some sort of wrongdoing, while reports are truth-claims about which the speaker has direct first-personal knowledge. Each has a distinct performative structure even though their surface grammar can at times be indistinguishable.

Reports help make up the fabric of our daily lives. "I woke up and made coffee" is a familiar report about the start of the speaker's day. Reports are truth-claims about something the speaker has particular standing to talk about, i.e., the things the speaker has knowledge of via direct first-personal experience. Presumably, anyone in a similar position would assent to the report, and so the truth-claims offered are agent-neutral in their outputs. The difference between reports and declaratives lies in the relation between the speaker and the subject. Reports require a first-personal relation between the speaker and the topic on which they are speaking. For example, if a person reads a book about sailing, they may give a report on *that book*, but they cannot report on sailing itself—they must *go sailing* to be able to report back on it. Thus, reports are agent-relative in their inputs, but agent-neutral in their outputs. These utterances have propositional content, but only those with first-personal access to that content have the standing to issue the report. However, anyone overhearing the report may incorporate that propositional content into their knowledge of the world. Having heard a report, an audience member⁷ may go on to issue declaratives about its content.

Reports instantiate a particular second-personal relationship between the speaker and the audience. As Richard Moran has argued, in telling, a speaker offers the one told their

word, and vouchsafes the truth of the content of the telling.⁸ Reports take this a step further, assuring the audience that the speaker has direct, first-personal experience of what they are disclosing. That is, reports are an invitation to trust, both in terms of trusting that the content of the account is true and in terms of trusting *in* the speaker themselves in terms of their first-personal experience of that propositional content. This raises the stakes of a report, and disbelief in the report is subsequently a stronger repudiation than it is for tellings generally. A person may fail as a knower by, say, being too gullible and telling others unwarranted and false information, but failure as a reporter means that person failed to aptly perceive their own lived experiences.

Accusations, on the other hand, aim to hold someone to account for their wrongdoing. The wrongdoing can range from the trivial, “You took my pen!” to the serious, “She sexually assaulted me.” While accusations can be leveled by the person wronged or by another, and can be directed to the accused directly or offered to a third party, at their core, they seek to hold the accused accountable for what they have done.

Accusations are about wrongdoing. This is definitional: a statement simply *isn't* an accusation unless it is about norm violation in some way. While just which norms are the focus of the accusation may vary depending on context, accusations are a way to call attention to those norms and the accused's failure to properly follow them. However, accusations do more than point out wrongdoing; they are a call to hold the accused accountable for their transgression. Just what form this holding to account takes may vary, but it requires a material, enacted response that directly engages with the wrongdoer. It must go beyond merely eliciting reactive attitudes from the audience; thinking “wow, he's an asshole” in response to a news story on a politician's deep seated corruption simply isn't holding that politician to account in any meaningful way; if reactive attitudes are all that happen, the accusation has not been properly taken up. Instead, there must be some sort of genuine material engagement with the wrongdoer in which they are sanctioned for what they have done.

This is why accusations paradigmatically must identify a wrongdoer. (Note that “I was robbed, but I don't know who did it,” isn't an accusation. It's a statement *about* a crime, but it doesn't yet accuse any specific person or persons of committing that crime.) Generally, this means naming the person who broke the norm or committed the crime, though sometimes it might mean naming a company or group. Accusations require some amount of corroborating evidence in order to proceed, but the accusation itself initiates this process of holding the wrongdoer accountable to the norm they have broken.

Accusations and reports are often indecipherable at the level of surface grammar. “That person raped me,” could fall into either category. This makes it easy for intended reports to be taken up and treated as accusations, and vice versa. Yet, accusations and reports differ, both in terms of what sort of normative relation is initiated and instantiated by the speech act, how the audience ought to respond, and what disregard for the speech act signifies.

II. THE SPEECH ACTS OF #METOO

It may be tempting to view #MeToo as a movement comprised of accusations. After all, the social media posts of #MeToo identified sexual violations the speaker was subject to and the movement overall is about identifying and combating sexual violence. It seems reasonable that holding perpetrators to account for the sexual violations they enacted would be at the core of the movement. If #MeToo is centrally about accusations of sexual violation, then the proper uptake of these speech acts is to initiate the process of holding perpetrators to account for their actions. Since accusations do require corroborating evidence, this might mean opening investigations either in the workplace, through the law, or via some other mechanism. How many perpetrators have been fired, prosecuted, or otherwise held materially accountable for the sexual violations they enacted would thus be a key metric of whether #MeToo succeeded in its aim. But we know that relative to the number of stories of sexual violation shared, there have been few official sanctions against perpetrators. If #MeToo is a movement centered on accusations, then it's pretty clear that the movement failed to achieve the ends of its speech acts.

Certainly, at least some of the utterances from the movement fall into the category of accusations, most clearly those that named the perpetrator and called on others to do something about it. But most do not fit the structure of an accusation. The paradigmatic post “#metoo” doesn't identify a perpetrator or give any of the details needed to hold someone to account in the material way that accusations call for. The only way to take this as an accusation is to treat it as a particularly incompetent and ill-formed one—one which never had a chance of succeeding in doing the work of an accusation. Treating the speech acts of #MeToo as accusations serves as a way of shifting the goal posts of the movement so that it was structurally unable to succeed. Moreover, it does so in such a way that puts the responsibility of that failure on the survivors who spoke out—it treats the failure to sanction assailants as the survivors' fault, since they were the ones to issue such poorly structured accusations. This reading of the speech acts of #MeToo demeans the agency and competency of the survivors who spoke out with these posts. Rather than assuming that survivors utterly failed at the basic mechanics of accusations, we ought to recognize that survivors were doing something else with their speech.

#MeToo is not a movement of accusations, but rather of reports. #MeToo was centrally and powerfully about survivors issuing truth-claims about first-personal experiences of sexual violation. Some survivors shared detailed stories, while some merely wrote “#metoo.” This statement, which made up the core of the movement, is a report in its most pure form: it's a first-personal truth-claim about the world, stripped down of all other description or information. Survivors asserted their own experiences. Note that the movement didn't center on declaratives: Facts and statistics about sexual violence may have been shared to give context to survivors' speech, but they weren't the heart of the movement, nor the purpose of the claims. #MeToo did something more than circulate information about sexual violence: it was about sharing first-personal

experiences of sexual violation. It was centrally, specifically, and importantly about reports.

Understanding what sort of speech acts #MeToo was comprised of shifts the metric by which we ought to evaluate the movement. Rather than focusing primarily on prosecutions or other forms of sanctions, we ought to look at what it means for these reports to receive the appropriate uptake. Taking up a report means taking the speaker to have the proper first-personal relation to the topic they are discussing. By issuing reports, speakers called for trust from their audiences, both in the truth of their testimony and also trust that survivors *are able* to speak to their own life experiences. Recognizing the reports of #MeToo means placing survivors as active epistemic and discursive agents within their own narratives.

III. UNSILENCED SPEECH

Feminist philosophers have long noted the ways survivors' speech has been silenced and dismissed. One way this happens is when survivors' testimony about the violations they've endured is rendered *unspeakable*.⁹ Survivors might say the words, recounting their violation, but audiences don't give the uptake needed for that speech to have any impact. That is, audiences don't take survivors to be saying anything at all. In order to successfully perform a speech act, speakers must receive uptake from their audiences. Audiences draw on linguistic conventions, social norms, and contextual features of the interaction to properly take up speech acts. Broad social conventions render survivors' speech unintelligible as any kind of speech act. Note that this is a common response to men's testimony about being subject to sexual violation: awkward silence, perhaps some laughter, and then changing the subject. Lacking a framework to govern how to interpret this speech, the audience doesn't take it as anything at all.

Of course, survivors' speech isn't always wholly silenced. But when survivors are taken as doing *something* with their speech, they often still labor under the dual burdens of epistemic and discursive injustice. Epistemic injustice, in which survivors aren't taken as credible *knowers*, leads to dismissing survivors as lying or mistaken (see responses to Christine Blasey Ford's Congressional testimony for a striking instance of this).¹⁰ Again, social conventions structure audiences to disregard survivors' accounts of their sexual violation, this time on the basis of survivors' status as either untrustworthy or incompetent epistemic agents. Either way, audiences take themselves to be justified in dismissing the survivor's account since it's not rooted in true facts about the world.

Other times, survivors aren't dismissed due to presumed lack of credibility, but rather are taken as issuing a wholly different sort of speech act altogether: instead of making assertions about the world, survivors are taken as issuing *expressives*.¹¹ These wholly personal speech acts express the speaker's emotional state, but don't call on the audience to recognize any propositional content about features of the world. When a speaker says, "My boss is sexually harassing me," the response they receive is, "I'm so sorry you're feeling that way." To the degree the audience member feels called on to do anything further in response to this,

they might follow up with, "What can I do to make you feel better?" But expressives don't make truth-claims about the world, and so the audience isn't called on to respond with anything more substantive. This is a form of what Kukla terms "discursive injustice," where a disempowered speaker is unable to do with their speech as they properly ought to be able to do, and instead their speech is taken up in a way that reinforces their disempowerment.¹² Treating survivors as hyperemotional and denying them the possibility of factual statements about the world further solidifies survivors' lack of social power.

The #MeToo movement signifies a substantial shift in the dominant interpretation and uptake given to survivors' speech. Quite clearly, survivors were taken as doing *something* when they spoke; the speech didn't simply disappear into a void. The more than 19 million Tweets in just one year, the thinkpieces, rallies, conferences, and continued attention to #MeToo all illustrate that at least some sort of uptake was secured. Moreover, survivors weren't merely taken as venting about their internal emotional states; the participants' speech wasn't simply dismissed as hysterical or angry. Rather, they were generally taken as issuing a statement that had some sort of propositional content. Neither were the participants all simply dismissed as lying or confused or misguided. While, of course, many people did ignore or disregard the movement, #MeToo seems to mark a shift in the dominant frameworks that governs how we take up survivors' speech. Overwhelmingly, survivors were recognized as speaking to their own experiences. Of course, all the traditional forms of dismissal are still occurring; the traditional forms of silencing have not been wholly supplanted by this new interpretive framework. But #MeToo heralds an important addition to our social norms, where the statement "me too" is now readily legible as a report on the speaker's lived experiences. Within #MeToo, we saw survivors' speech become speakable in a way it traditionally hasn't been.

Even the resistance to #MeToo illustrates this shift. The typical pushback to the movement has centered on asking if #MeToo "has gone too far." The thought here tends to be that innocuous interactions are being interpreted as sexual harassment or assault; we've gone "too far" in what we label as sexual violations. But the crucial point to note is that even the detractors recognize the survivors' speech as the action it is. They are giving uptake to survivors' speech as a having propositional content about the survivor's own experiences; the dismissal comes at the level of how survivors label those experiences. While survivors may still be dismissed, they're no longer being entirely silenced.

CONCLUSION

If we misconstrue the movement as being comprised primarily of accusations, this puts the focus on whether or not sanctions have occurred. With this focus it's easy to think the movement was ineffectual. But the survivors of #MeToo were doing something different. They were exerting ownership over their experiences and were calling on others to recognize both their authority over and the concrete reality of those experiences. Once we recognize what sort of action the participants in #MeToo were doing, we can see the substantive ways in which they succeeded.

We are still working out how to respond to the reports issued in #MeToo. There needs to be large social change so that sexual violations are no longer supported and enabled by the institutions in which many of us are immersed. We need to find better ways to hold perpetrators accountable for the sexual violations they enact, and we need to do so far more consistently. But in order for those changes to occur, we first needed people to recognize a *need* for them to occur. #MeToo heralded a shift in dominant interpretations of survivors' speech, where survivors were able to assert authority over lived experiences of sexual violation. This opened up the possibility for audiences to recognize how prevalent these experiences are and the need to make substantive changes to combat these violations. #MeToo marked a change in the possibilities available for what survivors can do with their speech. Though there's still much work to be done, this is an important success that we ought to celebrate.

NOTES

1. CBS News 2017.
2. Anderson and Toor, "How Social Media Users Have Discussed Sexual Harassment Since #MeToo Went Viral."
3. I'm going to follow standard practice and use "survivor" to refer to a person who has been subject to sexual violation. I recognize, though, that not everyone with such experiences self-identifies this way—sometimes because they do not consider their experience to be egregious enough to warrant the term "survivor" and sometimes because "survivor" seems to imply that the experience is in the past, while their trauma moves with them through life. I also recognize that this is, to an extent, begging the question—using this term presupposes that the speaker really has been sexually violated. This last part is deliberate; given the widespread dismissal of experiences of sexual violation here I opt to err on the side of belief.
4. Kukla and Lance, *Yo! and Lo! The Pragmatic Topography of the Space of Reasons*.
5. *Ibid.*
6. Derrida, "Structure, Sign, and Play in the Discourse of the Human Sciences."
7. "Audience member" here is again maximally broad, and refers to anyone who hears, reads, sees, or otherwise takes up the speech act.
8. Moran, "Getting Told and Being Believed."
9. Langton, "Speech Acts and Unspeakable Acts."
10. Dotson, "Tracking Epistemic Violence, Tracking Practices of Silencing"; and Fricker, *Epistemic Injustice: Power and the Ethics of Knowing*.
11. Kukla, "Performative Force, Convention, and Discursive Injustice."
12. *Ibid.*

BIBLIOGRAPHY

Anderson, Monica, and Skye Toor. "How Social Media Users Have Discussed Sexual Harassment Since #MeToo Went Viral." *Pew Research Center*. October 11, 2018. Accessed June 2019. <https://www.pewresearch.org/fact-tank/2018/10/11/how-social-media-users-have-discussed-sexual-harassment-since-metoo-went-viral/>.

CBS News. "More Than 12M 'Me Too' Facebook Posts, Comments, Reactions in 24 Hours." *CBS News*. October 17, 2017. Accessed June 2019. <https://www.cbsnews.com/news/metoo-more-than-12-million-facebook-posts-comments-reactions-24-hours/>.

Derrida, Jacques. "Structure, Sign, and Play in the Discourse of the Human Sciences." In *Writing and Difference*, translated by Alan Bass, 232–50. Chicago: University of Chicago Press, 1966/1978.

Dotson, Kristie. "Tracking Epistemic Violence, Tracking Practices of Silencing." *Hypatia* 26, no. 2 (2011): 236–57.

Fricker, Miranda. *Epistemic Injustice: Power & the Ethics of Knowing*. New York, NY: Oxford University Press, 2007.

Kukla, Rebecca. "Performative Force, Convention, and Discursive Injustice." *Hypatia* 29, no. 2 (2012): 440–57.

Kukla, Rebecca, and Mark Lance. *Yo! and Lo! The Pragmatic Topography of the Space of Reasons*. Cambridge, MA: Harvard University Press, 2009.

Langton, Rae. "Speech Acts and Unspeakable Acts." *Philosophical Affairs* 22, no. 4 (1993): 293–330.

Moran, Richard. "Getting Told and Being Believed." *Philosophers' Imprint* 5, no. 5 (2005): 1–29.

#MeToo?

Lori Watson
UNIVERSITY OF SAN DIEGO

INTRODUCTION

The #MeToo movement has moved us closer to what no law or prior social movement has yet to do: it has created significant social acknowledgment and *belief* of the pervasiveness of sexual harassment and sexual violence. Prior interventions, notably sexual harassment law and accompanying consciousness-raising, created the context for survivors to speak, defined their injury as one of inequality, and paved the way for the #MeToo movement to emerge as powerfully as it has. Against that backdrop, #MeToo has become a forceful social movement shifting the burden of proof away from survivors and placing it on perpetrators, where it always belonged. This shift from a kind of a priori suspicion of survivors' testimony to a predisposition to believe them marks a moment with incredible potential for transformation of the culture of disbelief and degradation of survivors.

Yet, as is so common with social movements aiming to increase the standing of subordinated groups, there is an undercurrent of a backlash. #Himtoo and other forms of victim blaming and denial are ever present in conservative circles and social media. And the fact that Brett Kavanaugh was confirmed to the Supreme Court of the United States amidst multiple accusations of sexual violence and harassment and in the face of the very powerful testimony of Christine Blasey Ford was a stark reminder that male power will consolidate itself when real power is at stake. This backlash, however, hasn't fully undermined and overtaken the power of #MeToo, and that suggests that the cultural shift has legs that may well underwrite the social transformation feminists have been working toward for some time.

Still, there is work to do to sustain this momentum and broaden its reach. Intersectional inequalities, including those on the basis of race, class, sexuality, disability status, as well as gender expression and identity mean that the power of #MeToo is unequally distributed. The aforementioned credibility boost attaches more or less to survivors' testimony depending upon their social location. The more socially subordinated, the more inequality on the basis of group membership, the less credibility or support is extended. The potential power of #MeToo depends on equality within it; hierarchies of standing that underwrite

social inequality such that only some women finally get their moment of justice, whether legal or social, must be dismantled for the reality of the #MeToo moment to unfold. Articulating and calling attention to the experiences of the most socially marginalized, across all forms of intersecting inequalities, is essential for the power of #MeToo to bring about the form of social change to which its visionaries and the rest of us aspire.

Arguably, the primary consequence of #MeToo has been to expose both the commonality and pervasiveness of heterosexual sexual violence and harassment. #MeToo has brought the analysis of heterosexual sexuality as structured by relations of domination and subordination, power and powerlessness, constituted by male power to coerce, extract, and demand sexual access to women as a social entitlement, into public social consciousness in a broader fashion than previous social movements were able to do. Same-sex sexual harassment and violence has played a significantly lesser role in the #MeToo movement. To the extent that same-sex sexual violence and harassment have received attention, or survivors have spoken out, it has largely been men revealing such abuse by other men. Same-sex sexual harassment or sexual violence between women has not been a part of the #MeToo movement. It is fair to say that, as far as #MeToo goes, woman-on-woman sexual violence and harassment has not yet found expression.

There may be important social-psychological reasons for this lacuna: First, as a sheer numbers game, there are likely far fewer of such instances. But they are not zero. And yet, same-sex sexual harassment and violence between women has not found a home in the #MeToo movement. Second, as is well known, members of socially subordinated and marginalized groups are often reluctant to call out or report on other members of such groups; they often fear, rightly, that reports of “bad behavior” on some members of the group will be used to taint and smear the entire group. The anticipation of social cost of such a backlash often prevents such reporting. Third, those subject to it may not readily conceptualize woman-on-woman sexual harassment or sexual violence as “like” male forms of sexual harassment and violence. Fourth, fear and shame of public ridicule, harassment, or worse often keep the most socially unequal persons from speaking their truth.

The aim of my comments is to shed some light on same-sex sexual harassment between women. In particular, I want to explore the gendered dynamics that shape and define such harassment, revealing the layered power dynamics at work. The fact that the harassment or sexual violence may occur between same-sex persons doesn’t thereby make it any less gendered. Of course, there is a wide variety of gender expression within the lesbian, bi, trans, queer, and questioning community. And one need not be self-identified as a member of our community to sexually harass a woman; straight women, too, harass women or gender queer persons in sexual ways. Perhaps contrary to common assumptions, same-sex sexual harassment does, in fact, have gendered structure. Social scripts of masculinity and femininity mediate and define the relational status of the women, though perhaps with some less rigidity than

opposite-sex cases of sexual harassment. The structure of gender might be most clear in cases in which the harasser is “the masculine one” defined in relation to “the feminine one.” But even in cases when the sexual harasser is marked “the feminine one” and the harassed “the masculine-yet-woman,” in my experience and observation, gendered power dynamics mediate the interaction. The masculine-identified woman occupies a space of sometimes being conferred forms of male power, but always a woman, and this unique gender constellation can reveal how gender power tracks and reinforces social inequality on the basis of sex.

BECAUSE OF SEX

Sexual harassment is “[g]enerally understood as unwelcome verbal, visual, or sexual conduct, typically in a context of inequality.”¹ Though, “[h]arassment does not have to be of a sexual nature, however, can include offensive remarks about a person’s sex. For example, it is illegal to harass a woman by making offensive comments about women in general.”² Key to the legal concept of sexual harassment as a form of sex discrimination is that the harassment is “because of the sex” of the harassed. In other words, but for the sex of the victim, the harassment would not have occurred. And yet put another way, sexual harassment happens, has the structure it does, “because of the victim’s gender.”³ However, precisely how to understand and define “because of sex” has proven less than straightforward in legal cases, and its doctrinal evolution took some time. It wasn’t until 1998 that the Supreme Court of the United States recognized that same-sex sexual harassment was “because of sex” and thus a prohibited form of sex discrimination under the auspices of Title VII. In *Oncale v. Sundowner Offshore Services, Inc.*, the Court held that same-sex sexual harassment is a violation of sex equality rights. Leading up to this case and subsequent to its holding, some have argued that same-sex sexual harassment should be understood as sexual orientation discrimination rather than sex discrimination, despite the fact that there is no federal prohibition against same-sex discrimination. Perhaps this suggestion is grounded in thinking through the various constellations of man-on-man sexual harassment, the underlying thought being that given male privilege, sexual harassment directed at men by other men, even if sexual in content or on the basis of sex-stereotyping, isn’t sex discrimination, it must be something else. This thought is mistaken. Perhaps, too, a similar thought underlies misrecognition of the gendered forms of woman-on-woman sexual harassment. What *Oncale* embraced, as legal theorists, most notably Catharine A. MacKinnon, had been advancing for some time, is that under conditions of inequality, sex and gender are socially fused such that gender power dynamics are not determined by biology, but rather, the social structure of gender, and its dominant form, heterosexuality.⁴

WOMAN-ON-WOMAN SEXUAL HARASSMENT

As I aim to shed light on the gendered structure of woman-on-woman sexual harassment both as a phenomenon that exists and as a gendered, sex-based phenomenon, two questions guide my analysis here: Where is the power? And how is it gendered? Putting these questions together, my comments aim to reveal the forms of gendered power that

are often operative in woman-on-woman sexual harassment, revealing the ways in which same-sex sexual harassment is shot through with gendered forms of inequality. My examples will largely draw on harassment aimed at masculine-identified women. This is perhaps the less intuitive case, though for that reason I think it is particularly illuminating. Also, it is my experience. In addition, though, I think that this particular gendered position is revealing insofar as being socially read as masculinely gendered may serve to obfuscate the power dynamics at work when people like me are subjected to unwanted sexual advances or attention.

First, such harassment often rests on an implicit assumption of sexual availability, with all the gendered assumptions that attach to femininity and masculinity. Feminine-presenting women harassed by other women, whether the other women identify as feminine or masculine or neither, are sexualized in their femininity. As such, the harasser often positions herself vis-à-vis the harassed as seeking to gain access in ways that demand polite acceptance, demur passivity, or engagement that affirms the sexual value, attraction, or power of the harasser. The gendered power dynamics are shaped by the social scripts of masculinity and femininity irrespective of the same-sexness of the parties. The harassed qua subordinate is expected to receive unwelcomed advances with aplomb, if not immediate acceptance. The positional power of the harasser situates them as entitled to demand, force sexual interactions upon, or otherwise press unwanted sexual overtures upon their prey. Where the harassed is clearly normatively feminine⁵ and the harasser is read as more masculine presenting, this dynamic may be most clear. However, the same power dynamics are at play for “masculine”-identified women, where the harassment they receive is often from “feminine”-identified women.

The projection of masculinity onto women who present more masculinely often comes with the assumption that they are ever up for a sexual overture or encounter. Insofar as the social interpretation of masculine sexuality is projected onto us it carries with it the belief that masculine persons don’t turn down sexual offers from anyone, ever. However, there is an interesting twist on the way in which the assumption of sexual availability functions in the context of the masculine-identified woman. She is still a woman, on some level, or in some way. And so, the assumption that she play a caregiving role, attending to the needs and vulnerability of the woman making the sexual advance, is also often operative in these contexts. In my experience, the way this can play out is that the (feminine) woman making an unwanted sexual advance projects the sexual nature of the interaction upon the masculine-identified woman. So, for example, she may sexualize the masculinely identified woman and then follow up such a remark in ways that shift the responsibility of the sexual content of the interaction to the masculinely identified woman, such as making a sexual comment and following it up with “you are such a flirt.” The interaction is shot through with gendered scripts, and the power dynamics are interestingly marked by the combination of masculine-yet-woman sex status of the harassed. The projection of masculinity and accompanying norms onto women who are read as socially masculine,

or self-identify as such, does not serve to undermine the requirement of a complicit femininity in receiving the unwelcome overtures of other women. Thus, the position of the advancer, harasser, often suggests that the sexual invitation is a form of flattery and assumes, as such, it is or ought to be welcomed. Refusal or failing to give uptake to the harassment often comes with the explicit or implicit judgment that you are a bitch, or a tease, or uptight, or other similar responsibility-shifting judgments.

The point is that despite the same-sex status of the harasser and the harassed, and even the socially read or imputed gender presentation of either party, the power dynamics are such that social femininity, even if not conjoined with physical presentation of outward femininity, is imposed upon the harassed. Accept the overture, comply, be polite, affirm the harasser’s sexual status or be a bitch, a tease, or someone who can’t take a joke.

A second prominent dynamic in sexual harassment, both opposite-sexed and same-sexed harassment, is the eroticization of “otherness”: many members of intersecting subordinated groups report the ways in which their “difference”—inequality, in fact—is sexualized and tokenized as “exotic.” One form of such “exoticism” within lesbian, bi, and other non-normative sexualities is “female masculinity” as it has been called,⁶ or “butchness,” relying on the older term, or gender non-conformity such that masculinity and femininity are combined or shunned, say, through androgyny. Such “exotic otherness” is often fetishized among some women who are sexually interested in women with masculine or other forms of gender-non conforming appearance. Sometimes this is marked by questioning and curious women feeling free to say things like “if I were ever going to be with a woman, I would be attracted to or want to be with a woman like you.” This kind of unwelcomed advance also reflects the false belief that gays and lesbians are sexually available to any interested same-sex other—the false belief that gays and lesbians don’t have discriminating tastes or preferences, but are just willing to have sex with any other available interested party. Sometimes the projection of “exoticness” comes from other lesbians or bi women who want to announce or otherwise make clear their interest in “women like you” irrespective of any indication on your part as to whether you are interested in them or welcome such suggestions. The insertion of sexual comments or overtures targeting you, and quite independent of your wants, desires, and needs, or expressions of them, is one of the primary ways in which male sexuality is forced upon women. That women can take this position vis-à-vis other women doesn’t make it any less gendered. Sexuality is structured in the first instance by normative heterosexuality and male dominance, and the eroticization of “otherness”—social subordinates—is one manifestation of such power. This dynamic occurs across and within same-sex sexual harassment and the same power structure underlies it: eroticization of inequality as difference.

A third feature of the way in which gendered power relations are expressed in sexual harassment is the way in which persons who are harassed are often reduced to their sexuality. This experience is often characterized,

in part, by those harassed as a reduction of self to one's sexuality, such that one is seen and treated merely as a sexual object for the use and enjoyment of the harasser. This experience is common to women, even as it takes on particularized forms across and intersecting with other forms of inequality, such as race, class, nationality, and so on. Gays and lesbians, too, live in a world that frequently reduce them to their sexuality; they are seen and treated as their sexuality. This reduction of self to sexuality is no less prominent in same-sex sexual harassment, or when the harasser shares a sexual identity with the harassed, for the structure of such harassment involves unwanted sexual attention or demands on the harasser's terms. The power of the harasser to define, construct, and impose their vision of the sexual status and interests of the harassed is constitutive of sexual harassment per se. The specificity of such power dynamics in the case of same-sex sexual harassment in particular includes not just the reduction to sex-object but to a marginalized sexuality in a context in which much of the lived experience of such persons includes such reduction.

A further feature of sexual harassment, well documented in discussions of the phenomena and especially prominent in outing the behavior of harassers in the #MeToo movement, are the ways in which harassers count on the subordinated status of the harassed to protect them from accusations or accountability. To the extent that the woman experiencing harassment is also socially situated vis-à-vis the harasser through other layers of inequality, race, class, non-normative gender expression, her membership in these groups with intersecting forms of inequality serves to undermine her credibility and prop up her harasser. Her "difference"—i.e., unequal social position—works against her in positioning her as a target and, likely, in any subsequent attempt to call out or report her harasser. #MeToo has had some efficacy in shifting the responsibility back to where it belongs, upon perpetrators. Whether and how effective that shift may be for women experiencing same-sex sexual harassment is unknown. Gays and lesbians are at heightened risk for blame shifting, especially in the context of dominant ideologies that pathologize gay and lesbian sexuality.

A further harm of same-sex sexual harassment is the way in which it functions to deprive the harassed of something like "insider status," at least as it concerns gender group membership. What I have in mind here is same-sex sexual harassment, for example, between women, where the harassed is seeking community or connection with the harasser simply as a woman, or as a lesbian, or member of the broader LGBTQIA community, and that connection is thwarted because of the sexual harassment. The experience of harassment as sexualized object positions the harassed as outsider, in some sense, to the community or connection. She is not and cannot be in that context "just one of the girls." The deprivation of inclusion within a community from which one seeks refuge is an especially harmful form of further subordination.

Though the examples and categories of analysis offered here in no way aim to be exhaustive of the phenomena and experience of sexual harassment, whether understood broadly or in the more specific case of same-sex sexual

harassment, they identify some of the key gendered dynamics of such harassment. Whether the harasser is the same-sex as the harassed, the power dynamics, the imposition of unwelcomed advances or comments serve to construct the power position of the harasser while enacting the subordination of the harassed.

CONCLUSION

Part of my motivation for writing this paper came from a recent discussion with a similarly situated gender non-conforming friend; although I identify as a woman and lesbian and they do not, our social presentation of gender is very similar. I am often read socially as a man, and "pass" regularly even though I do not explicitly aim to pass as a man; I do, however, accept whatever gender assignment any stranger projects on to me without correcting them, partly for safety, partly out of exhaustion with the whole damn gender thing. In any case, the friend and I were discussing the ways in which we are routinely sexualized by women, sometimes sexually harassed, and as we were discussing it, we were sharing the ways in which our gender presentation is read in ways that both seem to serve to "authorize" the harassment and to structure it. And in that sharing, I came to fully think through and own all the various moments in which I have felt unwanted sexual attention, as a butch lesbian, which is how I most identify, and how unable I was to sometimes name it and other times to do anything effective about it. I engaged in all the forms of self-blame, self-shame, catering to the harasser, excusing her, working to make her feel more comfortable, when I avoided or declined the unwelcomed advances, and many other forms of, well, self-denial. All this from someone who identifies as feminist, has read the relevant literature, and would support and assist any other woman in calling out or otherwise recognizing sexual harassment for what it is. Thus, what my conversation with my friend gave me was my own #MeToo moment, and central to its possibility, for me, was the shared experiences of someone like me in the salient respects. I haven't known many "like me." Thus, my humble hope is to have shed some light on the phenomenon of same-sex sexual harassment in ways that resonate for women, or others, like me, as it were, such that #MeToo includes us.

ACKNOWLEDGMENTS

I would like to thank Christie Hartley, Robin Dembroff, and Sarah Clark Miller for their comments and discussion of this paper.

NOTES

1. Catharine A. MacKinnon, *Sex Equality*, 3rd ed. (Minneapolis, MN: Foundation Press, 2015), 1002.
2. Equal Employment Opportunity Commission, "Sexual Harassment" Definition, available at https://www.eeoc.gov/laws/types/sexual_harassment.cfm.
3. Catharine A. MacKinnon, "Afterword," *Directions in Sexual Harassment Law*, eds. Catharine A. MacKinnon and Reva B. Siegal (New Haven, CT: Yale University Press, 2004), 676.
4. See, for example, Catharine A. MacKinnon, *Toward A Feminist Theory of the State* (Cambridge, MA: Harvard University Press, 1989), where she writes, "In the concomitant sexual paradigm, the ruling norms of sexual attraction and expression are fused with gender identity formation and affirmation, such that sexuality equals heterosexuality equal the sexuality of (male) dominance and (female) submission" (131).

5. I use “normatively feminine” here rather than the more commonly accepted “cis” to denote persons for whom their gender expression aligns with the social interpretation of their bodies. For reasons that are too long and complex to enumerate here, I am not a fan of “cis” as it now stands in for persons who don’t reject, or accept in some sense, their gender assignment at birth as a contrast to trans or gender non-conforming persons. While I recognize that persons who live in the world in ways that are more comfortable, in some sense, with their assigned at birth gender can do enjoy forms of privilege that others lack, I think the reductive dichotomy of “cis” vs. “trans” is itself not especially helpful for diagnosing that privilege.
6. See, for example, Judith Halberstam, *Female Masculinity* (Durham, NC: Duke University Press, 1998).

#MeToo vs. Mea Culpa: On the Risks of Public Apologies

Alice MacLachlan
YORK UNIVERSITY

1. INTRODUCTION

One need hardly be an expert in public apologies to conclude that the 2017 #MeToo resurgence on social media produced some spectacularly bad ones.¹ Harvey Weinstein, the movie mogul at the center of the celebrity turn in #MeToo, began his apology by blaming his upbringing in the 1960s for his actions in the 1990s and beyond. He ended by ranting about the NRA and Donald Trump.² After actor Anthony Rapp accused fellow actor Kevin Spacey of assault when Rapp was a minor, Spacey chose to offer a hypothetical apology—“If I did behave as he describes”—only to then seize the moment and throw himself a coming out party, publicly identifying as a gay man for the first time.³ And after four women accused him of sexual misconduct, chef Mario Batali issued a letter apologizing to his friends, family, fans, and team—but not to his victims—and then threw in a recipe for pizza dough cinnamon rolls as a postscript.⁴

More predictably, there was a spate of the usual non-apologies and quasi-apologies, that is to say, performances that might *appear* to say sorry without ever quite admitting wrongdoing, taking responsibility, or acknowledging the impact of the speaker’s actions on their victims. Actor Dustin Hoffman offered regrets for “anything I might have done” while disputing all the things he was actually accused of doing;⁵ fellow actor James Franco denied the accusations but insisted he nevertheless supported his accusers’ newfound voice.⁶ YA author Sherman Alexie denied all the specifics while going on to admit, almost cryptically, “there are women telling the truth about my behaviour” in his apology, and so on, and so on.⁷

Bad apologies can be frustrating, infuriating, and profoundly painful—in some ways, they are more hurtful than outright denials of wrongdoing. Certainly, the ambivalent, obfuscating public statements described above deserve our critique and condemnation. But what if they had been better? That is, what if the #MeToo movement had produced, along with an upsurge of truth-telling and solidarity among survivors and allies, an equivalent rise in genuinely sincere

statements of responsibility and remorse by perpetrators and collaborators? I want to suggest that, while perhaps preferable to what we did get, this would not have been a straightforwardly happy ending. That is, I believe the risks of public apologies are not limited to the faults and flaws of bad apologies; in many ways, good ones are more insidious. As a result, my focus in this essay is not primarily how we might assess and compare individual apologies, or even what makes for a good #MeToo apology in general. Instead, I identify three risks to both good and bad public apologies and demonstrate how these risks intensify when it comes to the gender politics of #MeToo, arguing that the aims of #MeToo may be in tension with and even be undermined by the practice of public *mea culpae*.

2. SORRY STORIES

The #MeToo apologies I am discussing are *public* apologies.⁸ More specifically, they are *public-personal* apologies, rather than official or institutional apologies: public figures addressing their own individual actions. There now exists a respectable philosophical literature analyzing the nature and value of both public and private apologies and debating the standards by which we ought to evaluate them.⁹ Since my aim here is not to rank individual #MeToo apologies as apologies, but to note overarching trends in the *practice* of apologizing, I do not dedicate much space to the topic here. Instead, a broad or inclusive approach to what counts as an apology will serve useful for identifying the risks associated with characteristic instances of the practice.

An apology takes place in the aftermath of wrongful harm. One person (the apologizer) offers it to another (the recipient), who can accept or refuse it. Apologies express or imply that the actions in question took place and were wrongful and harmful, that the recipient was wrongfully harmed by them,¹⁰ and that the speaker takes responsibility for their actions *and* censures or disavows them. Many apologies also express a commitment to some remedy, reform, or repair.¹¹ Finally, apologies typically express some appropriate affect, e.g., sorrow or shame—an emotional expression of the wrongdoer’s self-evaluation and current moral commitments.

The value of a given apology will usually depend on whether the speaker is able to persuade her audience of the truth of these claims and—in particular—the depth and sincerity of her responsibility and disavowal, her appropriate understanding of how the victim has been hurt, as well as the reliability of her commitment to change.¹² Together, these indicate that the speaker is appropriately apologetic, that she has shifted her beliefs, attitudes, dispositions and behaviors in the relevant ways. Other factors may play a role too: the timing of the apology, the events precipitating it (and, indeed, following it, so that our assessment of an apology cannot be made conclusively in the moment), and the conditions under which it is offered. Finally, there is an undeniable aesthetic dimension to a good or even outstanding apology. As Mihaela Mihai puts it, an apology is an *imaginative* act—a recasting and re-imagining of the relationship between act, agent, and recipient—and so “a check-list model of apology fails to capture faithfully the kind of imaginative act that an apology must be.”¹³

Not all #MeToo apologies were bad apologies. None were perfect (I suspect few of us have given or received a perfect apology) but several met some basic threshold for all or most of these elements.¹⁴ But the problem with #MeToo apologies is not simply that they need to be better.

3. RISK #1: APOLOGIZING TO WHOM?

The first risk of apologizing in public is that the primary victims of wrong will be overlooked in the wrongdoer's effort to address and reassure multiple audiences and—in particular—the broader public on whom the institution or public figure relies (e.g., state apologies that aim to reassure the voting public or corporate apologies that aim to establish broader market trust).¹⁵ It is telling that the framing of many #MeToo apologies suggest that it would be *insufficient* merely to apologize to the woman or women in question; she is often displaced altogether in the apologizer's concern for (in Mario Batali's words) "family, friends, fans, and team" or is blurred into the impersonal, easily substitutable "anyone" who might have taken offense. Weinstein would only specify "colleagues." Other apologizers go out of the way to make it clear that while they are apologizing, they offer that apology to anyone *but* their actual accusers (for example, Sherman Alexie). The most important people in the equation are erased.

Even those apologies that do name and acknowledge the victim(s) may dilute that focus in paying equal attention to secondary victims and those affected not by the apologizer's actions but by the events following their public accusations. Take, for example, two #MeToo apologies that actually succeed along many of the axes I described above: those by Senator Al Franken and stand-up comedian Louis CK.

In November 2017 Leanne Tweeden, a news anchor, accused Franken of forcibly groping and kissing her on a USO tour (for which she provided photographic evidence). Franken's eventual public apology began by naming his victim directly, "The first thing I want to do is apologize: to Leeann," but then went on to immediately list "everyone else who was part of that tour, to everyone who has worked for me . . . everyone I represent . . . everyone who counts on me to be an ally and supporter and champion of women."¹⁶

Ten days earlier, Louis CK had issued a public statement, admitting, for the first time, that he would expose himself and masturbate in front of women. While the comedian never quite brought himself to say the words "sorry" or "apologize," he did begin by immediately acknowledging his victims and their truths: "I want to address the stories told to *The New York Times* by five women named Abby, Rebecca, Dana, Julia who felt able to name themselves and one who did not. These stories are true." But this direct acknowledgment was later diluted, first, by his extended meditation on the relationship between power and consent that dwelt a little too long on how universally he had been "admired" and, again, when CK concluded by saying, "I'd be remiss to exclude the hurt that I've brought on people who I work with and have worked with whose professional and personal lives have been impacted by all of this, including projects currently in production." CK went on to name the "hardship and anguish" faced by his manager

(who played a not-so-minor role in pressuring women to cover up the stories) and FX executives.¹⁷

While, on one level, this widened audience is perhaps noble—it recognizes that the unfolding drama is *also* something the perpetrator brought on themselves—it also has the rhetorical effect of treating the public shaming as a second, equally significant harm, even equating the two traumas. Indeed, in the predictable but disappointing backlash to #MeToo, several have suggested that being called out for sexual assault is a harm equivalent or greater than sexual assault itself. Take these remarks by comedian Norm MacDonald: "There are very few people that have gone through what they have, losing everything in a day. Of course, people will go, 'What about the victims?' But you know what? The victims didn't have to go through that."¹⁸ MacDonald is not an outlier; others in the industries most affected by #MeToo have made similar comparisons.

Because #MeToo apologies didn't arise directly following the *harm*, but after a later and highly public moment of accountability—one that probably felt like a bigger deal to the apologizer—they tend to address the latter as much as the former, almost as if the apologizer is being asked to weigh in on #MeToo as a whole, rather than their own actions, shifting their role from target to ally.¹⁹

4. RISK #2: FROM WOMEN'S LIVES TO MEN'S FEELINGS

The second risk in the growing practice of public apologies is their tendency to shift our attention from accuser to accused. Like so many women, I remember quite viscerally when #MeToo rippled across my social media in October 2017. It was invigorating and heartbreaking in equal parts to read familiar iterations of the same story: harassment, pressure, coercion, manipulation, assault, silencing, denial. Sexual violence was filling cyberspace, media spaces, and daily spaces—and not as the plot on a crime-of-the-week TV show or murder podcast, as a depressing set of statistics, or other familiar tropes. Rather, it was being told in the voices of women, genderqueer, non-binary, and other marginalized folk, about the *experiences* of women and marginalized folk, and the very act of telling was taking on a new and collective power from those margins.²⁰

In many ways, the ability to call forth public acknowledgment—and even apology—was the most obvious (and sought after) manifestation of that power. Apologies will always matter insofar as victims and survivors *want* them, and insofar as they represent genuine accountability *to* those victims and survivors. But receiving the apologies of wrongdoers requires that victims pass the mic and the wrongdoer take center stage. Our attention moves from victim's stories to the wrongdoer's own accounting of what happened, why it happened, and how. In one sense, that first act of power—filling up public space with women's experiences of sexual violence, told in women's voices—was substantially diminished the minute perpetrators started talking.

And when perpetrators picked up the mic, they told us of their confusion and amnesia, remorse and shame, guilt and

sadness, who they were and weren't, what they would or would never do, what they resolved to do going forward, the ordeal they had experienced simply by being named, where they stood on the #MeToo movement and—in great detail—of their tremendous respect for women. The vast majority of #MeToo apologies focused on the apologizers themselves: their perceptions (“I don’t remember,” “I misread the situation,” “indications . . . it was consensual,” “I told myself it was ok”), their intentions (“that was not my intention,” “however unintentionally”) their emotions (“the hardest regret to live with,” “I am horrified and bewildered,” “I feel disgust with myself”) and their character (“it’s not reflective of who I am,” “I am not an assaulter,” “I have the utmost respect for women,” “I’m trying to do better”).

The spate of #MeToo apologies turned our attention away from what had happened to victims to what was now happening in the inner—and outer—lives of perpetrators. The apology became a report on the state of the soul of the apologizer: an inner assessment on which, coincidentally, the apologizer is a unique authority. Kate Manne has described her groundbreaking reconception of misogyny as a shift in focus from what men feel to what women face: a sociopolitical rather than a psychological phenomenon.²¹ The narrative shift in #MeToo story from protest storytelling to public apology effectively reversed this trend, drawing our attention away from what was actually happening to women in the world to what particular men did or did not remember, how they did or did not understand their own actions, and even what they felt and thought about the broader #MeToo movement—as when the apologies became opportunities to (re)assert one’s credentials as a feminist ally.

This shift in attention and, thus, in narrative control is worrying for a second reason: namely, it shifts who has authority over the story and how it gets remembered; who has *credibility*. And while it might seem that the person who publicly admits to groping or harassing or assaulting women—or even to being in such a drunken or drug-induced haze that he can neither confirm nor deny these acts—would lack credibility, this isn’t always the case for several reasons.

First, the perpetrators were largely known to the public, familiar and often beloved; for the most part their accusers were less so. The perpetrators were recognizable individuals, public figures, even anti-heroes. Their accusers were a collective and, increasingly, seen as a “mob.”²² Second, the #MeToo perpetrators who publicly apologized were powerful and successful men—mostly straight, mostly white, mostly able-bodied, almost entirely wealthy, and in many cases beloved, and therefore benefited from the privilege afforded to that demographic. They were speakers who possess what Kristie Dotson, Miranda Fricker, and others have identified as a “credibility excess,” a systematic bias in their favor in the economy of credibility.²³ Their accusers were—almost universally—members of subordinated groups who are likely taken as *less competent and less trustworthy*.

Third, the public apology—as genre—actually strengthens rather than weakens the moral and narrative credibility of

the apologizer. As Elizabeth Spelman writes, the apology is a vehicle for “vice nested in virtue”—since the apologizer both owns up to wrongdoing and disavows it—which allows the apologizer to “wrap herself in a glorious mantle of rehabilitation.”²⁴ Someone who is capable of apologizing for wrongdoing can claim (now) to know right from wrong, (now) to understand their responsibilities and their accountability, (now) to have appropriate commitments to make *right* their past wrongs and repair their damage, and (now) to express that commitment in appropriate affect. Someone capable of doing so in public can claim a certain degree of moral courage, too. In other words, the public apology allows the apologizer to position themselves on the side of the angels, to draw attention to their other *good works* (past and future), to perform and even *lead* our shared horror at sexual violence and inequality. Even the non-apology affords the apologizer room to be morally generous, allowing that while they certainly didn’t do *this* particular abhorrent thing, they can recognize their own lesser imperfections and take the opportunity to be better.

For example, Richard Dreyfuss issued a statement that almost, to my mind, ranked among the most thoughtful and reflective #MeToo responses out there, discussing the “complicated truth” of his life as “an asshole—the kind of performative masculine man my father modeled for me to be,” admitting he disrespected himself and others, was swept up in celebrity, that he must redefine what it is to be a man, that what he did—who he was—was not ok, that even now he is “playing catch-up” to what is necessary. He even acknowledged he needs to rethink what he thought was a mutual flirtation. And yet, buried in the middle of this thoughtful exhortation is a flat-out denial: “I emphatically deny ever ‘exposing’ myself to Jessica Teich, whom I have considered a friend for 30 years.” He acknowledges the effects of drugs, being drunk on power, that he is having to rethink every perception and relationship and interaction, that he *doesn’t know anything* except for this one thing: that he is not guilty of the thing he is actually accused of. In effect, Dreyfuss makes it the one thing he is prepared to bet on. His generosity thus becomes the perfect denial, more plausible than if he had insisted on his wider righteousness.²⁵

5. RISK #3: APOLOGY, EXILE, AND REDEMPTION

The narrative arc of the apology performs another redirection: #MeToo apologies turn our attention toward the question of male *redemption*. The offer of apology effectively marks the end of a period of wrongdoing, denial, and demands for accountability, and the beginning of a new arc of redemption—what we might call the first step in a comeback—the first premise of a syllogism that shifts next to forgiveness and concludes with absolution and rebirth.

After the first spate of credible #MeToo apologies, think pieces started emerging musing: “Do the men of #MeToo deserve to be forgiven?”²⁶ “How to find room for forgiveness in the #MeToo movement?”²⁷ “Does forgiveness have a place in the #MeToo movement?”²⁸ These were not media assessments of the forgiveness already offered by victims, but critiques of the victims’ apparent *failure* to do so. In other words, the salient public judgment was no

longer “Have these men been held accountable?” but—subtly different—“Are they candidates for redemption?” What was being assessed was no longer the experiences and the pain of women victims, but the experiences and the pain of male perpetrators: Had they suffered enough?

Again, it is useful here to turn to Kate Manne, who coined the term “himpathy” to describe our “tendency to pardon the hitherto historically dominant, especially when their currently down on their luck,”²⁹ noting especially “the excessive sympathy sometimes shown toward male perpetrators of sexual violence.”³⁰ There are few better illustrations of the distorting effects of himpathy on our assessment of comparative harms than the remarks by Norm MacDonald I referenced above: “There are very few people that have gone through what they have, losing everything in a day. Of course, people will go, ‘What about the victims?’ But you know what? The victims didn’t have to go through that’.”

And yet, only a small minority of #MeToo perpetrators have faced charges, and many of those have had the charges dropped; almost none have faced conviction—despite a preponderance of editorial cartoons that fashioned the hashtag into the barred windows of a cell. Many #MeToo perpetrators have had their careers impacted, but not at the rate women who face harassment have. A 2018 Marketplace-Edison research poll found 46 percent of women who have experienced harassment say it caused them to quit their job or switch careers; many of the #MeToo accusers specifically described the career-impacting and career-destroying effects of their harms, especially since the perpetrators were often their bosses or superiors.³¹ In other words, it is still far worse for one’s career to be assaulted than to assault. Moreover, powerful perpetrators are better cushioned in their fall.

It is remarkable how many perpetrators are prepared to position themselves as either already primed to repent or as capable of authoritatively stating they have already done the relevant moral work. For example, when pressed about accusations of onset harassment and bullying, actor Jeffrey Tambor simply said, “I’ve reckoned with this” and refused to engage further.³² Many #MeToo responses self-describe the apologizer as already having performed the requisite moral education. Newscaster Tom Ashbrook’s *Boston Globe* editorial was titled “Is there room for redemption?” and was written only a few months after he was fired for abusive behavior on the job. It opened with the line “Boston, can we talk?” and explicitly concluded that, at least in his case, there was indeed—or should be—room for redemption.³³ Journalists John Hockenberry and Jian Ghomeshi wrote long, rambling reflections on the uniquely painful experience of “Exile” (the title of Hockenberry’s piece in *Harper’s*) that lamented being banished from the world—where, by “world,” they presumably meant a uniquely privileged position of public influence.³⁴

In other words, the public apology has been taken by perpetrators to be the starting whistle for a limited period of exile, after which the tide will have turned, the past will be buried, and things can return to what they were. Less than a year after their respective public disgraces, both Aziz

Ansari and Louis CK have done stand-up sets in which they reference their #MeToo moments first obliquely and then directly, only to critique the movement for going too far, demanding too much, and talking too long. Implicit in each of these “returns” is the assumption that the apology—once offered—the last word on the matter. There is no attempt to discern and reflect on the fact that, as Lauren McKeon puts it in her *Walrus* article, “the women are not done talking yet.”³⁵ Once the perpetrator has spoken and the apology been issued, such talk is no longer publicly read as a moving uprising but rather, the ranting of an embittered mob.

6. CONCLUSION

My aim in this discussion has not been prescriptive—“stop demanding and offering apologies!”—but cautiously diagnostic. As the practice of public apology becomes further entrenched in public perceptions of #MeToo, we need to become increasingly savvy consumers of public apologies. Doing so requires that we move beyond the question of whether or not a given instance is a “good” one; we need more than “good” apologies.

In the cases described above, I have demonstrated how public apologies, both good-enough and bad, risk displacing the victims of harm, unmooring the question of accountability from its relational underpinning. They also direct our attention away from the suffering of the accuser, and toward the very present, publicly displayed, suffering of the accused. In doing so, they effectively rewrite #MeToo stories, complete with a new protagonist and a new narrative arc.

Indeed, analysis of #MeToo apologies draws our attention to two competing and potentially incompatible narratives of accountability and repair. The first—embraced by the #MeToo movement itself—presents accountability and repair as a question of reckoning, even revolution. It frames the misogyny of endemic assault and violence in Manne’s terms: as a social-political phenomenon that predominantly impacts and affects women and marginalized people. This model is also relational: In telling their stories, #MeToo survivors situate themselves in relation to that broader phenomenon and to other survivors (quite literally, *me too*). The harms in question are repaired when the moral landscape is changed, and the social conditions have shifted; when perpetrators are reliably and systematically held accountable, and victims are reliably and systematically believed and protected.

Most #MeToo apologizers, on the other hand—and those who support, defend, and enable them—draw on a performative and even *purifying* model of accountability. On this model, the repair at issue is a matter of soul-searching and penitence; what needs fixing is interior and achieved through self-reflection and sustained—but not permanent—remorse. Repair takes place when the wrongdoer is (in their own lights) sufficiently sadder and better. So long as #MeToo apologies continue to enact the second model of repair, even the best of them will sit in tension with and even undermine the very goals of the #MeToo movement itself.

NOTES

1. #MeToo is a grassroots movement founded by black feminist activist Tarana Burke in 2006, whose aim is to support survivors of sexual violence—particularly Black women and girls—through community-based activism, with a goal of “build[ing] a community of advocates, driven by survivors, who will be at the forefront of creating solutions to interrupt sexual violence in their communities” (“Me Too. History & Vision” from the website of the Me Too Movement). In many ways #MeToo the movement has been coopted and distorted by #MeToo the cultural moment, when the hashtag went viral in 2017 after a wave of public accusations against media mogul Harvey Weinstein.
2. See “Statement from Harvey Weinstein,” *The New York Times*, October 5, 2017.
3. See “Kevin Spacey Apologizes for Alleged Sex Assault with a Minor,” CNN, October 31, 2017.
4. See “Mario Batali’s Misconduct Apology Came with a Recipe,” *TIME*, December 16, 2017.
5. See “Dustin Hoffman Apologizes After Allegations That He Sexually Harassed a 17-year-old Intern in 1985,” *The Los Angeles Times*, November 1, 2017.
6. “James Franco says accusations of sexual misconduct ‘are not accurate,’” *VOX*, January 10, 2018;
7. “Author Sherman Alexie Apologizes Amid Anonymous Allegations,” *The Guardian*, March 1, 2018.
8. Not all #MeToo apologies are public apologies; some of the more moving accounts of apology in the #MeToo era have been women’s private stories of being contacted by a former abuser or assaulter, and having that abuse or assault acknowledged for the first time. Indeed, non-celebrity #MeToo apologies are the focus of my current work-in-progress on this topic.
9. See, for example, Thompson, “The Apology Paradox”; Smith, *I Was Wrong*; Martin, “Owning Up and Lowering Down”; MacLachlan, “Trust Me, I’m Sorry,”; and Russell, “The Paradox of Apology.”
10. She might not be the only person harmed, in the case of multiple victims.
11. Many excellent apologies will open up this commitment to the recipient, asking rather than deciding what is to be done. This element is not always present; some repeat offenders may not be able to commit sincerely, and yet wish to apologize—in part for this very failure.
12. If the apologizer couches her apology in excuses and self-justifications, for example, and brings up past grievances to settle the score, the recipient may have good reason to doubt her apologetic sincerity.
13. Mihai, “When the State Says ‘Sorry,’” 209.
14. Those interested in an example of a good #MeToo apology are invited to listen to or read comedy writer Dan Harmon’s apology to his fellow comedy writer and former colleague, Megan Ganz. After Ganz called out Harmon on Twitter for past harassment, Harmon dedicated seven minutes of his podcast to addressing, admitting, elaborating, and reflecting on Ganz’s accusations. (See Bryn McIvor, “Dan Harmon Apology,” YouTube.) Following his performance, Ganz publicly wrote the following: “Yes, I only listened because I expected an apology. But what I didn’t expect was the relief I’d feel just hearing him say these things actually happened. I didn’t dream it. I’m not crazy. Ironic that the only person who could give me that comfort is the one person I’d never ask. Please listen to it. It’s only seven minutes long, but it is a masterclass in How to Apologize. He’s not rationalizing or justifying or making excuses. He doesn’t just vaguely acknowledge some general wrongdoing in the past. He gives a full account” (@MeganGanz, Twitter, January 11, 2018). See also “Megan Ganz on Dan Harmon’s Apology: ‘I Felt Vindicated,’” *The New York Times*, January 20, 2018.
15. See MacLachlan, “Trust Me, I’m Sorry,” for further elaboration and discussion of this risk.
16. “Read Al Franken’s Apology Following Accusation of Groping and Kissing Without Consent,” CNN, November 17, 2017.
17. “Louis CK Responds to Accusations: ‘These Stories Are True,’” *The New York Times*, November 10, 2017.
18. “Norm MacDonald Apologizes After #MeToo Comments,” CNN, September 12, 2018.
19. Franken said, “what people think of me in light of this is far less important than what people think of women who continue to come forward to tell their stories. They deserve to be heard and believed. And they deserve to know that I am their ally and supporter” (“Read Al Franken’s Apology Following Accusation of Groping and Kissing Without Consent,” CNN, November 17, 2017). James Franco said, “But I completely support people coming out and being able to have a voice, because they didn’t have a voice for so long. So I don’t want to shut them down in any way” (“James Franco says accusations of sexual misconduct ‘are not accurate,’” *VOX*, January 10, 2018). Also, Aziz Ansari declared, “I continue to support the movement that is happening in our culture. It is necessary and long overdue” (Aziz Ansari Responds to Sexual Misconduct Allegations Against Him,” *VOX*, January 15, 2018).
20. While the voices of #MeToo were overwhelmingly women, I am wary of using the term exclusively—and want to honor, in particular, prominent male #MeToo voices Terry Crews and Anthony Rapp (both of whom spoke up from a marginalized position).
21. See Ezra Klein, “This Conversation Will Change How You Think about Misogyny,” *The Ezra Klein Show*, January 31, 2019.
22. See “#MeToo is Mob Justice, Not Social Justice,” *The Los Angeles Times*, September 26, 2018; “Terry Gilliam Says #MeToo Movement Has Transformed Into ‘Mob Rule,’” *Variety*, March 2016, 2018; “#MeToo has Morphed into a Career-Destroying Angry Mob,” *New York Post*, September 22, 2018.
23. See Fricker, *Epistemic Injustice*; Dotson, “Tracking Epistemic Violence, Tracking Practices of Silencing; Medina, “The Relevance of Credibility Excess in a Proportional View of Epistemic Injustice”; and Yap, “Credibility Excess and the Social Imaginary in Cases of Sexual Assault.”
24. Spelman, *Repair*, 96–97.
25. See “Richard Dreyfuss Responds to Sexual Harassment Allegations: ‘I Am Not an Assaulter,’” *Entertainment Weekly*, November 10, 2017.
26. *The Globe and Mail*, September 6, 2018.
27. *TIME*, February 26, 2018.
28. *The Japan Times*, January 14, 2019.
29. See Manne, *Down Girl*, 184.
30. *Ibid.*, 197.
31. See “Jian Ghomeshi and the Perils of the Non-Apology Apology,” *The Walrus*, September 19, 2018.
32. “‘Arrested Development’: We Sat Down With the Cast. It Got Raw,” *The New York Times*, May 23, 2018.
33. “Tom Ashbrook Asks: Is There Room for Redemption?” *The Boston Globe*, April 10, 2018.
34. “Exile,” *Harper’s Magazine*, October 2018; see also “Jian Ghomeshi, John Hockenberry, and the Laws of Patriarchal Physics,” *The New Yorker*, September 17, 2018.
35. “Jian Ghomeshi and the Perils of the Non-Apology Apology,” *The Walrus*, September 19, 2018.
36. Previous versions of this material were presented as part of a panel on #MeToo and Responses to Everyday Oppression at the 2019 meeting of the APA Central Division, and at the 2019 University of Toronto and York Undergraduate Research Conferences, where I benefited tremendously from the thoughtful engagement of each audience. I am especially grateful for the contributions and feedback of my fellow APA panelists, Clair Morrissey and Yolonda Wilson. I also wish to thank Daniel Groll, Susanne Sreedhar, and my research assistant, Olivia Sultanescu, for their assistance in drafting this essay, and especially Lauren Freeman for the opportunity to write it, and her encouragement and editorial advice as I did so.

BIBLIOGRAPHY

Ashbrook, Tom. "Tom Ashbrook Asks: Is There Room for Redemption?" *The Boston Globe*. April 10, 2018, sec. Tom Ashbrook asks. <https://www.bostonglobe.com/opinion/2018/04/10/tom-ashbrook-asks-there-room-for-redemption/IL437IlybamQgTNmTOYcWL/story.html>.

Bromwich, Jonah Engel. "Megan Ganz on Dan Harmon's Apology: 'I Felt Vindicated.'" *The New York Times*. January 20, 2018, sec. Arts. <https://www.nytimes.com/2018/01/13/arts/dan-harmon-megan-ganz.html>.

Chuba, Kirsten. "Terry Gilliam Says #MeToo Movement Has Transformed Into 'Mob Rule.'" *Variety*. March 16, 2018. <https://variety.com/2018/film/news/terry-gilliam-metoo-mob-rule-1202728819/>.

CNN. "Read Al Franken's Apology (Full Statement)." November 17, 2017. <https://www.cnn.com/2017/11/16/politics/al-franken-apology/index.html>.

Cowen, Tyler. "Does Forgiveness Have a Place in the #MeToo Movement?" *The Japan Times*, January 14, 2019. <https://www.japantimes.co.jp/opinion/2019/01/14/commentary/world-commentary/forgiveness-place-metoo-movement/>.

Deb, Sopan. "'Arrested Development': We Sat Down With the Cast. It Got Raw." *The New York Times*, May 23, 2018, sec. Arts. <https://www.nytimes.com/2018/05/23/arts/television/arrested-development-netflix-interview-jeffrey-tambor.html>.

Dotson, Kristie. "Tracking Epistemic Violence, Tracking Practices of Silencing." *Hypatia* 26, no. 2 (2011): 236–57.

Ducharme, Jamie. "Mario Batali's Misconduct Apology Came with a Recipe." *Time*. December 16, 2017. <http://time.com/5067633/mario-batali-cinnamon-rolls-apology/>.

Filipovic, Jill. "How to Find Room for Forgiveness in the #MeToo Movement." *Time*. February 26, 2018. <http://time.com/5160427/forgiving-sexual-assault-perpetrators-metoo/>.

Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. New York: Oxford University Press, 2007.

Ganz, Megan (@MeganGranz). "Please Listen to It. It's Only Seven Minutes Long..." Twitter, January 11, 2018, 12:41 am, <https://twitter.com/meganganz/status/951373403625177088?lang=en>.

Grady, Constance. "James Franco Says New Accusations of Sexual Misconduct 'Are Not Accurate.'" *Vox*. January 10, 2018. <https://www.vox.com/culture/2018/1/10/16873464/james-franco-denies-sexual-misconduct-allegations>.

Hockenberry, John. "Exile." *Harper's Magazine*. October 2018. <https://harpers.org/archive/2018/10/exile-4/>.

Klein, Ezra. "This Conversation Will Change How You Understand Misogyny." Interview with Kate Manne, *The Ezra Klein Show*, January 31, 2019. <https://www.vox.com/ezra-klein-show-podcast>.

Levin, Sam. "Author Sherman Alexie Apologizes amid Anonymous Allegations." *The Guardian*. March 1, 2018. <https://www.theguardian.com/world/2018/mar/01/sherman-alexie-author-apologizes-amid-sexual-misconduct-allegations>.

The Los Angeles Times. "#MeToo Is Mob Justice, Not Social Justice." Accessed April 7, 2019. <https://www.latimes.com/opinion/readersreact/la-ol-le-metoo-mob-justice-20180926-story.html>.

MacLachlan, Alice. "'Trust Me, I'm Sorry': The Paradox of Public Apologies." *The Monist* 98 (2015): 4.

Manne, Kate. *Down Girl: The Logic of Misogyny*. New York: Oxford University Press, 2017.

Martin, Adrienne M. "Owning Up and Lowering Down: The Power of Apology." *The Journal of Philosophy*. 107 (2010): 10.

McIvor, Bryn. "Dan Harmon Apology." YouTube video, 9:14, January 13, 2018, featuring Dan Harmon's public apology to Megan Ganz in the *Harmontown* podcast, episode 272. <https://www.youtube.com/watch?v=SyKA8ARBSkw>.

McKeon, Lauren. "Jian Ghomeshi and the Perils of the Non-Apology Apology." *The Walrus*, September 19, 2018. <https://thewalrus.ca/jian-ghomeshi-and-the-perils-of-the-non-apology-apology/>.

Medina, José. "The Relevance of Credibility Excess in a Proportional View of Epistemic Injustice: Differential Epistemic Authority and the Social Imaginary." *Social Epistemology* 25, no. 1 (2011): 15–35.

"Me Too. History & Vision." *Me Too Movement*. Accessed April 6, 2019. <https://metoomvmt.org/about/>.

Mihai, Mihaela. "When the State Says 'Sorry': State Apologies as Exemplary Political Judgments." *Journal of Political Philosophy* 21 (2013): 2.

Oh, Deborah. "Do the Men of #MeToo Deserve to Be Forgiven?" *The Globe and Mail*. September 6, 2018. <https://www.theglobeandmail.com/opinion/article-do-the-men-of-metoo-deserve-to-be-forgiven/>.

Park, Madison. "Kevin Spacey Apologizes for Alleged Sex Assault with a Minor." *CNN*. October 31, 2017. <https://www.cnn.com/2017/10/30/entertainment/kevin-spacey-allegations-anthony-rapp/index.html>.

Respers France, Lisa. "Norm Macdonald Apologizes after #MeToo Comments." *CNN*. September 12, 2018. <https://www.cnn.com/2018/09/12/entertainment/norm-macdonald-metoo-apology/index.html>.

Rice, Lynette. "Richard Dreyfuss Responds to Sexual Harassment Allegations: 'I Am Not an Assaulter.'" *Entertainment Weekly*, November 10, 2017. <https://ew.com/news/2017/11/10/richard-dreyfuss-responds-harassment-allegations/>.

Rottenberg, Josh. "Dustin Hoffman Apologizes after Allegations That He Sexually Harassed a 17-Year-Old Intern in 1985." *The Los Angeles Times*. November 1, 2017. <https://www.latimes.com/entertainment/la-et-entertainment-news-updates-dustin-hoffman-apologizes-after-being-1509561081-htmllstory.html>.

Russell, Francey. "The Paradox of Apology." Manuscript, 2019.

Smith, Kyle. "#MeToo Has Morphed into a Career-Destroying Angry Mob." *New York Post*. September 22, 2018. <https://nypost.com/2018/09/22/metoo-has-morphed-into-a-career-destroying-angry-mob/>.

Smith, Michelle R. "In Age of #MeToo, Can There Be Forgiveness, Second Chances?" *Associated Press*. April 15, 2018. <https://apnews.com/c69e3a82417f4af280376fdc10cf8651>.

Smith, Nick. *I Was Wrong: The Meaning of Apologies*. New York: Cambridge University Press, 2008.

Spelman, Elizabeth V. *Repair: The Impulse to Restore in a Fragile World*. Boston, MA: Beacon Press, 2002.

Stewart, Emily. "Aziz Ansari Responds to Sexual Misconduct Allegations against Him." *Vox*. January 15, 2018. <https://www.vox.com/identities/2018/1/15/16893468/aziz-ansari-allegations>.

The New York Times. "Statement From Harvey Weinstein." October 5, 2017. <https://www.nytimes.com/interactive/2017/10/05/us/statement-from-harvey-weinstein.html>.

The New York Times. "Louis C.K. Responds to Accusations: 'These Stories Are True.'" November 11, 2017. <https://www.nytimes.com/2017/11/10/arts/television/louis-ck-statement.html>.

Thompson, Janna. "The Apology Paradox." *The Philosophical Quarterly* (1950-) 50, no. 201 (2000): 470–75.

Tolentino, Jia. "Jian Ghomeshi, John Hockenberry, and the Laws of Patriarchal Physics." September 17, 2018. <https://www.newyorker.com/culture/cultural-comment/jian-ghomeshi-john-hockenberry-and-the-laws-of-patriarchal-physics>.

Yap, Audrey S. "Credibility Excess and the Social Imaginary in Cases of Sexual Assault." *Feminist Philosophy Quarterly* 3 (2017): 4.

Women, Work, and Power: Envisaging the Radical Potential of #MeToo¹

Robin Zheng
YALE-NUS COLLEGE

Remarkably and unmistakably, a striking number of watershed moments in the post-Trump era have been orchestrated by women. The day after the President's inauguration, an astonishing 4 million people—over 1 percent of the US population, thought to be the largest one-day political demonstration in the country's history²—participated in the Women's March on Washington, in over 673 cities across the country and around the globe.³ Equally astonishing, the next year saw public school teachers—77%

of whom are women⁴—beginning to walk out on the biggest educator strikes in a quarter-century,⁵ most notably in such conservative bastions as West Virginia, Oklahoma, Arizona, North Carolina, and Kentucky (dubbed by some the “Red State Rebellion”) but also in solidly Democratic states like California and Colorado. And then, of course, there was #MeToo.

In this essay, my interest is in thinking through #MeToo as a form of political struggle aimed at social change. In Section I, I analyze the movement alongside the Women’s March and the teacher’s strike wave, situating it between two poles of structured organizing and mass protest. This comparison highlights the imperative for movements to translate *symbolic* solidary power, which shift discourses and norms, into exercises of power that can alter *structural* conditions in a more fundamental way. Thus, in Section II, I argue that while #MeToo has been highly successful in disrupting sexist mores and patriarchal norms on the cultural front, it must match this with a commitment to transforming the fundamental material conditions that enable men’s dominance over women.⁶ Eliminating sexual harassment—particularly of non-elite women—requires addressing the underlying job insecurity, poor working conditions, and economic vulnerability that threatens almost all workers. In short, #MeToo must go radical.

I.

The Women’s March and the teachers’ strike wave exemplify what are sometimes thought of as two opposing poles of collective action: *structured organizing* and *mass protest*. Both represent modes of building the power of the people, i.e., the power of masses of “Have-Nots” who otherwise lack wealth, political influence, and the coercive backing of state-sanctioned violence.⁷ But they take very different forms. Building up grassroots power into long-lasting, disciplined, membership-based organizations is the aim of *structured organizing*, as manifest in much community and labor organizing.⁸ From this perspective, engaging in mass action without establishing organizational structure is like “skipping the first two [acts of a play], in which case there is no play, nothing but confrontation for confrontation’s sake—a flare-up and back to darkness.”⁹ By contrast, critics argue that social movements like the Depression-era labor upsurges and the Civil Rights Movement achieved their largest gains *before* individuals were folded into formally structured organizations, and lost them after.¹⁰ Large bureaucratic organizations are prone to co-optation and complacency. They are often inclined to restrain the disruptive potential of unorganized *mass protests*—that is, the spontaneous and broad-based defiance of masses of unorganized people—which can actually force the hand of elites.¹¹ I will not try to adjudicate between these views here, particularly since both types of action arguably play complementary roles in bringing about social change.¹² Without discounting the critical role of mass protest, however, I place greater emphasis on the importance of structured organizing, since it is here that I think #MeToo should learn the greatest lessons.

To set the scene, I begin by explaining some important differences between the Women’s March and the teachers’ strikes. The most obvious hallmark of the Women’s March

was its ability to energize and bring people out into the streets. The sheer size and scale of the marches—as broadcast via images of entire airplanes brimming with women’s marchers on their way to Washington, DC, solidarity protests as far-flung as Antarctica, profiles of mother-daughter marchers and first-time marchers ranging from businesswomen to poets—commanded wonder. It was an undeniable display of symbolic power; as a friend put it on Facebook, it was “the political statement we’[d] been waiting for” since November 8, 2016. For many, especially the marchers, it was deeply emotional: feeling oneself in the company of so many others was healing and fortifying. It inspired hope and replenished motivations to act in defense of a better future. In short, the Women’s March changed the conversation—boldly signaling that there would be resistance to a Trump agenda—while channeling collective anger, defiance, and determination on a massive scale.

The March thus demonstrates several features—in its etiology, nature, and effects—that are characteristic of “New Left” mass protests that coalesced in the 1960s–1970s around social identities such as race, gender, and sexuality (among other issues, like opposition to the Vietnam war). It was conceived independently on the night of Trump’s election by a retired attorney in Hawaii and a fashion entrepreneur in New York, who teamed up with seasoned activists after their Facebook events unexpectedly attracted 100,000 participants in a single day.¹³ The multi-sited, broad-based origins of the March thus epitomize what social movement activists call a “moment of the whirlwind,” that is, an atypical burst of intensive, spontaneous, and decentralized mobilization—beyond the control of any single individual or organization—that erupts after a dramatic public event.¹⁴ Moreover, the deliberately open-ended, “big tent” nature of the call for women to march encouraged an unusually broad array of participants. Indeed, even politically conservative women, also horrified by Trump’s sexist remarks on the campaign trail, participated in order to provide what they described as a “pro-life feminist voice;” others felt “misunderstood and frustrated” when the official organizers of the march released a pro-choice platform and dissolved their initial partnership with an anti-abortion organization.¹⁵ Still, the success of the March demonstrated the power of calls for “sisterhood” and “solidarity” based on social identity, and scores of participants—through selfies and social media—celebrated the personally transformative experiential dimensions of participating in mass demonstration. When Trump unrolled his notorious “Muslim Ban” in the weeks following the march, he was met by tremendously energetic protesters (and immigration lawyers willing to work *pro bono*) who rallied their forces at airports across the country.

Three years on, the Women’s March has also displayed some of the characteristic weaknesses of mass protest. Many individual participants—recruited via social media—were unyoked to local or more longstanding political organizations, and therefore unlikely to take further action after completing their personal emotional catharsis. As time wore on, moreover, the urgency of resisting the Trump presidency began to wear off, due to “protest fatigue” and the onward grind of daily life. Perhaps most importantly,

the movement did not formulate a clear list of sufficiently specific demands or concrete goals, and it was unclear how to move forward.

None of this is intended to detract from the major accomplishments of the Women's March. The same features that weakened the March's long-term prospects were crucial for its rousing symbolic appeal and ability to spur unprecedented numbers of people into the political arena, which arguably paved the way for both #MeToo and the teachers' strikes (along with other groundswells of action such as the March for Science and protests against family separations at the border). Changing the conversation is a key step in changing social conditions, as I discuss in Section II with respect to #MeToo, because disruption functions to loosen the grip of unreflective ideology and subject it to questioning.¹⁶ But it is not sufficient. Momentary disruptions can be brushed past, and powerful elites will not be swayed unless they are under pressure, that is, unless and until their self-interest is actually at stake.¹⁷

What I want to stress here is that the collective power of the people—the ability of groups to apply pressure in the absence of control over wealth, political influence, or coercive state backing—is held purely in virtue of two things: their numbers, and the fact that their acquiescence (and labor) is indispensable for social functioning. Both historically and in the present, structured organizing has played a key role in unlocking this kind of collective power, which brings me now to the teachers' strike wave. Although teaching—itself a form of gendered “carework”—is predominantly performed by women,¹⁸ and although the teachers amassed large numbers at their rallies, the logic of what were effectively *industrial strikes* differed greatly from that of the Women's March. For their genesis, the teachers' strikes relied on the existence of established, nationally networked labor unions¹⁹ such as the American Federation of Teachers and the National Education Association. The union membership, often recruited laboriously over years in the workplace or through face-to-face door knocking, had deep roots and strong support in their communities; during strikes, they coordinated free meals and childcare together with local businesses, churches, and non-profits. The strikes themselves were the result²⁰ of careful deliberation stemming from a formalized collective decision-making procedure (i.e., strike authorization votes), and each teachers' union struck for a highly specific list of demands—no changes to healthcare premiums, restoration of state funding, the employment of nurses and librarians, etc.—generated by the bargaining process. More experienced unions intentionally studied analyses and tactics of previous strikes, and they exercised tightly regimented discipline even on the picket lines, where attendance was taken and reported to the central strike committee. In virtue of the bargaining cycle, there are clear “next steps” to take after striking: enforcing the agreed-upon contract, filing grievances for violations, and building membership for the next round of bargaining when the contract expires. (Moreover, strikes themselves can escalate insofar as other industries join in or refuse to cross picket lines.)

Most importantly, the strikers did not demonstrate power only symbolically but *exercised it* by withholding their labor: the work stoppages cost untold millions and forced recalcitrant policymakers to make genuine concessions. And though each union's demands were specific to their work situations, the underlying threats—evaporating funding, increasing privatization of public services, attacks on collective bargaining rights—were common to teachers nationwide, as well as other workers. In West Virginia, rank-and-file members of the union defied their leaders by refusing to accept a deal that gave teachers a 5 percent raise while only giving 3 percent to other public workers, perpetuating a wildcat strike until the demand was secured for all.²¹ In Los Angeles, teachers demanded²²—and won—an end to racial profiling in schools, a legal defense fund for students and parents at risk of deportation, and the construction of green spaces, in addition to basic pay raises and the elimination of a clause that routinely allowed class size caps to be overruled.²³ Impressively, they secured a temporary moratorium on charter schools,²⁴ as part of a growing statewide and national consciousness that the privatization of education via explosive growth in charter schools amounts to an undermining of the public good for the benefit of wealthy profiteers. The teachers' deployment of “bargaining for the common good” demonstrates the prodigious potential for industrial action, even by a historically weakened labor movement, to force social change across multiple dimensions and levels. Through the long process of education and consciousness-raising, cultivating internal and external relationships, and democratic decision-making that is necessary for structured organizing to produce mobilization, workers thereby build up their collective capacity to win larger and longer-lasting structural changes to their shared social conditions.

In light of the foregoing discussion, I propose to understand the #MeToo movement as lying somewhere on the spectrum between these two poles. In its origins, of course, #MeToo began as a moment of mass protest: after a damning exposé covering decades of allegations against director Harvey Weinstein, actress Alyssa Milano's tweet in support of the women who reported him went viral on October 15, 2017. The #MeToo hashtag garnered 825,000 uses on Twitter and 4.7 million interactions on Facebook within 24 hours; indeed, more than 45% of Facebook users in the US were estimated to be friends with someone who posted “Me too.”²⁵ Without doubt, #MeToo changed the conversation by, in the words of Milano's tweet, “giv[ing] people a sense of the magnitude of the problem.” A day later, she identified community organizer Tarana Burke as the founder of an earlier #MeToo movement, and the “moment” began to consolidate into a more unified “movement.” Milano, meeting Burke in person for the first time on the *Today* show in December, clasped hands with her while professing that they had developed a friendship involving multiple texts a day. Their interview came on the heels of a *TIME* magazine article declaring “the Silence Breakers” Person of the Year, which conspicuously joined the experiences of (mostly White) celebrity and professional women of varying political persuasions to those of (non-White) women employed as a hotel worker, dishwasher, strawberry picker, housekeeper, and office assistant.²⁶ In the first week of January 2018, a number of actresses

coordinated a headline-grabbing action at the Golden Globes Awards in which they wore black, brought longtime activists of color such as Burke as their dates, and rehearsed a well-organized message: that sexual harassment must be eliminated “not just in this industry, but every industry,” and that supporters should donate to their legal defense fund.²⁷ By the end of the month, they had raised \$21 million for the fund,²⁸ which is being administered by the National Women’s Law Center to provide free legal assistance, with priorities for LGBTQ people and people of color, as well as those in low-income or non-traditional employment and those facing retaliation.

In tandem with such high-profile symbolic actions, the incipient power of #MeToo grew apparent as time passed. One year later, over 200 prominent men had been fired from their positions (about half of which were then filled by women),²⁹ while estimates of public allegations ranged between twice to four times that number,³⁰ across all different industries: news media, government and politics, art and entertainment, finance, education, and medicine. State legislatures passed a number of laws in response, especially those targeting non-disclosure agreements, rape kits, and statutes of limitation.³¹ #MeToo thus typifies what activists call the toppling of “pillars of support,” i.e., changes in prevailing opinion within major social institutions such as the media, courts, government, corporations, and religious communities, which represents a crucial pathway toward social change.³²

The political potential of #MeToo is also evident in the logic of its demands. The problem of sexual harassment lends itself to being treated fundamentally as a *workplace issue*. By this I do not mean to imply that harassment does not occur outside the workplace. But the workplace by nature is a highly specific context in which people are first and foremost required to perform productive tasks; social and non-productive activities are condoned only insofar as they contribute to or do not hinder productivity. In this sense (for better or worse) it has traditionally been deemed a “public” space, defined against “private” activities such as sexual intimacy and reproduction. It is no wonder, then, that sexual harassment is paradigmatically thought of—and legally prohibited as—a workplace phenomenon. This connection to work means that the #MeToo movement has built-in possibilities for leveraging the strategies of structured organizing found in the labor movement. Indeed, #MeToo has already demonstrated an extraordinary capacity for forging common interests between women ranging from the most vulnerable to the most privileged. Less than a month after #MeToo went viral, the Alianza Nacional de Campesinas (National Alliance of Women Farmworkers) published an open letter of solidarity with the actresses. At the same time, #MeToo galvanized the most elite and well-paid of white-collar workers: in November 2018, Google employees conducted a global walkout across 60 percent of its offices. Prompted by the news that major executives were secretly given cushy exit packages after allegations of sexual misconduct, they issued a list of demands for better accountability, transparency, and wage equality that even included a brief mention of contingent workers.³³ In the next section, I argue that #MeToo must tap into this radical potential and build itself from the support of a broad

coalition of workers, if it is to achieve its aims of eliminating sexual harassment against (*all*, and not just some) women.

II.

Notwithstanding #MeToo’s considerable victories, much remains to be done. Unsurprisingly, given #MeToo’s origins in the world of Hollywood celebrities, its effects have primarily been felt in what Patricia Hill Collins calls the “hegemonic domain of power,” that is, in the body of collectively shared images, narratives, tropes, concepts, and values that perform the ideological function of rationalizing an unjust status quo.³⁴ For many men, #MeToo shattered the illusion that sexual harassment is caused by a few “bad apples,” incontrovertibly establishing that our society systemically and ubiquitously enables perpetrators to commit wrongdoing and remain protected from consequences. For many women, #MeToo reduced the feelings of shame, self-doubt, and guilt that often plague targets of harassment and prevent them from reporting. The sheer scale of the thousands who began to “break the silence”—not only on social media, but through filing reports—vividly demonstrated to individual women that they were not alone,³⁵ which thereby introduced two complementary shifts in the narrative. First, it resolved women’s doubts about their own experiences: if so many others experienced harassment, then it was not attributable to naivety, carelessness, overreaction, or some other fault of their own. Simultaneously, it forced the public to finally recognize that if women were not the problem (not “asking for it”), then *men* were. The importance of these changes in perspective should not be underestimated. By validating women’s experiences of harassment, and by holding men accountable for perpetrating harassment on a global scale, #MeToo profoundly altered the possibilities for women to report, and more importantly, to be believed—both necessary steps for resolving and preventing future cases of harassment.

Unfortunately, these changes are not as easily incorporated into the “structural domain of power,” that is, the large-scale social institutions that configure individuals’ life prospects such as law, government, markets, media, and systems of education, healthcare, housing, etc.³⁶ According to Carol Moody, president of the country’s oldest legal advocacy group for women: “Everybody thinks some massive, massive change in laws has happened with the #MeToo movement. But it hasn’t. Nobody will disagree something needs to be done, but the devil’s in the details.”³⁷ Changing the law takes time; and the more substantial the change (or the more it threatens powerful interests), the harder it is to pass. Moreover, merely changing the law on the books is insufficient for improving conditions on the ground. Catharine MacKinnon writes:

It is widely thought that when something is legally prohibited, it more or less stops. This may be true for exceptional acts, but it is not true for pervasive practices like sexual harassment, including rape, that are built into structural social hierarchies. . . . If the same cultural inequalities are permitted to operate in law as in the behavior the law prohibits, equalizing attempts—such as sexual harassment law—will be systemically resisted. . . . [T]oday’s

survivors are initiating consequences none of them could have gotten through any lawsuit—in part because the laws do not permit relief against individual perpetrators, but more because they are being believed and valued as the law seldom has. Women have been saying these things forever. It is the response to them that has changed.³⁸

The lesson here, then, is that #MeToo must *structurally* solidify the gains it has made in altering the wider, extra-legal “cultural inequalities” that allow powerful individuals to ignore the law.

Hegemonic gains, however, are easy to reverse if rooted only in public opinion, as was proved all too clearly by the devastating confirmation trial of Supreme Court Justice Brett Kavanaugh. The trial not only showed that the testimony of even an extremely competent, compelling, socially respectable White witness could be overridden, but also that #MeToo’s narrative was vulnerable to backlash and distortion. While Trump declared that it was a “very scary time for young men in America,” 40 percent of Americans stated that the #MeToo movement had gone too far.³⁹ These developments expose the limitations of a strategy aimed at taking down individual bad actors one at a time. To be sure, punishing perpetrators is undeniably important. It is needed for targets of harassment to heal, long overdue to them on grounds of justice, and vital for the shift in norms I described earlier. But no matter how many big fish get skewered, their demise in itself cannot constitute the structural transformation that is ultimately needed. So long as the conditions of unequal power that enable men to harass women remain in place, harassment will continue (and the most powerful and deplorable of perpetrators will continue to get away with it).

To better understand this, I suggest that we must, in the words of bell hooks, bring “from margin to center” the experiences of poor and working-class women of color.⁴⁰ We must examine the conditions that turn them into deliberate targets of harassment and rape. What becomes evident in these cases is that sexual harassers intentionally use women’s subordinate gender and class status against them, coercing women into sexual activity—and then into silence—by threatening to have them deported or deprived of their jobs if they do not comply or if they consider reporting.⁴¹ The situation is compounded by the fact that the very nature of the work itself—picking vegetables in large fields, scrubbing floors in hallways after hours, cleaning people’s homes and hotel rooms, etc.—requires women to be in situations where they are isolated and unsupervised. This makes it easy for men to physically overpower them, and it also undermines women’s chances of making credible reports without eyewitnesses.⁴² To reiterate an earlier point: no matter how stringent the legal protections against harassment may be, they will be of no use unless enforced. The only way to truly protect these women from harassment is to reduce the stark disparities in power that exist between them and their employers.

This lesson has been brilliantly understood by the Fair Food Program (FFP), an initiative founded by the Coalition of Immokalee Workers (CIW). The CIW is an organization

of farmworkers drawing on popular education techniques pioneered by peasant and community organizers in Latin America. Where rates of sexual harassment and assault amongst women farmworkers have been revealed to be as high as 80 percent elsewhere, only a single case of sexual harassment involving physical contact has been found since 2013 on farms participating in the FFP.⁴³ This extraordinary achievement has been accomplished by enormously empowering FFP farmworkers in multiple areas. The FFP works by establishing legally binding contracts between the CIW and major corporate buyers such as Taco Bell, McDonald’s, and Walmart (fourteen in all so far), which the CIW identified as the primary cause of poor working conditions due to their combined ability to force down the price of produce, which in turn suppresses workers’ wages and incentivizes exploitative labor practices.⁴⁴ Participating corporations, which were initially pressured to sign on due to consumer boycotts, are subject to two requirements. First, they must pay a “one penny per pound” premium toward farmworkers’ wages, which has generated a wage increase of 50–70 percent,⁴⁵ and second, they must cease purchases from any farms that are found to violate the FFP Code of Conduct, as determined by an independent monitoring organization called the Fair Food Standards Council (FFSC).⁴⁶ The Code of Conduct, which was developed according to workers’ own knowledge of exploitative practices and over which the CIW maintains veto power, includes zero tolerance for forced labor, child labor, and sexual harassment;⁴⁷ between 2011 and 2017, nine such cases of sexual harassment involving physical contact were found, all of which resulted in offenders being terminated and a ban from future FFP employment of up to two years.⁴⁸

The success of the FFP depends on its aggressive, multi-level approach to enforcement. First, worker-to-worker education programs for all 35,000 FFP workers—compensated by the hour, repeated on multiple occasions, and attended by farm managers to indicate buy-in—serve to inform employees of their rights (including protection from retaliation) and the reporting process, which creates an unparalleled on-the-ground monitoring system far outstripping the capacity of government investigators and commercial auditors. Second, the FFSC implements a 24/7 hotline and reporting system that typically resolves complaints within two weeks; cases of retaliation that involve the firing or denial of work to an employee, for instance, are resolved by immediate termination or public reprimand of the retaliator. Third, the FFSC conducts audits in which it interviews at least half of the workers on every farm; notably, these audits are performed by the same personnel who answer the hotlines and are therefore uniquely positioned to detect patterns of repeat offense.⁴⁹ In sum, the FFP provides farmworkers with virtually unheard-of rights and protections against their employers, enforced by the threat of losing sales to corporate buyers that are legally bound to cease purchases from FFSC-sanctioned farms.

The FFP’s consumer-worker alliance thus represents one concrete model according to which #MeToo could pursue structured organizing. Ultimately, the model depends on both the ability of consumers to pressure corporations into joining the program and the ability of the FFSC to carry out

its aggressive enforcement program. Unfortunately, both of these have inherent weaknesses, and may be difficult to scale. Consumer boycotts are hard to sustain outside industries like food and clothing, which are particularly likely to gain traction due to politically active constituencies like college students. Corporations may steadfastly refuse to sign FFP contracts even in the face of consumer pressure (as Wendy's has consistently done), or may eventually decide to simply back out.⁵⁰ Insofar as suppliers can still sell their products to non-participating corporations—and so long as there remain populations of potential workers willing to tolerate poor working conditions⁵¹—the power of the FFP is diminished. Moreover, the FFSC itself is a non-profit organization funded in large part by philanthropic organizations like the Kresge and Kellogg Foundations,⁵² which are by their nature utterly dependent on private capital and hence unlikely to support initiatives that represent any real, large-scale threat to employers' domination over workers. This is not in any way to retract my support from the FFP and similar initiatives, but only to recognize their limitations.

The case of farmworkers and the FFP lucidly illustrates the fundamental connection between sexual harassment and women's economic vulnerability. At the end of the day, the only failsafe way to empower workers is to ensure that they have viable exit options—that is, an ability to find stable employment elsewhere. Unless marginalized women attain job security and freedom from the fear of poverty, the underlying material conditions of their lives will continue to perpetuate their economic vulnerability—and, hence, their vulnerability to harassment. In short, as socialist feminists have long argued, the problem of sexual harassment against women cannot be solved without confronting the larger problem of capitalist exploitation.

Let me be clear: there are two distinct arguments here. The first, it must be emphasized, is *not* an instance of what Angela Harris calls "nuance theory," in which, for example, cursory consideration of the plights of women of color "reduces racial oppression to a mere intensifier of gender oppression" that demonstrates more clearly how *all* women suffer.⁵³ My point here is that the sexual harassment faced by working-class women on the job, and by Black, Asian, and Latina women subordinated on the basis of race, ethnicity, and immigrant status (who are also disproportionately poor and working-class), is qualitatively different from that faced by middle-class White women. The only way to address their vulnerability to sexual harassment is to fundamentally restructure the capitalist economic order in such a way that they are no longer *economically* vulnerable. If #MeToo's proponents are serious about supporting the working-class women of color with whom they have dramatically and symbolically joined forces, then they must be serious about working against capitalism as we know it.

That said, there are undeniable resonances that should not be ignored. First of all, the material dimensions of coercion are present for many women even outside the lowest-paid, menial jobs. Aspiring actresses more often than not also struggle to make ends meet, and many middle-class women are far from being financially independent. Furthermore, the working conditions of increasing numbers of women

in traditionally well-paid and/or prestigious industries are being eroded by the "gig economy," or by two-tier systems in which stable employment is restricted to a tiny minority while the majority works precariously in contingent positions; academia, media, and the arts represent particularly clear examples of the latter. Moreover, gender oppression is not wholly reducible to class exploitation: prevailing gender norms, even after #MeToo, make it difficult even for women of means to adequately resist or report sexual harassment. And much sexual harassment occurs outside the workplace. But gender norms—according to which women should be sexually available (yet appropriately modest), physically attractive (yet not overly so), compliant and cooperative (though that prevents them from demonstrating leadership), instinctively nurturing (though that signals lack of commitment to their careers), and so on—are unlikely to change, or stay changed, without more wholesale structural changes in the sexual division of labor between men and women as a whole. In this sense, the fates of all women are connected, though not equivalent. Specifying precisely what would be needed to dismantle the gender norms that enable sexual harassment is not something I can address within the confines of this essay, but it is clear that the requisite changes would need to radically alter the fundamental social organization of our world.

Whether #MeToo can manage to realize its radical potential, however, remains very much in the balance. For such a task is not at all easy, and is yet to even appear on the agendas of many ardent #MeToo supporters. I suspect that if it does, it will be due in no small part to interweaving the movement with a much larger confluence of anti-oppressive forces that are currently gathering on multiple fronts. If #MeToo wants to achieve its own goals, it should welcome these anti-racist, anti-heteronormative, anti-ableist, anti-fascist, pro-workers' movements as natural and necessary allies.

NOTES

1. This work is supported by a Ministry of Education (MOE) Tier 1 funded grant (IG17-SR103) from Yale-NUS College.
2. Chenoweth and J. Pressman, "This Is What We Learned by Counting the Women's Marches."
3. Przybyla and Schouten, "At 2.6 Million Strong, Women's Marches Crush Expectations."
4. Bhattacharya, "Women Are Leading the Wave of Strikes in America. Here's Why."
5. Kerrissey, "Teacher Strike Wave: By the Numbers."
6. Of course, men suffer from sexual harassment as well. However, given space constraints, I focus here on dynamics particular to men's harassment of women.
7. Alinsky, *Rules for Radicals: A Practical Primer for Realistic Radicals*.
8. *Ibid.*; Engler and Engler, *This Is an Uprising: How Nonviolent Revolt Is Shaping the Twenty-First Century*.
9. Alinsky, *Rules for Radicals*, xix.
10. Piven and Cloward, *Poor People's Movements: Why They Succeed, How They Fail*.
11. *Ibid.*
12. Engler and Engler, *This Is an Uprising*.

13. Hintz-Zambrano, "Bob Bland of the Women's March on Raising Children in the Resistance, & Much More"; Woerner, "Who Started the March? One Woman."
14. Engler and Engler, *This Is an Uprising*, 188.
15. Stein, "Is There a Place at the Women's March for Women Who Are Politically Opposed to Abortion?"
16. Haslanger, *I—Culture and Critique*.
17. Cf. Alinsky, *Rules for Radicals*, xxxiv.
18. Bhattacharya, "Women Are Leading the Wave of Strikes in America. Here's Why."
19. Importantly, however, the willingness of rank-and-file teachers in West Virginia to act against their official union leadership (discussed briefly below) was key to securing their eventual victory (Blanc, "Betting on the Working Class").
20. While some strikes were openly inspired by the West Virginia teachers, the UTLA strikes in Los Angeles had been in preparation for years.
21. Petersen-Smith, "West Virginia Put Class Struggle Back on the Map."
22. The original demands even included an ambitious call for building affordable housing, the lack of which is one of the primary causes of teachers' leaving the field (La Ganga and Chabria, "Teachers' Strike Fueled by Bay Area Housing Crisis: 'They Can't Afford Oakland'"), on unused land.
23. UCLA Center X, "Bargaining for the Common Good: Alex Caputo-Pearl."
24. Stokes, "LAUSD Calls For Charter Moratorium After Teachers Strike."
25. Santiago and Criss, "An Activist, A Little Girl and the Heartbreaking Origin of 'Me Too'."
26. Zacharek, Dockterman, and Sweetland Edwards, "Person of the Year 2017: The Silence Breakers."
27. Sen, "The Lefty Critique of #TimesUp Is Tired and Self-Defeating."
28. Seale, "What Has #MeToo Actually Changed?"
29. Carlsen et al., "#MeToo Brought Down 201 Powerful Men. Nearly Half of Their Replacements Are Women."
30. Griffin, Recht, and Green, "#MeToo: One Year Later."
31. However, most legal changes remained relatively limited in scope—see (Kelly and Hegarty, "#MeToo Was a Culture Shock. But Changing Laws Will Take More Than a Year") discussion in the next section.
32. Engler and Engler, *This Is an Uprising*, 107.
33. Stapleton et al., "We're the Organizers of the Google Walkout. Here Are Our Demands."
34. Collins, *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*.
35. As Piven and Cloward (*Poor People's Movements: Why They Succeed, How They Fail*, 12) write of social movements in the 1930s and 1960s: "The sheer scale of these dislocations helped to mute the sense of self-blame, predisposing men and women to view their plight as a collective one, and to blame their rulers for the destitution and disorganization they experienced."
36. Collins, *Black Feminist Thought*.
37. Kelly and Hegarty, "#MeToo Was a Culture Shock."
38. MacKinnon, "#MeToo Has Done What the Law Could Not."
39. Smith, "On #MeToo, Americans More Divided By Party Than Gender."
40. hooks, *Feminist Theory: From Margin to Center*.
41. Yeung, *In a Day's Work: The Fight to End Sexual Violence Against America's Most Vulnerable Workers*.
42. Ibid.
43. Fair Food Standards Council, *Fair Food 2017 Annual Report*.
44. Asbed and Hitov, "Preventing Forced Labor in Corporate Supply Chains: The Fair Food Program and Worker-driven Social Responsibility."
45. Dearing, "The Fair Food Program."
46. Asbed and Hitov, "Preventing Forced Labor in Corporate Supply Chains."
47. Ibid.
48. Fair Food Standards Council, *Fair Food 2017 Annual Report*.
49. Asbed and Hitov, "Preventing Forced Labor in Corporate Supply Chains."
50. I owe this argument to David Smith.
51. I am indebted to Xing Xia for discussion of this point.
52. Williams, "Fair Food Standards Council Oversees Compliance with Fair Food Agreement."
53. Harris, "Race and Essentialism in Feminist Legal Theory," 595.

REFERENCES

Alinsky, S. D. *Rules for Radicals: A Practical Primer for Realistic Radicals*. New York: Vintage Books, 1989.

Asbed, G., and S. Hitov. "Preventing Forced Labor in Corporate Supply Chains: The Fair Food Program and Worker-driven Social Responsibility." *Wake Forest L. Rev.* 52 (2017): 497.

Bhattacharya, T. "Women Are Leading the Wave of Strikes in America. Here's Why." *The Guardian*. April 10, 2017. Retrieved from <https://www.theguardian.com/commentisfree/2018/apr/10/women-teachers-strikes-america>.

Blanc, E. "Betting on the Working Class." *Jacobin*. May 29, 2017.

Carlsen, A., M. Salam, C. Cain Miller, D. Lu, A. Ngu, J. K. Patel, and Z. Wichter. "#MeToo Brought Down 201 Powerful Men. Nearly Half of Their Replacements Are Women." *The New York Times*. October 29, 2018. Retrieved from <https://www.nytimes.com/interactive/2018/10/23/us/metoo-replacements.html>.

Chenoweth, E., and J. Pressman. "This Is What We Learned by Counting the Women's Marches." *The Washington Post*. February 7, 2017. Retrieved from <https://www.washingtonpost.com/news/monkey-cage/wp/2017/02/07/this-is-what-we-learned-by-counting-the-womens-marches/>.

Collins, P. H. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. New York: Routledge, 1990/2000.

Dearing, P. "The Fair Food Program." 2017. Retrieved from *Harvard Business Review*: <https://www.bridgespan.org/fair-food-program>.

Engler, M., and P. Engler. *This Is an Uprising: How Nonviolent Revolt Is Shaping the Twenty-First Century*. New York: Nation Books, 2016.

Fair Food Standards Council, *Fair Food 2017 Annual Report*. 2017. Retrieved from <http://www.fairfoodstandards.org/reports/>.

Griffin, R., H. Recht, and J. Green. "#MeToo: One Year Later." *Bloomberg News*. October 5, 2018. Retrieved from <https://www.bloomberg.com/graphics/2018-me-too-anniversary/>.

Harris, A. "Race and Essentialism in Feminist Legal Theory." In *Critical Race Feminism*, edited by A. K. Wing, 11–26. New York: New York University Press, 1997.

Haslanger, S. "I—Culture and Critique." Paper presented at the Aristotelian Society Supplementary Volume. 2017.

Hintz-Zambrano, K. "Bob Bland of the Women's March on Raising Children in the Resistance, & Much More." *Mother*. September 1, 2017.

hooks, b. *Feminist Theory: From Margin to Center*. Boston, MA: South End Press, 1984.

Kelly, C., and A. Hegarty. "#MeToo Was a Culture Shock. But Changing Laws Will Take More Than a Year." *USA Today*. October 5, 2018. Retrieved from <https://www.usatoday.com/story/news/investigations/2018/10/04/metoo-me-too-sexual-assault-survivors-rights-bill/1074976002/>.

Kerrissey, J. "Teacher Strike Wave: By the Numbers." *Labor Notes*. 2018. Retrieved from <https://labornotes.org/blogs/2018/10/teacher-strike-wave-numbers>.

La Ganga, M. L., and A. Chabria. "Teachers' Strike Fueled by Bay Area Housing Crisis: 'They Can't Afford Oakland.'" *The Los Angeles Times*. February 21, 2019. Retrieved from <https://www.latimes.com/local/education/la-me-edu-oakland-strike-2019-story.html>.

MacKinnon, C. A. "#MeToo Has Done What the Law Could Not." *New York Times*. February 4, 2018. Retrieved from <https://www.nytimes.com/2018/02/04/opinion/metoo-law-legal-system.html>.

Petersen-Smith, K. "West Virginia Put Class Struggle Back on the Map." *Socialist Worker*. March 15, 2018. Retrieved from <https://socialistworker.org/2018/03/15/west-virginia-put-class-struggle-back-on-the-map>.

Piven, F. F., and R. A. Cloward. *Poor People's Movements: Why They Succeed, How They Fail*. New York: Vintage Books, 1979.

Przybyla, H. M., and F. Schouten. "At 2.6 Million Strong, Women's Marches Crush Expectations." *USA Today*. January 21, 2017. Retrieved from <https://www.usatoday.com/story/news/politics/2017/01/21/womens-march-aims-start-movement-trump-inauguration/96864158/>.

Santiago, C., and D. Criss. "An Activist, A Little Girl and the Heartbreaking Origin of 'Me Too'." *CNN.com*. October 17, 2017. Retrieved from <https://edition.cnn.com/2017/10/17/us/me-too-tarana-burke-origin-trnd/index.html>.

Seale, R. "What Has #MeToo Actually Changed?" *BBC News*. May 12, 2018. Retrieved from <https://www.bbc.com/news/world-44045291>.

Sen, R. "The Lefty Critique of #TimesUp Is Tired and Self-Defeating." *The Nation*. 2018.

Smith, T. "#MeToo, Americans More Divided By Party Than Gender." *NPR News*. October 31, 2018. Retrieved from <https://www.npr.org/2018/10/31/662178315/on-metoo-americans-more-divided-by-party-than-gender>.

Stapleton, C., T. Gupta, M. Whittaker, C. O'Neil-Hart, S. Parker, E. Anderson, and A. Gaber. "We're the Organizers of the Google Walkout. Here Are Our Demands." *New York*. October 1, 2018. Retrieved from <https://www.thecut.com/2018/11/google-walkout-organizers-explain-demands.html>.

Stein, P. "Is There a Place at the Women's March for Women Who Are Politically Opposed to Abortion?" *The Washington Post*. January 18, 2017. Retrieved from https://www.washingtonpost.com/local/social-issues/is-there-a-place-for-anti-abortion-women-at-the-womens-march-on-washington/2017/01/17/2e6a2da8-dcbd-11e6-acdf-14da832ae861_story.html.

Stokes, K. "LAUSD Calls For Charter Moratorium After Teachers Strike." *LAist*. January 29, 2019. Retrieved from https://laist.com/2019/01/29/lausd_board_asks_state_to_temporarily_half_new_charter_schools_from_opening.php.

UCLA Center X. "Bargaining for the Common Good: Alex Caputo-Pearl." *Just Talk: Voices of Education & Justice*. September 28, 2017. Retrieved from <https://www.utla.net/news/bargaining-common-good-alex-caputo-pearl>.

Williams, A. B. "Fair Food Standards Council Oversees Compliance with Fair Food Agreement." *The News-Press*. June 16, 2012. Retrieved from <https://www.nesri.org/news/2012/06/fair-food-standards-council-oversees-compliance-with-fair-food-agreement>.

Woerner, M. "Who Started the March? One Woman." *The Los Angeles Times*. 2017. Retrieved from <https://www.latimes.com/nation/la-na-pol-womens-march-live-who-started-the-march-one-1485033621.htmlstory.html>.

Yeung, B. *In a Day's Work: The Fight to End Sexual Violence Against America's Most Vulnerable Workers*. New York: The New Press, 2018.

Zacharek, S., E. Dockterman, and H. Sweetland Edwards. "Person of the Year 2017: The Silence Breakers." *TIME*. October 15, 2017.

Field Notes on Conference Climate: A Decade with the Philosophy of Science Association Women's Caucus

Julia R. S. Bursten
UNIVERSITY OF KENTUCKY

Between 2008 and 2018, registration at the biennial meeting of the Philosophy of Science Association doubled. During that same interval, attendance at meetings of the Philosophy of Science Association Women's Caucus grew nearly fourfold. It is still not a large percentage of the meeting attendees—a quick back-of-the-envelope calculation revealed that just over 20 percent of 2018 PSA registrants attended the 2018 Women's Caucus meeting—but there is a marked difference between what forty people can accomplish and what one hundred fifty can.

I served as co-chair of the Caucus from 2015 to 2018, and many of our most successful initiatives aimed to make a more humane conference environment for all participants, rather than focusing solely on the experience of women at the conference. This focus was not preplanned or coordinated by the co-chairs; rather, it emerged from the interests and energies of Caucus members. Upon reflection, it is clear that this focus both came from and contributed to my desire to see gender parity become not just a reality, but an unremarkable one, at future PSA meetings.

I attended my first PSA meeting in 2008, a few months after beginning graduate school. I would not have gone to a conference at all at that stage, but it was in the city where I lived, and my department was looking for volunteers to help manage the registration desk. Plus, classes were cancelled, since most of my professors were involved in the conference in some way.

The conference was a transformative experience for me. I was still adjusting to the idea that philosophy of physics was something that one could just *do*, as a career, and it filled me with hope and excitement to enter sessions on quantum entanglement and statistical mechanics and find crowded rooms. I remember noticing at the time, without much reflection, how easy it was to get a seat in the women's restroom even at the end of standing-room-only talks.

That year, when I showed up early on Saturday morning to attend the PSA Women's Caucus breakfast meeting (having failed to convince the other women in my graduate cohort to join me at 7:45 a.m. on a weekend) I figured the forty-person attendance rate was due mostly to the timing. After all, my undergraduate thesis advisor, a philosopher of science, had been a woman, and another woman had taught my first philosophy of science class. My graduate department was chaired by a woman. Three other women entered my graduate program with me. Sure, I was the only woman enrolled in my philosophy of quantum mechanics seminar, but sometimes weird flukes happen. It wasn't until the next PSA meeting in 2010 that I noticed the pattern and drew the connection between the pleasantly un-crowded

public restrooms and the quantity of leftover pastries at the Women's Caucus breakfast.

The point of this personal anecdote is not to convince the readers of this newsletter that philosophy, or philosophy of science, has a gender representation issue. That is a given. And, for my part, my experience as a woman in philosophy of science got much worse before it got better, in all the mundane and terrifying ways gender-based discrimination and harassment are wont to do. The point is that once the wool came off my eyes, I wanted nothing more than to have it back. Not to go back to my ignorance, but rather to find my way to a place where I could just *do* my philosophy, without constantly scratching at the itchy, ill-made sweater of self-conscious awareness of my gender every time I presented work or submitted a paper. So, when I was approached for the position of early-career co-chair of the PSA Women's Caucus in 2014, I figured that getting involved in activism for women philosophers of science may at least soften, if not reshape, the sweater.

When my term came to an end this winter, I reflected on a number of the efforts I undertook with my co-chairs and committee heads. The ones I am most proud of are all about softening the sweater, trying to make it possible to forget for a few minutes or hours that the women philosophers of science at PSA meetings were *women* philosophers of science. Selfishly, perhaps, they were geared toward recasting the conference as I first saw it in 2008, with the naïve enthusiasm made possible by a career that, despite its brevity, had prominently featured other women doing philosophy of science.

Many of these initiatives are not specific to women. Rather, many of them aimed to make a more humane conference environment for all involved, and in so doing, make it easier for women to exist in the conference space—and to exist as philosophers, rather than as *women* philosophers. For instance, beginning before my tenure, the Caucus spearheaded an initiative to distribute dependent-care grants for meeting participants. In 2016 and 2018, the Caucus was also involved in the governing board's decision to provide on-site childcare. We developed a mentoring program for scholars in philosophy of science from underrepresented groups. We compared notes with other coalitions that supported the work of minoritized groups in related areas, such as Minorities and Philosophy and the History of Science Society Women's Caucus, and we saw how few of our problems were unique to the PSA. We invited allies to our business meetings, and they came.

In 2016, we sponsored the first Women's Caucus Prize Symposium, a special session selected by a committee of our membership as an exemplar of quality philosophy of science done by women, responding to women philosophers of science, and about issues of interest to Caucus members. The Caucus has sponsored two prize symposia now, one on animal cognition and one on climate change. Each symposium prominently featured both early-career and established scholars of multiple genders, and each has centered on a cutting-edge topic. Each has hosted at least one scientist, and each has aimed to advance philosophical methodology as well as subject matter. Both

symposia have been delivered to packed rooms, and I have had to wait in line in the restroom after both of them. Both times, standing in those lines nudged awake an awareness of my gender. However, unlike in 2010, the reminders were welcome and revitalizing, largely because the awareness had been allowed to lie dormant while hours of philosophy had happened around it. Coming back to awareness from that perspective was pleasantly bracing, a cool breeze to remind you why you brought a sweater in the first place.

In advance of the most recent PSA meeting in 2018, we assembled [a flyer on Bystander Intervention](#) and distributed it in all attendees' registration packets. The flyer reviews basic information on what constitutes sexual harassment and discriminatory behavior, and it offers a variety of quick, in-the-moment strategies to reroute an instance of discrimination and support the targeted person. Like many of our other initiatives, it is not geared solely toward the experience of women at the conference; it addresses discrimination due to other aspects of personal identity as well. The flyer was developed primarily by Janet Stembwedel, the head of our professional climate committee, and approved by the Caucus at large prior to distribution.

We assembled this flyer because we were not able to get an enforceable sexual harassment policy on the books for the 2018 meeting, due to logistical and budgetary constraints. There was a conduct statement that attendees were asked to sign upon online registration, and at the conclusion of the 2018 meeting, the Governing Board approved the plan to implement an enforceable policy for 2020. We have no data on how many registrants read the flyer or found it useful, and I harbor few illusions. It was a flyer in a registration packet. But it was a flyer that showed up in every PSA registration packet, indicating that the association at large has a vested interest in combatting discriminatory behavior at PSA meetings.

Two anecdotes from the meeting suggest that the culture is changing in more substantial ways. The first has to do directly with the flyers. During the meeting, a relatively well-known philosopher of science wrote a blog post on a not-quite-anonymized personal blog decrying the mentoring initiative and the distribution of bystander intervention flyers. As of the time of this writing, the blog post still has no comments, while a Caucus member's social media post about the blog has over seventy comments variously pointing to logical flaws in the argument, offering biting ripostes, and generating a sense of solidarity among women attendees and allies at the meeting and beyond. Yes, it's just something that happened on social media, but as the past few years have shown, social media can be a powerful force.

Second, I met up with an old friend during the meeting, someone who I met at an undergraduate experience prior to entering graduate school. We are both on the tenure track now. When I first met him, he was not particularly interested in the problem of diversity in philosophy. Not that he was opposed to it; it just wasn't on his radar. He told me about a casual experiment he and his colleague had been conducting during the meeting. They were tracking data on how many men versus how many women

raised their hands during question periods, how often each gender was called on, and other information that would track patterns of behavior consistent with gender-based discrimination during question periods. The past decade had transformed my friend into someone who felt a pressing need to do something about the environment for underrepresented groups in philosophy of science. My friend and I are both still (hopefully) early in our careers, and the changes we have both undergone since beginning graduate school give me hope to imagine what we might find, and what we might create, at philosophy conferences thirty or forty years from now.

Although my term as a co-chair of the Women's Caucus has ended, I hope that as the Caucus continues to grow and evolve, it will make more contact with broader efforts to diversify philosophy in the APA as well. None of the initiatives discussed here were targeted solely at improving the climate for women in philosophy of science. The Women's Caucus has some of those initiatives too. However, the more obviously intersectional efforts discussed here have played a significant part in creating a PSA meeting with a climate that is more welcoming to a wider net, more generally. There are more women and more people of color coming to meetings, as well as more accommodations available for people with children and people with special needs. With a little luck and a lot of effort, and with the continued support of a governing board that has chosen to recognize and prioritize the need for coordinated efforts to improve diversity in the PSA, I can only hope the situation will continue to improve. It would be nice to leave the sweater at home someday.

BOOK REVIEWS

What Is Rape? Social Theory and Conceptual Analysis

Hilkje Charlotte Hänel (Bielefeld: Transcript-Verlag, 2018). 320 pp. \$100.00. ISBN 978-3837-64434-0.

Reviewed by Caleb Ward
 STONY BROOK UNIVERSITY (SUNY)

Over the past forty years, philosophical discussions of rape have focused on two sometimes-overlapping aims. With key insights about power and social structure, feminist social and political theorists (e.g., Susan Brownmiller, Carole Pateman, Claudia Card, Ann Cahill) have sought to produce more adequate accounts of the causes, social functions, and lived effects of rape and to articulate viable avenues for resistance and societal transformation. Meanwhile, moral and legal philosophers—both those who explicitly adopt feminist commitments (e.g., Susan Estrich, Lois Pineau, Keith Burgess-Jackson, Jean Hampton, Joan McGregor) and those who don't (e.g., David Archard, Alan Wertheimer, Donald Dripps)—have asked a different set of questions: What is the wrong of rape? How can legal definitions reflect this wrong? How ought culpability to be determined?

With *What Is Rape?* Hilkje Hänel enters the debate from a refreshingly new angle, leveraging contemporary critical theory and social ontology to analyze how the phenomenon of rape is conceptualized, upstream from the normative and definitional concerns of moral and legal theory. She expands the project of feminist structural critique to look beyond rape's causes and effects to the question of how the conceptualization of "rape" is itself shaped by power relations and socially available epistemic resources. Her careful analysis produces a convincing social theory of how rape fits into sexist ideology and a rigorous, constructive account of how rape ought to be conceptualized—not only to reflect the complexity of the phenomenon but also to account for "rape" as a concept susceptible to ideological distortion.

While the category of rape is widely contested in legal and popular discourses, Hänel diagnoses a common, often implicit working understanding of rape shaped by defective beliefs and judgments colloquially known as "rape myths." Rape myths are a particular kind of cultural narrative that "shape our understandings of sexual activities and sexualized violence," particularly in ways that "legitimize male entitlement to a female body" (35). Rape myths serve both explanatory and justificatory functions in shaping attitudes and interpretations of sex; in other words, they shape both how people interpret the factual events of a sexual encounter (e.g., "it wasn't really rape") and how they apportion blame for harms (e.g., "she/I was asking for it"). It is characteristic of rape myths that they are generally false but widely held either implicitly or explicitly to be true and that they circulate in both "everyday depictions and symbols" and by means of "everyday language practices" (44–45). Given their ready-to-hand cultural accessibility, Hänel argues that the distortionary effects of rape myths are especially great when they stand in as "indicator properties" to help us interpret situations where other evidence or information is lacking or where alternative explanations appear less salient.

Hänel argues that despite increasing institutional uptake of more adequate formal definitions of rape (replacing force with lack of consent, recognizing marital rape, etc.) the pervasive influence of rape myths fuels a dominant working understanding of rape that privileges aggravated stranger rape as the paradigm case, thereby failing to track the wide range of phenomena—date rape, acquaintance rape, war rape, rape in prisons, etc.—that ought to be included in the category. This descriptive failure of the concept has significant normative effects in that the inadequate conceptualization of rape hinders victims' abilities to make experiences intelligible to themselves and to others. Even when formal definitions of rape evolve, the influence of rape myths on the dominant working understanding can produce an "institutional mismatch" between such definitions and their application. Victims' experiences are read as "not really rape" if they don't match the dominant working understanding, which undermines victims' ability both to gain private support and to be believed by institutional gatekeepers of justice.

Hänel argues, following Sally Haslanger (2012), that the injustices stemming from the failure of the concept to

track reality can and ought to be ameliorated by enhancing the descriptive power of the concept in politically useful ways. However, rather than simply pursue a more refined definition of rape—the project of much feminist moral and legal theory—Hänel argues for a reconsideration of the conceptual architecture used to understand the phenomenon. To respond to the distortions of rape myths, she develops a methodology she calls *emancipatory amelioration*, which builds on Haslanger’s conceptual amelioration by formulating a concept of rape that is not only more descriptively adequate but also “fruitful for overcoming the de facto distorted conceptions of the concept and the equally problematic usage of the term” (24). This project requires two steps: (a) a systematic social theory that locates the phenomenon and its conceptualization within social structures, and (b) a reformulation of the concept to better capture the phenomenon and to become more resistant to the distortions that characterize the status quo.

Hänel’s social theory identifies rape as a social practice (in the technical sense) within Haslanger’s (2017) detailed conceptual map of social structures. With case studies and a careful attentiveness to debates in social philosophy, Hänel argues that rape can be described as an *accepted* social practice within the social structure because many of its forms are persistently misrepresented and justified according to available schemas (namely, rape myths), and because institutional responses distribute resources to perpetrators of rape (e.g., the ability to get away with it) while denying resources, including testimonial credibility, from victims.

Hänel adopts insights from well-established feminist critiques, but she also adds complexity to the classic argument that rape is a feature of patriarchal social control. She argues that the social practice of rape and the way it is popularly conceptualized cannot be understood without examining the sexist ideology—beliefs, attitudes, and practices—of which it is a part. Drawing omnivorously from Haslanger (2017), Barbara Fields (1990), Tommie Shelby (2003), and the Frankfurt School (Geuss 1981), she develops a nuanced view of how different social groups are susceptible to the epistemic distortions that mask rape’s injustice. Locating rape within an ideology has the benefit of acknowledging the many social functions of rape beyond the unidirectional domination of women by men (e.g., destruction of communities in war, enforcement of white supremacy), which helps explain why women are not the only targets of rape and men are not alone in reproducing rape myths. All social groups can be located within the ideology, and each plays a role in its perpetuation. Crucially, the ideology explanation of rape also lends itself to the tools of immanent critique, suggesting that contradictions emerge within the social structure to create possibilities for resistance from within.

How should rape be conceptualized given its vulnerability to distortion within a sexist ideology? Rather than settle the contested conceptualization of rape with a definition based on necessary and sufficient essential characteristics, Hänel proposes that rape be understood as a “cluster concept,” with disjunctively, rather than conjunctively, necessary

criteria. (She grounds the theory of cluster concepts with an extensive discussion of Wittgenstein’s (1997[1965]) concept of “family resemblance” as a tool for describing social phenomena.) The ten salient criteria she selects for the concept of rape are drawn from empirical observation and social theory and articulated through a wide range of case studies. They include sexual activity, violence, means of physical coercion, means of psychological coercion, lack of consent, context of social vulnerability, and others.

Hänel astutely recognizes, however, that these criteria are not binary features cleanly present or absent in particular cases. Rather, they are each time actualized in varying degrees, which motivates Hänel to add a further dimension of depth to her cluster concept. This is what she calls a “core” to the cluster, where an event that has many criteria actualized to the greatest degree is closer to the core—and therefore most clearly rape—while an event that has few or no criteria to a high degree and others to a low degree will be further from the core. In the latter case, an event may still appropriately fall under the category of rape, but it may also be more adequately included under a neighboring concept, such as another form of sexual mistreatment (e.g., sexual harassment) or a form of morally permissible sex (e.g., high-risk consensual sex).

Hänel’s approach brings several advantages. By examining the concept of rape apart from its juridical uses, she avoids drawing a sharp line between rape and not-rape, which in turn allows examination of how grey areas can be present not only within rape but also between rape and neighboring concepts. This appreciation for grey areas is also well-served by the variability of degree she builds into each criterion—an ambitious solution to the problem of conceptualizing the multifaceted character of rape, even if the complexity of the solution at times exceeds the metaphors offered to aid the reader’s visualization. This approach leaves open substantially more space for discussion and contestation of the concept of rape, allowing a more reflexive stance toward genuine ambiguity and toward shifts in intuitions, such as those that come about in response to survivor movements like #MeToo.

Hänel’s model views the “consent” criterion, for example, as always actualized in degrees, which enables the significance of an ambiguous expression of consent to be evaluated within a wider consideration of the features and context of an encounter. While standard philosophical discussions of consent acknowledge that context can affect consent’s “moral validity,” Hänel avoids defining rape according to a lack of morally valid consent alone. By taking into account a wider range of features within a given encounter, she seeks to weaken the potential power of rape myths over how consent is understood and evaluated. Here, however, her discussion would be strengthened and clarified by a more detailed explanation of how her cluster account departs from standard definitions of rape according to valid consent. Hänel allows that the criteria she identifies might be present in varying degrees in any particular rape, but it is not self-evident how she would respond to the argument that some criteria are simply more important to determining whether something is rape—namely, non-consent and sexual penetration of any kind.

As methodology, Hänel's approach to rape as a normative core-cluster concept can be applied to other social phenomena characterized by grey areas, ideological obscurations, and interference from neighboring concepts. (In footnotes she draws interesting parallels with critical analyses of race and racism.) She usefully reminds us of the need for concepts to accommodate shifting social realities, such as changing cultural norms and practices, and to allow productive contestation to continue. While legal scholars have work to do to figure out whether this flexibility can influence the law—Hänel argues that it can indirectly—critical theory and activist practice will benefit immediately from the move to decouple a concept's adequacy from its authority to resolve all contestation.

The book ends with a return to common normative questions, exploring the implications of Hänel's account for holding perpetrators responsible and for enacting solidarity with accusers. Specifically, Hänel considers the problems posed by a "cognitive deficit" (i.e., ignorance-based) interpretation of the ideological distortions surrounding rape. If ideological distortion is straightforwardly a category of ignorance, locating rape within a pervasive ideology might be taken to give perpetrators an excuse for their actions (i.e., because they act out of ignorance) and to undermine the epistemic authority of victims' testimony. Hänel takes a generous approach to perpetrator ignorance, gesturing toward a model of accountability and restorative justice that can "adequately and productively confront ideological beliefs and result in a learning process," even if it softens blame toward perpetrators whose social positions within the ideology (e.g., as cis-males) encourage ignorance as to the harmfulness of their actions (205). Moral philosophers steeped in debates about the nature of responsibility will probably have qualms with Hänel's specific disaggregation of blameworthiness from accountability, but they would do well to remember that her proposal is political—what is to be done?—rather than a metaethical claim about moral responsibility as such.

For responding to victims, Hänel prescribes solidarity that affirms the validity of victims' subjective experiences and recognizes that, due to their position within a sexist ideology, they are likely to suffer hermeneutical injustice as well as testimonial injustice (Fricker 2007). However, she stops short of attributing to victims epistemic privilege to wholly define the truth of what took place and its appropriate moral interpretation. Her position is that victims ought to be recognized as full epistemic authorities on their experiences of harm, but the questions of what objectively happened and whether an encounter should be defined as rape often require further interpretation. This argument is a close neighbor to Linda Martín Alcoff's argument in *Rape and Resistance* (2018)—published when *What Is Rape?* had already gone to press—and it could be revisited and strengthened in light of Alcoff's insights. Hänel's argument would benefit from Alcoff's point that experiences must also be interpreted—empowerment and solidarity require enhancing survivors' epistemic resources—and from Alcoff's claim that we can recognize epistemic *privilege* of survivors without granting absolute epistemic *authority*. On balance, Hänel's prescriptions in the final chapter reflect a laudable commitment to hold the emancipatory aspect of

her project accountable to feminist activism, but her claims will require a more sustained treatment—perhaps in her next book?—to satisfy both philosophers and activists who focus on these particular issues.

The only true weak point in this ambitious book is the introduction, which doesn't do justice to the breadth or depth of the project and could do more to set up the analysis that follows. In particular, the introduction could have provided more transparency regarding the operative concept of rape that undergirds Hänel's conceptual critique; the basic commitments she names are only peripherally connected to the primary argument of the first two chapters. Moving some of the methodological points forward from chapter three could have grounded the critique of standard conceptualizations of rape, making the argument more convincing by acknowledging that the author's own social position shapes her critical intuitions about the concept (Cf. Alcoff 2018).

Still, Hänel's arguments throughout *What Is Rape?* are characterized by fine-grained distinctions and careful attention to how concepts are produced and used in social reality. Her writing is paced for clarity and marked by abundant signposting. Perhaps the most widely accessible contribution for a general audience will be Hänel's insightful account of the nature and function of rape myths. However, specialists in social philosophy will also benefit from her methodological innovation for analyzing ambiguous and complex social phenomena. The nuance and flexibility of her approach to the topic of rape is welcome, and its importance will only increase as popular intuitions about rape continue to shift and as prevailing sexual values and practices across societies continue to be questioned.

REFERENCES

- Alcoff, Linda Martín. *Rape and Resistance*. Cambridge: Polity Press, 2018.
- Fields, Barbara. "Slavery, Race and Ideology in the United States of America." *New Left Review*, no. 181 (1990): 95–118.
- Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press, 2007.
- Geuss, Raymond. *The Idea of Critical Theory: Habermas and the Frankfurt School*. Cambridge: Cambridge University Press, 1981.
- Haslanger, Sally. *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press, 2012.
- . *Critical Theory and Social Practice*. Spinoza Lectures. Assen: Koninklijke Van Gorcum, 2017.
- Shelby, Tommie. "Ideology, Racism, and Critical Social Theory." *Philosophical Forum* 34, no. 2 (2003): 153–188.
- Wittgenstein, Ludwig. *The Blue and Brown Books*. Oxford: Blackwell, 1997[1965].

Debating Pornography

A. Altman and L. Watson (Oxford University Press, 2018).

Reviewed by Mari Mikkola

UNIVERSITY OF OXFORD AND SOMERVILLE COLLEGE

This book is part of OUP's *Debating Ethics*-Series, where authors defend opposing views on pressing ethical,

political, and legal issues. As the title suggests, this edition is about pornography. Andrew Altman first offers a defense of the right to buy, sell, and use pornography made with consenting adults on the basis of adults' right to sexual autonomy. Lori Watson provides the opposing view: she advances a sex equality argument for the view that pornography ought to be restricted under civil rights law. Both Altman and Watson hold that in order to make a reasonable case for legally restricting pornography, we must demonstrate how pornography threatens and undermines women's equal standing as citizens. However, they disagree on whether pornography systematically does so: Altman takes the empirical evidence to establish this to be lacking, while Watson holds that the evidence for pornography systemically interfering with equality is overwhelming. Both parts of the book (around 150 pages each) offer ample and detailed background to the relevant legal and ethical matters, and readers without a robust background in these debates will gain excellent introductions to them. The book is written lucidly with wonderful clarity, and yet without compromising philosophical rigour and depth. It is a rich work that offers much to both beginners and seasoned experts in the debates. Both Altman and Watson discuss pornography philosophically with real-world considerations in mind, drawing on legal cases along with psychological and social scientific research on pornography's effects on consumers and the wider society. This is deeply compelling and (I maintain) precisely the correct methodology when thinking about whether and how to restrict pornography legally.

Altman starts his contribution (Chapter 1) by discussing the common view that legally regulating pornography hinges on free speech concerns. He rejects the view that pornography's permissibility is about freedom of expression; rather, governmental efforts to restrict pornographic materials is an issue about sexual autonomy understood as "a moral right of adult individuals" (17). Legal restrictions that block access to "material that graphically depicts sex" (17) violate that right, though legal regulations as to where and how much materials may publicly be displayed and advertised do not. (This highlights a frequent misunderstanding about the liberal pornography position: that they promote an "anything goes" laissez-faire attitude to pornography.)

Altman's view presupposes that sexual autonomy is a basic liberty that is not on a par with all other liberties; it is part of "a small group of liberties that carry special weight" (29). This being the case, in order for (adult) pornography restrictions to be justified, its proponents must make good on the claim that pornography causes harm to women at a social, rather than an individual level. Altman is convinced that pornography production and consumption has harmed individual women, but finds the evidence on the social level "very weak" (24). Moreover, Altman continues, even if there were strong evidence of harms caused on the social macro-level, this still would not *eo ipso* justify pornography restrictions. He draws a parallel between alcohol and pornography consumption: the social-level evidence that alcohol consumption causes violence (both sexual and non-sexual) against women is strong and apparently "much stronger than the evidence that pornography consumption

causes violence against women" (25). Nonetheless, there are no legal restrictions to block the production, purchase, and consumption of alcohol for adults—something they have a right to since drinking is "an exercise of personal liberty" (25). However, pornography use isn't just "an exercise of personal liberty" but the exercise of our freedom to shape our lives as sexual beings (25)—it is an exercise of our *basic* liberty to sexual autonomy. This renders the right to pornography even stronger than the right to consume alcohol (the former being grounded in a morally weightier liberty than the latter). Hence if adults have a right to drink, they have a right to pornography consumption. Altman doesn't think that the right to pornography is absolute. But given the strength of this right as a matter of sexual autonomy, he holds, arguments to restrict pornography on the basis of societal harms must be on very sturdy grounds. And the empirical case for these harms is allegedly too weak to do the work. That is, although our basic liberty to sexual autonomy doesn't generate permissions to exercise that liberty in whatever way we want, unless anti-pornography advocates can demonstrate that pornography production and consumption generates systematic harms to women, the right to pornography remains intact since it is an aspect of the right to sexual autonomy. What supports the view that the right to pornography is part of our right to sexual autonomy is that pornography is used for sex (section 2.4.). Pornography consumption then, for Altman, is a kind of sex.

Let me pause for a moment to comment on this part of Altman's argument. I am happy to accept that we have a basic liberty to sexual autonomy. (Although both Watson and myself find Altman's explication of this curious: for him, lacking control of one's sex life signals a degraded social status "because it means that others have social permissions to spill their disgusting sticky substances onto or into you without your consent" (45).) I am less convinced, though, that there is a right to pornography on the grounds Altman cites (pornography is used for sex). Many things are used for sex: vibrators, dildos, lube, etc. Their exercise is also an expression of our sexual autonomy. Following Altman, then, we seemingly have a right to dildos, lubricant, and a whole array of sex toys. I suspect that many people think this generates rights too cheaply and we should not dilute rights-talk to such an extent. And if the right to pornography is "cheap" in this way, it is doing little work apart from serving a rhetorical purpose in being provocative insofar as the talk of "a right to pornography" is typically taken to connote something stronger and to engender different expectations.

Moreover, although it is certainly a common view that pornography is used as "masturbation materials," empirical research suggests (perhaps surprisingly) that pornography consumption is grounded in a variety of motives, where being used for sex is one among many. Men consume pornography for (e.g.) entertainment, escape, or release.¹ Many young male consumers self-report consuming pornography as an educational guide (see, e.g., the 2013 report from the UK Office of the Children's Commissioner). Because pornography is not always used for sex but apparently sometimes just to pass the time or to find out about sex, I am not convinced it should be viewed as a kind of sex. Of course, Altman might retort and claim there

is a difference between *viewing* pornography (which is done out of various reasons and isn't a kind of sex) and *using* pornography (where pornography just is a sexual surrogate). This, however, seems to presuppose that what it is to use or consume pornography is to view it while masturbating. The idea that pornography use is a kind of sex, I contend, now becomes even less convincing: the kind of sex going on is masturbation—pornography just happens to be there in the background.

In Chapter 3, Altman considers empirical evidence for harm that pornography supposedly does to women's equal citizenship. He accepts that this evidence establishes correlations between pornography use and "aggressive behaviors and attitudes toward women" (85). Still, causal connections remain to be established. Altman goes through various studies that allegedly fail to establish harm at a population level. In Chapter 4, Altman discusses and rejects the famous MacKinnon-Dworkin ordinance against pornography on the grounds that the empirical evidence of systematic harm is lacking (I will discuss the ordinance shortly since Watson's contribution defends it). The final substantive chapter, Chapter 5, briefly rejects the idea of pornography as subordinating speech. A well-known anti-pornography position (championed by Rae Langton in addition to MacKinnon and Dworkin) is that pornography does not merely depict subordination; it is women's subordination. This is because pornography is said to "speak authoritatively"—a claim Altman is skeptical of. His concerns, on the one hand, turn on the idea that pornography says anything at all and, on the other, that pornography is authoritative. He accepts that some men take pornography to be authoritative, but goes on to claim that these men can hardly be considered "representative of men in society generally" (129). By way of evidence, Altman offers a personal anecdote of watching a pornographic film as a naïve college student and *even then* thinking it was not authoritative about sex insofar as the film was utterly ludicrous. He further notes that he would be very surprised if anyone in the audience thought any differently. This isn't the only part where Altman appeals to anecdotal personal experiences as a pornography consumer. In section 2.3, he discusses the feminist anti-pornography claim that pornography eroticizes inequality. Altman rejects this claim, holding instead that pornography's consumers drawn to materials that do so must already find inequality and dominance/submission sexy. If they did not find such depictions arousing, they wouldn't seek to view them: "If men do not find sex as dominance arousing to begin with, then I do not see how exposure to pornography makes such sex arousing for them" (72). Again, he offers evidence for this from his own pornography consumption.

Altman's discussion is admirably frank, and he notes that perhaps his experiences have been idiosyncratic. Still, the methodological problems here run deep. For one thing, plenty of other pornography consumers self-report having initially found some deeply inegalitarian pornography disgusting and disturbing, having only viewed such materials originally out of curiosity and to see what all the fuss is about. (I have seen this claim made in print and have heard it in conversation with pornography-consuming men.) Upon longer exposure, they report coming to enjoy

such materials. There is, of course, much that Altman can say against this: we can, for instance, speculate that those men unconsciously harbored inegalitarian thoughts to begin with, which predisposed them to finding such materials ultimately arousing. Be that as it may, to hold that male sexuality works in the way Altman suggests or that there is such a thing as uniform "male sexuality" to begin with is highly suspect in my view. Furthermore, the appeal to personal anecdotes and intuitive speculation leaves me with little grounds to find Altman's case convincing. Watson also critiques the idea that pornography use doesn't affect consumers' psychologies: given the robust empirical evidence that media affects our beliefs, attitudes, and behaviors (just think of advertisements), it is strange to think that pornography doesn't affect consumers. "How is it," Watson asks, "that somehow pornography, almost alone, is an exception to the general knowledge about how images and words impact behavior?" (232) According to her, both common sense and empirical evidence run counter to the idea that pornography has little effect on people's behavior.

In her contribution, Lori Watson takes a sex equality approach to pornography, whereby pornography is "a key social location of the subordination of women to men" (151). She takes evidence for this claim to come from scientific empirical studies of pornography's harms, testimony of women and girls harmed through pornography, and from seeing how pornography eroticizes inequality and affects women's social status. As part of her approach (see Chapter 7), Watson explains and defends MacKinnon and Dworkin's civil rights ordinances by way of legally regulating pornography's production and consumption. Note at the outset that legally regulating something isn't equivalent to banning or censoring it—nor does Watson advocate legislation that makes pornography production and consumption a criminal offense or subject to censorship. Rather, certain materials are civilly actionable provided that they fulfill two conditions. First, they fall under a narrow conception of *pornography* as

the graphic sexually explicit subordination of women through pictures and words that also includes women dehumanized as sexual objects, things, or commodities; enjoying pain or humiliation or rape; being tied up, cut up, mutilated, bruised, or physically hurt; in postures of sexual submission or servility or display; reduced to body parts, penetrated by objects or animals, or presented in scenarios of degradation, injury, torture; shown as filthy or inferior; bleeding, bruised, or hurt in a context that makes these conditions sexual.²

Second, in addition to meeting this definition of *pornography*, a harm specified in the ordinances must be proven before a judge, jury, or other legal administrative body. These include coercion into a pornographic performance, forcing pornography on a person, assault or physical attack due to pornography, and discrimination by trafficking in pornography. In the first part of her contribution, Watson helpfully clarifies MacKinnon and Dworkin's ordinances against common misconceptions, which is both valuable and overdue. In other words, *contra* critics, some work isn't

legally actionable just in satisfying the first condition (the definition of *pornography*). Demonstration of harm is also necessary. And Watson holds that the empirical evidence of pornography's harms is overwhelming.

In Chapter 8, Watson goes on to discuss how the pornography industry harms. She maintains that various pornographic genres not only eroticize gender inequality but also racial, trans-identity-based, age, queer, and familial inequalities insofar as "the 'sexiness' is in the inequalities, the hierarchy between the paired groups" (213). Moreover, she takes there to be production and consumption harms well-documented empirically. For one thing, evidence suggests that pornography consumption increases the acceptance of rape myths and callous attitudes towards women. Hence:

Insofar as consumers of pornography occupy the positions of police investigator, judge, and jury member, and consumption of certain forms of pornography . . . is demonstrably shown to increase acceptance of rape myths and the acceptance of interpersonal violence, then pornography is implicated in the systematic denial of women's equal civil rights to have a fair and impartial investigation and adjudication of crimes against them. (225)

Watson then discusses the evidential strength of social-level studies. Contra Altman, she thinks that the evidence is strong. One extremely important point Watson raises centers on the claim that pornography consumption is causally connected to the occurrence of rape. This is something that critics of anti-pornography feminism and proponents of a right to pornography typically home in on: Since there is no conclusive proof of such a causal connection, the anti-pornography position is a nonstarter. Watson rightly critiques this view. Of course feminists are not merely concerned with grave sexualized violence, and there are many harms that anti-pornography feminism has identified. Once we cast the net more widely and look at harms that are not merely to do with rape, the empirical evidence paints a rather different picture to that held by critics of anti-pornography feminism.

Nonetheless, I am not convinced that Watson's treatment of the empirical evidence persuades critics. The debate between her and Altman follows familiar lines: both sides cite more or less the same evidence, where one side finds mere correlation while the other finds clear causation. This is one way I take pornography debates in philosophy to be intractable: our interpretation of the available evidence turns too much on our prior normative views to produce genuinely fruitful results. Bluntly put, interlocutors in pornography debates tend to find evidence in empirical research for the views that they antecedently hold. I am skeptical that the moral or legal case for or against pornography can be settled by looking at allegedly systematic harms to women, even though the evidence cited by both Altman and Watson is illuminating and important. It is also striking that in the quote above Watson claims pornography to be "implicated in the systematic denial of women's equal civil rights." At the start of her

contribution, she claims pornography to be "a key social location of the subordination of women to men" (151). The idea that pornography is implicated in gender subordination is much weaker and something I suspect Altman would not reject. He doesn't think that pornography is harmless; he seems rather to hold that evidence shows pornography not to be a key social location of subordination and hence for our right to sexual autonomy (of which pornography use is supposedly an aspect) to be overriding when thinking about pornography's legal regulation.

Chapter 9 charts out and ultimately rejects various defenses of pornography. These are the free speech defense, the feminist pro-porn position, and the argument from sexual autonomy. Watson convincingly (I hold) critiques the first and third defenses. However, I found her treatment of the second defense less compelling. She starts by focusing on the supposed feminist attacks of the MacKinnon-Dworkin ordinances. At the time (1980s), Feminists Against Censorship Taskforce in the United States prominently campaigned against anti-pornography positions. Watson discusses FACT's opposition to the ordinances and rightly points out that it grossly misrepresented the ordinances and the available empirical evidence. I am in utter agreement with Watson that this and some other supposedly feminist defenses of pornography in the 1980s-90s fail. However, Watson also rejects newer feminist pornography positions. In fact, she doesn't appear to accept that feminism can genuinely be compatible with a pro-pornography stance. One of the most celebrated self-proclaimed feminist pornography books is *The Feminist Porn Book* (2013). According to self-proclaimed feminist pornographers, the central features of feminist and mainstream pornography come apart in terms of what is represented and depicted, and how production is organized. Feminist pornography is said to be a genre that

uses sexually explicit imagery to contest and complicate dominant representations of gender, sexuality, race, ethnicity, class, ability, age, body type, and other identity markers. It explores concepts of desire, agency, power, beauty, and pleasure at their most confounding and difficult, including pleasure within and across inequality, in the face of injustice, and against the limits of gender hierarchy and both heteronormativity and homonormativity.³

Watson, however, rejects that this vision of pornography is genuinely emancipatory. She holds that just because some natal females take up and enact dominant sexual roles, this isn't emancipatory for them or women in general. On her approach, "relationships of domination and subordination, including sexually scripted roles in which someone is the dominator and another the subordinate, do not constitute liberation for anyone, especially the people at the bottom of the hierarchy. They remain hierarchical, hence unequal" (265). Supposed feminist pornography retains the sort of inequality and lack of freedom that serves women's subordination, hence not being genuinely feminist.

I wholeheartedly agree with Watson that not all self-proclaimed feminist pornography *therefore* counts as

feminist. But I find her treatment of contemporary (at least putative) feminist pornography to be somewhat uncharitable and too quick.⁴ Moreover, I am troubled by Watson's suggestion that any hierarchies found in sexuality and sexual scripts is *eo ipso* subordinating. Underlying Watson's discussion are seemingly strong normative views about sexuality and what constitutes "good" sex. For instance, she holds that pornography undermines sexual autonomy in that it "limits, undermines, and restricts the development of authentic sexuality, for both women and men, and interferes with interpersonal intimacy" (287). Although she doesn't draw on this, Watson's views reflect those of Gail Dines, a well-known anti-pornography activist and writer, whose work Watson elsewhere cites approvingly. Dines claims that in "pornland"

an authentic sexuality—one that develops organically out of life experiences, one's peer group, personality traits, family and community affiliations—is replaced by generic porn sexuality limited in creativity and lacking any sense of love, respect, or connection to another human being.⁵

Given what we know about coercive mechanisms that families, communities, and even the law exercise to make people conform to heteronormative sexual expectations, I see little reason to believe that an authentic sexuality would "organically" develop in the course of sexual maturation without pornography. Furthermore, Watson's view reflects something else Dines writes: "Missing from porn is anything that looks or feels remotely like intimacy and connection, the two ingredients that make sex interesting and exciting in the real world."⁶ I suspect that not everyone agrees with Watson and Dines about the role of intimacy and what makes sex interesting in the real world—and (I contend) it is possible to disagree without being in thrall to a sexist sexual ideology. Contra Watson, I think that there is more to be said in favor of some pornography and that pornography can be a genuine force for liberation. Of course, Watson can retort noting that she is merely interested in materials that fall under the narrow conception of *pornography* and sexually explicit liberatory materials don't count as such. But then her discussion of the above putative feminist pornography looks puzzling. In order for this discussion to be intelligible, Watson must think that these sorts of materials fall under the narrow conception she is operating with. This move, I contend, is not obviously warranted, though.

Debates about what we should legally do about pornography (if anything at all) are at an impasse and, despite the admirable cases both Altman and Watson make for their views, I remain unconvinced that the book resolves this impasse—or even comes close to resolving it. But I do not see this as a weakness. Rather, the richness of the book will (I am sure) instigate many further discussions that are both interesting and sorely in need of taking place.

NOTES

1. Matthew B. Ezzell, "Pornography Makes a Man: The Impact of Pornography as a Component of Gender and Sexual Socialization."
2. Catharine A. MacKinnon, *Feminism Unmodified*, 176.

3. Taormino et al., *The Feminist Porn Book: The Politics of Producing Pleasure*, 9–10.
4. For more, see chapter 7 of my *Pornography: A Philosophical Introduction*.
5. Gail Dines, *Pornland: How Porn has Hijacked our Sexuality*, xi.
6. *Ibid.*, 68.

REFERENCES

Dines, G. *Pornland: How Porn has Hijacked our Sexuality*. Boston: Beacon Press, 2011.

Ezzell, Matthew B. "Pornography Makes a Man: The Impact of Pornography as a Component of Gender and Sexual Socialization." In *The Philosophy of Pornography: Contemporary Perspectives*, edited by Lindsay Coleman and Jacob M. Held, 17–34. Lanham: Rowman and Littlefield, 2014.

MacKinnon, Catharine A. *Feminism Unmodified*. Cambridge, MA: Harvard University Press, 1987.

Mikkola, M. *Pornography: A Philosophical Introduction*. NY: Oxford University Press, 2019.

Taormino, Tristan, Celine Parreñas Shimizu, Constance Penley, and Mireille Miller-Young, eds. *The Feminist Porn Book: The Politics of Producing Pleasure*. New York: The Feminist Press at CUNY, 2013.

Foucault and Feminist Philosophy of Disability

Shelley L. Tremain (Ann Arbor, MI: University of Michigan Press, 2017).

Reviewed by Catherine Clune-Taylor
SAN DIEGO STATE UNIVERSITY

In 2002, Rosemarie Garland-Thomson called for the integration of disability as a "category of analysis and a system of representation" into feminist theory, arguing that such work would generatively enrich and transform both feminist theory and disability studies. Further, it would clearly demarcate a vibrant field of academic study that was already producing critical interventions with regards to both—feminist disability studies.¹ In 2017's *Foucault and Feminist Philosophy of Disability*, Shelley L. Tremain performs the urgent work of providing (and motivating) a clear feminist philosophy of disability to theoretically underwrite such a field, while making clear the "transgressive potential" of her account. The book builds upon Tremain's Foucauldian account of the constitution of impairment and disability first introduced in 2001's "On the Government of Disability" in which the author followed Butler's (1990) lead in challenging both the sex/gender binary—and the nature/culture binary onto which it mapped—through the revelation of sex as "gender all along." With her (2001) excavation of impairment as, similarly, "disability all along," Tremain provided the foundation for critical disability studies as a field of analysis characterized by a rejection of impairment as a natural fact of the body. This was the first critical alternative to the British Social Model and its assertion of an impairment/disability binary drawn along nature/culture lines.² In *Foucault and Feminist Philosophy of Disability*, we are provided a full articulation of Tremain's feminist account of disability as "an apparatus of productive force relations rather than as a personal characteristic, an identity, a difference, or a form of social oppression," and

of impairment as the naturalized effect and mechanism of that apparatus, brought about through the historical emergence of a “diagnostic style of reasoning” previously available only in piecemeal (204).³ Near the beginning of her first chapter, Tremain states that her book is a call for a “conceptual revolution” (borrowing Rae Langton’s term)—that is, a “critical ontology of what philosophers think disability is, of what they think about how disability is produced, and of what their current thinking about disability does” (3). While the call she issues in the form of arguments regarding the historically and culturally specific constitution of both impairment and disability and of the emergence of the academic discipline of bioethics as a “technology of government” may have particular import for feminist philosophers, philosophers of disability, social epistemologists, and bioethicists, it aims at the reform of philosophy as a discipline in general.

The book’s arguments span what Tremain calls “two distinct, but interrelated and mutually constitutive spheres,” outlined within the first chapter of the book, titled “Groundwork for A Feminist Philosophy of Disability.” The first is a “reconstructive-conceptual sphere” in which she provides a historicist and relativist account of the constitution of both disability *and* impairment as a corrective to what she identifies as the “standard philosophical view” pervasive throughout the field that disability is a natural, pre-discursive, objectively disadvantageous human trait that should not only be prevented or corrected, but is also philosophically uninteresting (2). The second is a “metaphilosophical sphere” which connects the exclusion of disabled philosophers from the field to this standard view. As Tremain writes, this standard philosophical view regarding disability which she dispels, and the “underrepresentation of disabled philosophers within the profession are inseparably embedded in the institutional infrastructure of the discipline, mutually constitutive and mutually reinforcing” (2–3). And disabled philosophers are—indeed—exceptional in their marginalization from professional philosophy—a discipline (in)famous for its overwhelming homogeneity. While women make up about 25 percent of full-time philosophy faculty in the United States and 30–35 percent of them in Canada, philosophers of color are estimated to make up less than 10 percent of that population in North America, and Black philosophers only about 2 percent (29). Despite disabled individuals making up 22 percent of the general population, recent surveys indicate disabled philosophers only make up about 2 percent of faculty members in the United States and less than 1% in Canada. Professionally, then, philosophy is—as Tremain writes—“made up almost entirely of nondisabled white people” (29).

The underrepresentation of disabled folks in philosophy can be at least partially attributed to the employment discrimination these individuals are subject to within the academy, and at large, for which Tremain provides compelling empirical evidence. However, she argues that the “grievous unemployment and poverty that accrue to disabled people” are the result of the apparatus of disability itself, which naturalizes impairment and (among other things), renders higher education inaccessible to many disabled folks by relying upon an “individualized, medicalized,

and privatized conception of disability,” that positions their requirements for flourishing as “special needs” (30). Tremain notes that this “accommodationist” understanding of disability is echoed within the philosophical canon, reminiscent—as it is—of Rawls’s “set[ting] aside” disabled folks’ entitlement claims to be only considered after those in the original position arrived at their blueprint for a just society for “normal and cooperating” people (31).⁴ This conception of disability, which underwrites and supports the standard philosophical understanding of disability (and impairment) as philosophically uninteresting, positions valid engagements with disability as those which properly construe it as a primarily biomedical issue. This, in turn, legitimates sequestering considerations of disability in (always already) devalued areas of “applied” philosophy, and nonideal theoretical fields like applied ethics, or as subfields of bioethics and cognitive science at best, in ways that disincentivize specialization as well as hiring.

This constitutive (mis)construal of disability—for description is never neutral, but rather always both constitutive and prescriptive—is reified by and reflected within professional output on the topic. As Tremain notes, most feminists philosophers who take up disability and most philosophers of disability “do not rigorously question the metaphysical and epistemological status of disability, but rather advance ethical and political positions that largely assume the self-evidence of that status” (2). This is despite the work of many feminist scholars—particularly feminist philosophers of science and social epistemologists—stressing the importance of demographics and the sociopolitical nature of knowledge production, pointing to, for example, the constitutive nature of categorization, the sociopolitical effects of speech acts, and the need to explore identity intersectionally, etc. Tremain introduces the term “ableist exceptionism” to refer to the phenomenon persistent throughout the discipline, whereby disability, “because it is assumed to be a pre-discursive, natural, and politically neutral human characteristic . . . is uniquely excluded from the production and application of certain values, beliefs, principles, and actions that circulate in a political consciousness” (33). Evidence of ableist exceptionism can be seen throughout the profession: in the quotidian use (and defense) of ableist phrases (e.g., “blind review,” or “blind to the consequences”); in the theoretical reliance on ableist concepts and metaphors (“moral blindness,” “justice is blind,” and in the continued failure to include disability in analysis which otherwise admit (and often specifically study) the constitution of identity-based oppression.

Ableist exceptionism is perhaps most striking when reproduced in theorizing which aims specifically at highlighting the sociopolitical nature of knowledge production and its oppressive forces. For example, the lacuna in knowledge production that results from the failure of professional philosophers to dispense with the standard philosophical view of disability, and thus, to recognize and engage with disability as an object (and appropriate subject) of epistemic and metaphysical investigation, arguably constitutes a form of hermeneutical injustice on Fricker’s increasingly popular account of epistemic injustice.⁵ However, Tremain highlights the ableism within Fricker’s account itself, insofar as it reproduces the standard

philosophical understanding of disability as external to power relations (38–40). Indeed, Fricker uses disability (i.e., a medical condition that is largely undiagnosed and misunderstood at a particular historical moment) as an example to differentiate hermeneutical disadvantages that result in epistemic injustice from those which are merely bad luck. Tremain notes how striking it is that Fricker—like many feminist philosophers keenly aware of (and seeking to address) the political nature of knowledge production and its role in oppression—reproduces the same hermeneutical injustice she is attempting to articulate regarding disability in her articulation of it. If women experiencing postpartum depression are subject to hermeneutical injustice regarding their situation because of unequal social conditions “conducive to the relevant hermeneutical lacuna,” such that they are likely members of a “hermeneutically disadvantaged group,” then the same is true for those folks with undiagnosed and misunderstood medical conditions she invokes.⁶ Background conditions similarly shape “the public perceptions and authoritative epistemologies from which the negative social, political, interpersonal and economic consequences” Fricker describes as accruing to these folks, argues Tremain, in ways that “natural[ize], medical[ize], and depolitic[ize] these perceptions and epistemologies in ways that conceal their contingent and artifactual character” (40).

Tremain’s contribution in *Foucault and Feminist Philosophy of Disability*, then, is a politically astute, culturally specific, and historically situated account of the constitution of both impairment and disability of the type that is essential to those currently theorizing (and challenging) systems of gendered and racial oppression. Indeed, it is very clear why the book won the 2016 Tobin Siebers Prize for Disability Studies in the Humanities while in manuscript form. In providing it, she successfully undermines not only the standard philosophical understanding of disability, but also the multitude of theories and practices within the profession upon which it rests. Her identification of the ways in which the standard philosophical account of disability works to not only constrain philosophy as a *discipline* (that is, as a field of knowledge/power), but also to contribute to the oppression of disabled individuals beyond the boundaries of the professional field is of import to every practitioner in the field. Indeed, while efforts to address the overwhelming whiteness and cisgender maleness of professional philosophy have been inconsistently successful for at least some marginalized folks, disabled folks have thus far been left out of such efforts almost completely, in a manner that is, indeed, exceptional (and often mirrors the culture at large). In the same way that combating gendered and racialized oppression within the academy and beyond requires representative accounts of the unique emergence and functioning of its many forms (e.g., misogyny, transphobia, misogynoir, etc.), so too does social justice for disabled folks and for the treatment of disability require the kind of “historicist and relativist” account of disability Tremain provides. In outlining a new metaphysics of disability, Tremain reveals the epistemic, sociopolitical, and bioethical stakes of disability, both in itself, and as an object of analysis.

Thus, I would argue that this book is a necessary intervention for every professional philosopher, given that the overwhelming homogeneity of the field is, arguably, a problem for all of its members—both politically and epistemically. However, beyond those working in disability feminist theory, social epistemology, and bioethics, Tremain’s work in this book will be of particular interest to Foucault scholars as well. I would argue that Tremain is one of the most incisive and careful readers of Foucault in the field and her skill is on display throughout the book.

After briefly outlining her two spheres of argument in Chapter One, Tremain moves on in Chapter Two (“Power and Normalization”) to review Foucault’s insights regarding the historical emergence of biopower and the constitutive effects of normalization in order to “flesh out and fill in the sketch of this historically and culturally specific apparatus of power relations that effectively brings disability . . . into being as a problem”—the rigor and care she takes in her analysis is evident (47). In Chapters Three (“Historicizing and Relativizing Philosophy of Disability”) and Four (“Foucault, Feminism, Disability”), she takes up critiques regarding the liberatory potential of a Foucauldian approach to disability, feminist theory, and activism respectively, rebutting each and articulating the critical potential for each in turn. Her treatment in dispelling each of these claims is masterful, and her analysis of Foucault’s controversial treatment of Charles Jouy in particular moves the conversation regarding this case among feminist Foucauldians forward in important ways. Finally, the arguments presented in Chapter Five (“Bioethics as a Technology of Government”) have urgent implications for anyone in the field. In it, Tremain charts the emergence of bioethics as an academic discipline which “rationalizes the eugenics of a (neo)liberal governmentality,” making sense of (and further grounding for) feminist concerns regarding the use of atomistic notions of autonomy, and tendency to discount disabled folks’ assessments of their own quality of life (45).

The work as a whole, as well as specific arguments in it, feels uncannily timely within not only our larger biopolitical context, but also within the academy, where time-worn debates regarding social justice have reignited in slightly new forms. For example, Tremain’s arguments against using “people-first” language (“people with disabilities”) in favor of “disabled people” in order to emphasize the constitutive nature of power seem to me in line with an emerging trend among intersectional scholars to shift away from thinking about the identities that power constitutes to thinking about power itself in its technologies, its administration—that is, in the multitude of quotidian ways in which certain populations are “made live” while others are “let die.”⁷ Further, the critical feminist account of disability the book lays out strikes me as an articulation of the theoretical core of recent feminist work attempting “decolonize disability,” by situating the constitution of impairment and disability within the functions of colonial power and transnational politics, as well as global capitalism and its racialized ecological effects.

NOTES

1. Garland-Thomson, "Integrating Disability, Transforming Feminist Theory," 3.
2. Tremain, "On the Government of Disability," 632.
3. Tremain "This Is What a Historicist and Relativist Feminist Philosophy of Disability Looks Like"; Tramaïn, "The Biopolitics of Bioethics and Disability."
4. Rawls, *A Theory of Justice*.
5. Fricker, *Epistemic Injustice: Power and the Ethics of Knowing*.
6. *Ibid.*, 152–53.
7. Latham, "(Re)Making Sex: A Praxiography of the Gender Clinic"; Spade, *Normal Life: Administrative Violence, Critical Trans Politics, and the Limits of Law*.

REFERENCES

- Butler, Judith. *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge, 1990.
- Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford ; New York: Oxford University Press, 2007.
- Garland-Thomson, Rosemarie. "Integrating Disability, Transforming Feminist Theory." *NWSA Journal* 14, no. 3 (2002): 1–32.
- Latham, J. R. "(Re)Making Sex: A Praxiography of the Gender Clinic." *Feminist Theory* 18, no. 2 (2017): 177–204. <https://doi.org/10.1177/1464700117700051>.
- Puar, Jasbir K. *The Right to Maim: Debility, Capacity, Disability*. Duke University Press, 2017.
- Rawls, John. *A Theory of Justice*. Harvard University Press, 1971.
- Spade, Dean. *Normal Life: Administrative Violence, Critical Trans Politics, and the Limits of Law*. Duke University Press, 2015.
- Tremain, Shelley. "On the Government of Disability." *Social Theory and Practice* 27, no. 4 (2001): 617–36.
- . "The Biopolitics of Bioethics and Disability." *Journal of Bioethical Inquiry* 5, nos. 2–3 (2008): 101–6. <https://doi.org/10.1007/s11673-008-9105-6>.
- . "This Is What a Historicist and Relativist Feminist Philosophy of Disability Looks Like." *Foucault Studies* 19 (June 2015): 7. <https://doi.org/10.22439/fs.v0i19.4822>.
- . *Foucault and Feminist Philosophy of Disability*. University of Michigan Press, 2017.

NEWS FROM THE COMMITTEE ON THE STATUS OF WOMEN

On behalf of the committee, I am pleased to welcome the following new members, whose terms end in 2022:

Naomi Scheman
 Katie Kirkland
 Luciana Garbago
 And *ex officio*, Jill Gordon as the new Site Visit coordinator.

I also extend the entire committee's gratitude to the following outgoing members, whose terms ended as of July 1, 2019:

Charlotte Witt (chair)
 Amy Baehr
 Michael Rea
 Margaret Atherton
 And *ex officio*, Peggy DesAutels, the founder and the outgoing coordinator of the Site Visit program.

SEEKING NEW EDITOR OF BLOG SERIES

The CSW seeks a new editor, for a term starting in March 2020, of the Women in Philosophy series on the *Blog of the APA*. Thanks to inaugural editor Adriel Trott, this has been a successful series and a source of communal support in its first two years. The successful blog editor serves *ex officio* as a voting member of the Committee on the Status of Women. If interested in being considered for the position, email the chair at kathrynnorlock@gmail.com.

"ASK A SENIOR WOMAN PHILOSOPHER" IS A YEAR OLD!

As readers know, for over a year, our blog series editor Adriel Trott has offered a platform for voices and perspectives that are not often given space in the field by editing a series on Women in Philosophy for the APA's blog. The first installment of her mini-series, "Ask a Senior Woman Philosopher," was posted in August 2018, the second in December 2018, and the third in January 2019. See the July 2019 installment for [Jill Gordon's response](#) to "Dealing with Unwelcome or Inappropriate Attention." If you have a question for which you would like advice from a senior woman philosopher but don't have someone to ask or don't feel like you can ask the senior women philosophers you know, send your question to the series editor, Adriel M. Trott, at trotta@wabash.edu. Questions will be anonymized and a suitable respondent found. Topics in the wider series of Women in Philosophy thus far have included feminist philosophy conferences, Southern Black feminism, the work of the Graduate Student Council of the APA, the importance of having people who have experienced oppression working in relevant areas of philosophy, and a call to decolonize the philosophical canon, among other topics. At the CSW we noticed that the comment sections still tend to be populated by men, and often men who are telling the posters how to better think about diversity, so Adriel Trott has also worked with the APA on better moderation of the comments. If you are interested in supporting the series, consider submitting a post to the series editor (Adriel M. Trott at trotta@wabash.edu) or commenting on posts.

CSW POSTERS

Two new posters are available for purchase on the CSW website (<http://www.apaonlinecsw.org/>). The background for the "W" poster includes hundreds of names and images of contemporary women in philosophy. All the posters in this series are available for purchase at <http://www.zazzle.com/apacsw>.

COMMITTEE MEMBERS FOR 2019–2020

Adriel M. Trott (APA Blog Series Editor), Kathryn J. Norlock (Chair 2022), Rachel V. McKinnon (Member 2020), Julinna C. Oxley (Member 2020), Katie Stockdale (Member 2020), Nancy Bauer (Member 2021), Nicole J. Hassoun (Member 2021), Janet A. Kourany (Member 2021), Luciana S. Garbago (Member 2022), Katie Kirkland (Member 2022), Naomi Scheman (Member 2022), Lauren Freeman (Newsletter Editor), Jill Gordon (Site Visit Program Director)

ANNOUNCEMENTS

FEMINIST APPROACHES TO BIOETHICS (FAB) CONFERENCE 2020

University of Pennsylvania, Philadelphia

June 17–18, 2020

IAB: June 19–21, 2020.

The FAB conference is a part of the World Congress of Bioethics (IAB). The CFP is coming in a few weeks.

World Congress of Bioethics website: <http://www.bioethics.net/2018/12/iab-world-congress-of-bioethics-2020/>

If you have questions, please contact Anna Gotlib at agotlib@brooklyn.cuny.edu.

14TH ANNUAL PHILOSOPHIA CONFERENCE: A SOCIETY FOR CONTINENTAL FEMINISM – CALL FOR ABSTRACTS

Hosted by Vanderbilt University, Nashville TN
May 14–17, 2020

Plenary Speakers:

Kathryn Sophia Belle, Penn State University
Lisa Guenther, Queens University, Canada
Tracy Sharpley Whiting, Vanderbilt University

Plenary Panel: New Perspectives on Disability

Kim Q. Hall, Melinda Hall, Joel Reynolds, and Shelley Tremain

The conference will have two workshop streams: Rethinking Prisons and Rethinking Disability.

Submit abstracts (500–700 words), or panel proposals [panel abstract (500 words) plus panelists' abstracts (500–700 words)], on any topic related to Continental Feminism—very broadly construed—for the general program. Indicate on your abstract if you are applying to participate in a workshop. **Send abstracts to 14thphilosophia@gmail.com by December 15, 2019.** Lodging will be provided on a limited basis.

FEMINIST ETHICS AND SOCIAL THEORY (FEAST)

The Future of Feminist Ethics: Intersectionality, Epistemology, and Grace

Celebrating FEAST's 20th year

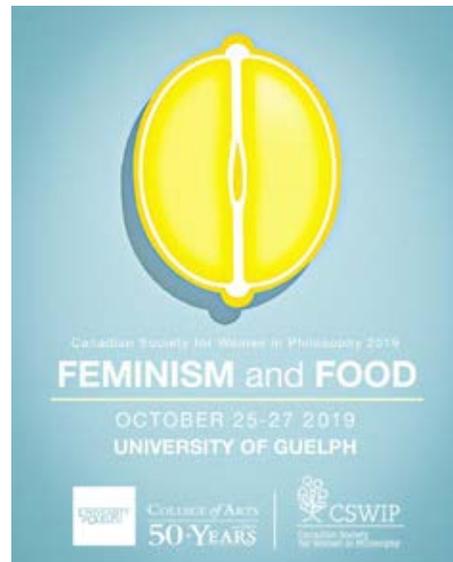
October 3–6, 2019

Sheraton Sand Key Resort in Clearwater Beach, FL

This year we're asking what challenges feminists continue to face and what new challenges have arisen since FEAST first began and how "revisiting feminist ethics" at this juncture might help feminists confront those challenges.

Thinking through how well we have engaged intersectionality and where we've failed; engaging with how our ethical and political lives are intertwined with our lives as knowers; articulating the grace—sometimes generous and sometimes critical—necessary to navigate this difficult world.

We are fortunate to be featuring a Keynote Conversation between Kristie Dotson and Britney Cooper and a Keynote by Talia Bettcher. We will also be holding a session in honor of Joan Callahan.



CONTRIBUTOR BIOS

Julia R. S. Bursten is an assistant professor of philosophy at the University of Kentucky and former co-chair of the Philosophy of Science Association Women's Caucus. Her research concerns questions of modeling and scale in nanoscience and other physical sciences. From 2011 to 2015 she served as the resident philosopher in the Millstone Nanosynthesis Laboratory at the University of Pittsburgh.

Catherine Clune-Taylor is assistant professor of feminist science and technology studies in the Department of Women's Studies at San Diego State University. Before that, she was a postdoctoral research associate in the Program in Gender and Sexuality Studies at Princeton University. In addition to a PhD in philosophy from the University of Alberta (and some other degrees in philosophy), she has a BMSc in microbiology and immunology from the University of Western Ontario. Clune-Taylor writes and teaches in the fields (and at the intersections) of feminist theory, philosophy of gender and sexuality, critical disability studies, and bioethics. She is currently at work on a book critically exploring the science, ethics, and biopolitics underwriting contemporary clinical, social, and political management of intersex and trans lives. She has published articles in *PhaenEx: Journal of Existential and Phenomenological Theory and Culture*, *The American*

Journal of Public Health, and *Hypatia* (forthcoming, November 2019).

Cassie Herbert is an assistant professor at Illinois State University, where she holds a joint appointment in Philosophy and Women's and Gender Studies. She received her PhD in philosophy from Georgetown. She works in social philosophy, philosophy of language, and philosophy of sex. Her research focuses on the pragmatics of slurs, how we use language to construct ingroups and outgroups, and on the harms of sexual violence. In her spare time, Cassie boxes, bikes, and takes great pleasure in making excessively intricate crafts.

Alice MacLachlan is an associate professor of philosophy at York University (Toronto, Canada), where she directs the graduate program in philosophy, and is co-editor of *Feminist Philosophy Quarterly*, an open-access, peer-reviewed, scholarly journal. She writes and teaches in moral, political, and feminist philosophy, focusing on philosophical issues arising in the aftermath of conflict: the nature and limits of forgiveness, the power and value of apologies, and the role that emotions like resentment and indignation play in reconciliation and repair. She has also worked in LGBTQ philosophy on topics ranging from the ethics of coming out to the changing nature of family. Her recent publications include "Fiduciary Duties and the Ethics of Public Apology" (*Journal of Applied Philosophy* 2018), "'Trust Me, I'm Sorry: The Paradox of Public Apology'" (*Monist* 2015), "Gender and the Public Apology" (*Transitional Justice Review* 2013), and "Closet Doors and Stage Lights: On the Goods of Out" (*Social Theory and Practice* 2012). She co-edited a collection titled *Justice, Responsibility, and Reconciliation in the Wake of Conflict* (Springer 2013) and she is currently completing a SSHRC-funded research project on the ethics and politics of civility.

Mari Mikkola is an associate professor of philosophy at University of Oxford and tutorial fellow of Somerville College. She is the author of two books (*The Wrong of Injustice: Dehumanization and Its Role in Feminist Philosophy* and *Pornography: A Philosophical Introduction*, both with Oxford University Press) and of several articles on feminist philosophy, social ontology, and pornography.

Sarah Clark Miller is associate professor of Philosophy and affiliate faculty in Bioethics and Women's, Gender, and Sexuality Studies at Pennsylvania State University. She is an ethicist who also works in social and political thought. She is past acting and associate director and current faculty affiliate of the Rock Ethics Institute. Her recent work includes *The Ethics of Need: Agency, Dignity, and Obligation* (Routledge) and articles on sexual violence, relational dignity, reproductive ethics, global responsibility, and harm and moral injury in journals such as *The Journal of Social Philosophy* and *Social Theory and Practice*. She is currently completing a book on sexual violence.

Miranda Pilipchuk is a PhD candidate in the Department of Philosophy at Villanova University, specializing in feminist theory, philosophy of law, and critical race theory. Her dissertation examines the relationship between public discourses surrounding sexual violence and logics of mass incarceration. She is the recipient of the Social Sciences and Humanities Research Council of Canada Doctoral Fellowship, the Sir James Lougheed Award of Distinction, the Villanova University Gender and Women's Studies Graduate Research Award, and is a two-time winner of the Greater Philadelphia Women's Studies Consortium Graduate Research Award. She served as the managing editor of *Hypatia: A Journal of Feminist Philosophy* for two and a half years, and she was an appointed member of the APA's inaugural Graduate Student Council, a PIKSI: Boston Teaching Fellow, and the founding member of Villanova's Minorities and Philosophy chapter. Her previous work has been published on the *Blog of the APA*.

Caleb Ward is a PhD candidate in philosophy at Stony Brook University (SUNY). His research is about sexual consent and responsibility in intimacy, a topic he approaches using feminist philosophy (especially critical theory and black feminist thought), as well as tools from both continental and analytic ethics. He has also co-edited two volumes on food ethics, including the *Routledge Handbook of Food Ethics* (2017).

Lori Watson is professor and chair of Philosophy at the University of San Diego; she is also an affiliate faculty member in the School of Law. Her books include *Equal Citizenship and Public Reason: A Feminist Political Liberalism*, co-authored with Christie Hartley (Oxford University Press, 2018); *Debating Pornography* with Andrew Altman (Oxford University Press, 2019); and *Debating Sex Work* with Jessica Flannigan (forthcoming).

Robin Zheng is an assistant professor of philosophy at Yale-NUS College. She specializes in ethics, moral psychology, and feminist and social philosophy. Her recent publications include "Why Yellow Fever Isn't Flattering: A Case Against Racial Fetish" in the *Journal of the American Philosophical Association*; "Bias, Structure, and Injustice: A Reply to Haslanger" in *Feminist Philosophy Quarterly*; "A Job for Philosophers: Causality, Responsibility, and Explaining Social Inequality" in *Dialogue: Canadian Philosophical Review*; and "Precarity is a Feminist Issue: Gender and Contingent Labor in the Academy" in *Hypatia: A Journal of Feminist Philosophy*.



APA NEWSLETTER ON

Hispanic/Latino Issues in Philosophy

CARLOS A. SÁNCHEZ AND LORI GALLEGOS DE CASTILLO, CO-EDITORS

VOLUME 19 | NUMBER 1 | FALL 2019

FROM THE EDITORS

Carlos Alberto Sánchez
SAN JOSE STATE UNIVERSITY

Lori Gallegos de Castillo
TEXAS STATE UNIVERSITY

For this issue of the newsletter, we have brought together a variety of essays that showcase the true diversity of philosophical projects occupying our Hispanic/Latinx philosophical community. We've divided this issue into three sections: the first, a special cluster edited by Stephanie Rivera Berruz, is comprised of papers delivered at the Latinx Philosophy Conference at Marquette University in May of this year; the second section is comprised of four articles dealing with issues of race and embodiment, Latinx thought, exile, and Latin American Marxism; and, finally, the third section includes an excerpt from a forthcoming book by Susana Nuccetelli and a book review of Linda Martín Alcoff's *Rape and Resistance*.

We begin this issue with a Special Cluster of essays that came out of a 2019 conference jointly hosted by the Fourth Latinx Philosophy Conference and the Society for Mexican American Philosophy at Marquette University. Stephanie Rivera Berruz, one of the conference organizers, introduces this exciting collection of papers.

In the following section, Omar Rivera offers a critique of José Carlos Mariátegui's aesthetics, namely, that its representative investments in indigeneity can essentialize Andean peoples and cultures in order to support Mariátegui's own revolutionary program. Then, reading Mariátegui against himself, Rivera approaches Mariátegui's *Seven Essays* by drawing from a notion of racial embodiment taken from W.E.B. DuBois, Frantz Fanon, and Linda Martín Alcoff. This counter-reading reveals a dynamic and contextual approach to racial embodiment, one that complicates Mariátegui's adherence to a racist representative aesthetics of liberation.

In that same section, the essay by Alejandro Vallega seeks to dismantle the traditional, reductionist treatment of Latino/a/x identities as a mere racial identity. He calls on us to engage the diversity of Latino/a/x lineages, histories, and experiences which, he argues, ontologically ground a rich and distinctive Latino/a/x thought and consciousness. Specifically, building on the work of Aníbal Quijano, Vallega shows how the temporality of Latin American experience unsettles the single history of progress established by the coloniality of power and knowledge.

Sergio Lomeli Gamboa's "Jose Revueltas' Marxism: A Struggle Against Orthodoxy" is an excellent introduction into the thought of this brilliant Mexican thinker. Known mostly for his fictional writing, Revueltas is also one of Marx's most original readers and a fierce critic of "Orthodox Marxism." Gamboa's reading here seeks not only to locate Revueltas in the center of Mexico's philosophical landscape of the last century, but also at the center of Mexico's cultural and political life. This care for the Mexican circumstance is symptomatic of Mexican philosophy as a whole, and the essay presented here is exemplary in its execution. To close off this section, Arturo Aguirre's short piece aims to think about the concept of "exile" as an operative concept capable of being deployed for the analysis of the modern dispossessed. Beginning with an "archeology" of the concept in Plato's *Laws*, Aguirre ends by suggesting that the concept of exile is intimately violent.

In the third section, we are pleased to offer our readers an excerpt from Susana Nuccetelli's forthcoming book on Latin American philosophy. In this intriguing essay, she considers some of Karl Marx's criticisms of Simón Bolívar. In particular, she focuses on the nature of Bolívar's sympathies for authoritarian government.

Finally, we close our newsletter with Juan Colomina-Alminana's review of Linda Martín Alcoff's *Rape and Resistance*, which centers on historical, institutional factors that allow and condone sexual violence and assault. Colomina-Alminana focuses his commentary on Alcoff's discussion of how to transform the unjust conditions that currently frame survivors' speech and the handling of that speech by others.

CALL FOR SUBMISSIONS

The *APA Newsletter on Hispanic/Latino Issues in Philosophy* is accepting contributions for the SPRING 2020 issue. Our readers are encouraged to submit original work on any topic related to Hispanic/Latino thought, broadly construed. We publish original, scholarly treatments, as well as reflections, book reviews, and interviews. Please prepare articles for anonymous review.

All submissions should be accompanied by a short biographical summary of the author. Electronic submissions are preferred. All submissions should be limited to 5,000 words (twenty double-spaced pages) and must follow the APA guidelines for gender-neutral language and *The Chicago Manual of Style* formatting. All articles submitted

to the newsletter undergo anonymous review by members of the Committee on Hispanics.

BOOK REVIEWS

Book reviews in any area of Hispanic/Latino philosophy, broadly construed, are welcome. Submissions should be accompanied by a short biographical summary of the author. Book reviews may be short (500 words) or long (1,500 words). Electronic submissions are preferred.

DEADLINES

Deadline for spring issue is November 15. Authors should expect a decision by January 15. Deadline for the fall issue is April 15. Authors should expect a decision by June 15.

Please send all articles, book reviews, queries, comments, or suggestions electronically to the editor, Carlos Alberto Sánchez, at carlos.sanchez@sjsu.edu, or by post: Department of Philosophy, San Jose State University, One Washington Sq., San Jose, CA 95192-0096.

FORMATTING GUIDELINES

The APA Newsletters adhere to *The Chicago Manual of Style*. Use as little formatting as possible. Details like page numbers, headers, footers, and columns will be added later. Use tabs instead of multiple spaces for indenting. Use italics instead of underlining. Use an "em dash" (—) instead of a double hyphen (--). Use endnotes instead of footnotes. Examples of proper endnote style: John Rawls, *A Theory of Justice* (Cambridge: Harvard University Press, 1971), 90. See Sally Haslanger, "Gender and Race: (What) Are They? (What) Do We Want Them to Be?" *Noûs* 34 (2000): 31–55.

SPECIAL CLUSTER

Latinx Philosophy Conference

Stephanie Rivera Berruz
MARQUETTE UNIVERSITY

In the spring (May 2–4, 2019), the Fourth Latinx Philosophy Conference partnered with the APA Society for Mexican American Philosophy (SMAP) for one full-length conference at Marquette University. The conference celebrated philosophical work by Latinx and Latin American philosophers on issues particularly relevant to Latinx and Latin American communities. The event was the first of its kind and successfully brought together scholars working on a broad range of topics and from a wide variety of philosophical and interdisciplinary traditions. The program featured the work of over a dozen scholars at different stages of their academic careers and was keynoteed by Jacqueline Martínez (Arizona State University) and José Medina (Northwestern University). Historically, the conference has intended to bring scholars together who otherwise would not have had an opportunity to intersect. In this vein, the special cluster featured in this issue showcases the possibilities of such intersection. The authors here further remind readers that our current political climate, particularly as it pertains to immigration, is material, lived, and with great consequence.

Allison Wolf's essay "Dying in Detention as an Example of Oppression" argues for the importance of a structural lens of oppression in order to better understand the reasons why Raul Morales, Moises Tino, Roxsana Hernández, and Raquel Calderon died while in US detention. Alan Chavoya's essay, "A Negative Path Toward Anti-Immigration Policy," maintains that current US non-racist immigration policy does little more than contribute to the myth of a post-racial social order while at the same time perpetuating mass amounts of racial injustice. As a result, Chavoya defends the possibilities of anti-racist immigration policy by using a negative conception of justice as the starting point. Considering issues of injustice and immigration on an international scale, "Immigration and International Justice" by Jorge M. Valadez explores the moral dimensions of the territorial rights of states as a mechanism for developing immigration policies that can address injustices globally. Juxtaposed, Wolf, Chavoya, and Valadez importantly remind us of the ways in which the intersection between immigration and criminalization are structurally embedded into our social fabrics, and they demand more careful and nuanced social and political analysis. Eric Bayruns Garcia and Damián Bravo Zamora shift attention toward the epistemic dimensions of our social fabric. In "Are Our Racial Concepts Necessarily Essentialist Due to Our Cognitive Nature?" Bayruns Garcia explores the epistemic content framed by the social constructionism of race as argued by Mallon and Kelley. He maintains that their position does not hold by drawing on historical references to racial concepts as well as the lived experiences of race in Dominican NY-based communities. Bravo Zamora demands "Epistemic Humility Now!" as he develops an account of epistemic humility that can appreciate a multiplicity of worldviews without collapsing into relativism.

In sum, the essays in this special cluster showcase the importance of spaces/events like the Latinx Philosophy Conference as well as the Society for Mexican American Philosophy. The topics taken up here demonstrate the need for the development of ideas that track the lived experiences of our communities. It is my sincere hope that projects like these will continue to pave the way for a burgeoning field of work rooted in demands for social justice.

Dying in Detention as an Example of Oppression

Allison Wolf
SIMPSON COLLEGE

INTRODUCTION

On December 8, 2018, seven-year-old Jakelin Caal Maquin died in US custody. Shortly thereafter, on Christmas Day, eight-year-old Felipe Alonzo-Gomez died from the flu, also while in US custody.¹ Given the widespread news coverage, many may have heard about the tragic deaths of these children. But let me introduce you to a few more.

First, Raul Ernesto Morales-Ramos:

On April 6, 2015, Raul Ernesto Morales-Ramos, a 44-year-old citizen of El Salvador, died at Palmdale Regional Medical Center in Palmdale, California of organ failure, with signs of widespread cancer. He had entered immigration custody four years earlier in March 2011. An ICE investigation into the death of Morales-Ramos found that the medical care he received at both facilities failed to meet applicable standards of care in numerous ways. Two independent medical experts, analyzing ICE's investigation for Human Rights Watch, agreed that he likely suffered from symptoms of cancer starting in 2013, but that the symptoms essentially went unaddressed for two years, until a month before he died. . . . Throughout this time, Morales-Ramos repeatedly begged for care. In February 2015, he submitted a grievance in which he wrote, "To who receives this. I am letting you know that I am very sick and they don't want to care for me. The nurse only gave me ibuprofen and that only alleviates me for a few hours. Let me know if you can help me."²

Then, there is Roxsana Hernández:

Roxsana Hernández, a 33-year-old trans woman seeking asylum in the US from Honduras, reportedly died from HIV-related complications following an alleged five-day detention in what's known by immigrant rights groups as the "ice box"—Ice [sic] detention facilities notorious for their freezing temperatures.³

During her first week in the United States, Roxy's body and spirit quickly deteriorated. . . . Soon after asking for asylum, Hernandez was initially held at a detention center by US Customs and Border Protection. She was cold, lacked adequate food or medical care, and was held in a [freezing] cell [known as an ice box] with the lights turned on 24 hours a day. A week later, ICE says she arrived at the transgender unit in the Cibola County Correctional Center in Milan, New Mexico—a privately run federal prison for men that contracts with ICE. But by the next day, she was taken to a local hospital with symptoms of pneumonia, dehydration and "complications associated with HIV," ICE says. Hours later, she was transferred to a hospital in Albuquerque where "she remained in the intensive care unit until her passing," the agency said.⁴

An autopsy report showed that she was also beaten in custody.⁵

The third person is Moises Tino-Lopez:

Moises Tino-Lopez, 23, had two seizures within nine days, each observed by staff and reported to the nurses on duty in the Hall County Correctional Center in Nebraska. He was not evaluated by a physician or sent to the hospital after the first seizure. During his second seizure, staff moved him to a mattress in a new cell, but he was not evaluated by a medical practitioner. About four hours after that seizure, he was found to be unresponsive, with his lips turning blue. He was sent to the hospital but never regained consciousness and died on September 19, 2016.⁶

And, finally, Raquel Calderon de Hildago:

A 36-year-old Guatemalan woman, Raquel Calderon de Hildago, died Sunday at Banner Casa Grande Medical Center, according to Immigration and Customs Enforcement officials at the Eloy Detention Center in Arizona. She was taken to the hospital after medical staff at the Eloy Detention Center called paramedics following a series of seizures, ICE officials said in a news release. The woman continued to experience seizures in the ambulance on the way to the hospital, ICE officials said.⁷ She was the 15th person to die at that facility.

According to an autopsy report, however, Calderon de Hildago died of blood clots "throughout all the lobes of the right lung," the result of a leg injury. Reportedly, Calderon de Hildago had difficulty walking at the time she was apprehended by the Border Patrol. The blood clots, which traveled from her leg to her lung, would have triggered the seizures, Dr. Gregory Hess, Pima County's chief medical examiner, said in an interview.⁸

In 2013, women prisoners in Eloy Detention Center went on hunger strike to protest life threatening conditions there.⁹ Calderon de Hildago died November 27, 2016.

Sadly, these six cases (the four just highlighted and the two children) are not anomalies. In 2019, three more children died in US Customs and Border custody, the most recent being a sixteen-year-old Guatemalan teenager, Carlos Hernandez Vasquez, who crossed the border in Texas.¹⁰ Just six days earlier, a Guatemalan toddler also died in US custody.¹¹ In 2018, at least twelve adults and three children died in US detention facilities.¹² In 2017, twelve people died in US detention facilities and at least ten women filed complaints against ICE for causing them to miscarry due to their treatment in custody. In fact, Human Rights Watch reports that between 2010 and 2018, seventy-four people have died in US detention facilities.¹³ Worse, over half of these deaths were both preventable and caused by inadequate medical care offered by Immigration and Customs Enforcement facilities.¹⁴

The purpose of this essay is not to argue that these deaths are wrong—I take that as obvious. Instead, I want to focus

on *why* they are wrong. More specifically, I argue against mainstream philosophers of immigration that the principal wrong in these circumstances is the rampant existence of human rights violations (though I concede that is clearly the case). I will suggest that the major wrong underlying this pattern of deplorable conditions in detention that lead to immigrants' deaths are best understood as the result of oppression. But first, I need to tell you a bit about US detention policy and the rights of detainees.

A BRIEF OVERVIEW OF US DETENTION POLICY

Migrants who are detained are not criminals; they have not committed a criminal act and or broken criminal law. So why do we detain them? Two reasons: "(1) to ensure individuals appear for immigration court proceedings" and (2) to "comply with removal orders."¹⁵ Now, because detained migrants have not broken criminal laws, they are not supposed to be housed in jails or with criminals. They are entitled to be housed in good conditions, with the ability, for example, to exercise, move freely within the facility, eat well, be safe, and have good health care.

Despite the fact that, for many of us, detaining migrants for these purposes seems relatively commonplace, the idea of detaining large numbers of immigrants in the United States while they await their immigration hearing or deportation is relatively new. For most of the twentieth century, the Border Patrol and Immigration and Naturalization Services (INS) did not focus on detention or deportation. Their job was to *regulate* migration, not to *punish* it.¹⁶ And so Border Patrol and INS were not focused on criminal prosecutions of immigration violations at all and, as sociologist Patricia Macías-Rojas explains, "Training manuals of the time instructed agents to avoid illegal detention" on the grounds that it violates the Constitution.¹⁷ Instead, if someone was caught crossing the border without papers, they were simply sent home. So what changed?

Macías-Rojas traces what she refers to as the "punitive turn" in immigrant detention to the 1980s and 1990s. In 1986, the Immigration Reform and Control Act introduced a criminal alien program designed to deport convicted felons.¹⁸ In 1988, the Anti-Drug Abuse Act created provisions for deporting noncitizens who committed aggravated felonies.¹⁹ In 1996, however, the game changed when the Illegal Immigration Reform and Immigrant Responsibility Act (IIRICA) was introduced to target criminal aliens as a major enforcement priority and expanded the list of aggravated felonies for which one could be deported to include petty theft, minor drug offenses, and DUIs.²⁰ It also required mandatory detention of noncitizens who had completed their prison sentences.²¹ Notice that increased deportation and detention was *never* about immigration *per se*; it was about drugs and crime.

And it stayed that way until September 2001 when the USA PATRIOT Act authorized indefinite detention of those suspected of being terrorists in the wake of the September 11, 2001, attacks. At that point, immigration matters, generally, and deportation and detention, in particular, were reframed as national security concerns. This led to the reorganizing of governmental operations and agencies to better emphasize this new focus. In 2003, the Immigration

and Naturalization Service was disbanded and Immigration and Customs Enforcement (ICE) was created to deal with internal enforcement via deportation and detention. By 2004, the Department of Human Services (DHS) started prosecuting those who entered the country without permission in order to minimize the practice of paroling eligible, mostly Central American, migrants from detention (i.e., catch and release), a policy that existed precisely because detention centers did not have enough beds.²² In order to support detention Congress gave increasing funds to maintain detention beds. In 2009, Congress mandated that at least 33,400 detention beds be available. Recently, the Trump Administration has ordered that those who are caught entering the country without permission be charged with a criminal violation for the first time in recent history. The result is that we need more beds but do not have them. In response, in its most recent budget, the Trump Administration pushed to expand the number of beds to more than 51,000.²³ Currently, the United States detains the highest number of migrants in the world, close to 450,000, most of whom are male and from Mexico and Central America.²⁴ We still do not have enough beds to detain all noncitizens, so "ICE officers order detention on a discretionary basis."²⁵

In detention, migrants (even if they commit a crime) have a right to health care. Moreover, most of those who are detained at the border are asylum seekers, following US and international law by crossing the US border and then petitioning for asylum. As such, they too have the right to a certain level of care, including medical care. The Inter-American Council on Human Rights of the Organization of American States (of which the United States is a part), for example, states that:

In border areas, it is the duty of state authorities to provide immediate assistance to migrants and asylum-seekers, and that this assistance must specifically include medical assistance, adequate food and water, clothing, blankets, personal hygiene supplies, and opportunity to rest. As part of operations to intercept migrants and control migration at international borders, it is the duty of state authorities to prioritize medical and health screenings for migrants and asylum-seekers. This implies that competent medical personnel must be present in places where migrants and asylum-seekers are intercepted or detained in order to examine them and refer them for further medical attention, including mental health referrals, when appropriate.²⁶

Similarly, the United Nations argues that immigrants—refugees, asylum seekers, or otherwise—have a right to health care. Article 12 of the International Covenant on Economic, Social and Cultural Rights, for example, maintains that:

States that are a party to the Covenant have an obligation to ensure the provision of: equal and timely access to basic preventive, curative, rehabilitative health services and health education; regular screening programmes; appropriate

treatment of prevalent diseases, illnesses, injuries and disabilities, preferably at community level; essential drugs; and appropriate mental health treatment and care. They also have an obligation to prevent, treat and control epidemic, endemic, occupational and other diseases. Under article 12.2 (d), they must create “conditions which would assure to all medical service and medical attention in the event of sickness.”²⁷

Even the United States government recognizes its obligations to provide health care to immigrants in detention, as evidenced by the fact that ICE maintains the ICE Health Service Corps (IHSC) to provide “direct care to approximately 13,500 detainees housed at 21 designated facilities throughout the Nation to include medical, dental, and mental health care, and public health services as well as medical case management and oversight for an additional 15,000 detainees housed at approximately 119 non-IHSC staffed detention facilities across the country.” And IHSC oversees the financial authorization and payment for off-site specialty and emergency care services for detainees in ICE custody.²⁸

Despite this, medical care and services for migrants is woefully inadequate. First, the numbers show this is the case—ICE’s own numbers show that they are equipped to provide care to 28,000 detainees, but there are hundreds of thousands in the system. Unsurprisingly, then, the American Civil Liberties Union reports: “ICE puts thousands of people’s health and lives at risk by failing to provide adequate medical care to the people it detains for weeks, months, and even years.”²⁹ As we saw in the cases described earlier, responses are delayed or inadequate, emergency responses botched and slow, and the quality of care is poor. As a result, migrants are suffering (and increasingly dying).

WHY IS THIS WRONG? OPPRESSION

There are many reasons that people would morally condemn the treatment of detainees in the United States as well as the increasing numbers of people getting sick and dying in US custody. Within the literature on immigration justice, the most common condemnation is framed as a violation of immigrants’ basic human rights. I think this intuition is correct for all sorts of reasons: This treatment violates Article 3 of the Universal Declaration for Human Rights, which states that “everyone has the right to life, liberty, and security of person,” as well as Article 9, which states that “no one shall be subjected to arbitrary arrest, detention, or exile.”³⁰ Some, such as philosopher Joseph Carens, argue that detentions, arrests, and most types of border control also violate Article 13, which guarantees the right of all to exit their country.³¹ Still, I think that viewing these injustices as human rights violations is not our best bet for a couple of reasons: (1) Because the human rights framework focuses on individual acts and incidents, which means it is hard to detect structural and systemic issues; and (2) because the human rights framework takes the target and perpetrator of the wrong to be individuals and not social groups or nations. As a result, the human rights approach to thinking about immigration justice and injustice will not actually reveal what I take to be the core immigration injustice in

these cases, namely, that the treatment of detainees and their resulting deaths are the result of oppression.

Let me state my position more clearly. I argue that immigration injustice should be defined as occurring when any immigration policy, practice, norm, or system perpetuates, reflects, or maintains oppression. US detention policies, practices, and systems perpetuate oppression in various ways. Therefore, they are unjust, and the core reason is their role in oppression, not the fact that they violate human rights (though they clearly do that too). But what is “oppression”? To help explain, I employ feminist philosopher Marilyn Frye’s famous metaphor of a bird cage. Imagine a birdcage. When we pay attention to the birdcage, we notice that the cage only functions when multiple wires are arranged in specific ways, namely, to “restrain, restrict, or prevent the [bird’s] motion or mobility.”³² In other words, the cage does not consist in random, unrelated wires just thrown together. To the contrary, the cage is formed by the specific way that wires are put together in relation to the others. Further, to be arranged in this specific way is neither random, accidental, or haphazard. The cage is designed to trap the bird and prevent it from escaping.

Oppression is similar. Under oppression, social and institutional norms, policies, practices, and structures work together to reduce, immobilize, and mold the oppressed specifically to restrict and impede their options. The cage does not simply work to trap the bird. Rather, it works to trap the bird in particular ways—namely, by creating double-binds. A double-bind exists when the oppressed’s “options are reduced to a very few and all of them expose one to penalty, censure, or deprivation.”³³ No matter what they do, the oppressed are vulnerable to facing negative consequences. A common example of a double-bind is young women’s sexual expression. If young women have sex, they are whores, but if they abstain, they are prudish, teases, and frigid. Regardless of their choice, they are vulnerable to censure. The experience of oppression is being caught up in a system that is designed to trap you in double-binds.

Frye’s birdcage metaphor raises a key question: Who or what does the bird represent? In the immigration context, the answer is almost always social groups, nations, nation-states, transnational communities (like Indigenous groups). So when someone finds herself in a “cage” and asks, “Why am I in this cage?” (i.e., Why am I caught in this double-bind?), the answer will not be that individual’s actions. Instead, the answer will be “You are in the cage because you are a member of a certain social group, nation, state, community, etc.” For example, in the US there is a practice of racial profiling Black and Latino men (and other men of color). Under these circumstances, there is nothing an individual Black or Latino man can do to avoid the police pulling him over—driving too fast will be cause, since one is speeding; driving too slow will be cause, since it is suspicious; driving the speed limit will be cause, since nobody drives the speed limit unless they are hiding something. In these cases, the man is not stopped because of his actions (all of which many white people do on a regular basis, including myself—much to my mother’s chagrin). He is stopped because he is Black or Latino. In

this way, one is oppressed because of their membership in a social group, nationality, citizenship, or community.

Of course, individuals are members of multiple groups, communities, and even nations. Many individuals can be both victims and perpetrators of oppression. Take me as an example. As a Jewish woman in Iowa, I am a victim of sexism, Anti-Semitism, and their intersections in a shockingly large number of ways. But, in Los Angeles, where I am primarily perceived as white (especially in the middle of these frigid winters that take all of the glorious color out of my skin), I also participate in and benefit from racism and classism. Both are true because of my belonging to multiple social groups.

To be clear, to say that these networks of systems and barriers are set up to restrict and immobilize the oppressed does not mean that oppression is conscious, calculating, or necessarily done by evil people with malice. In fact, while we cannot deny that this can occur, more often than not, oppression is carried out by well-meaning people upholding norms or enacting policies that are deeply problematic. The problem lies in the structures and norms being upheld in the policy, not the individual character flaws of those who create or enforce it. To put it more concretely, the problem is not that US Customs and Border Patrol officials are intentionally cruel as they deport people to Mexico and Central America, but rather the policies that they are enforcing, for example, requiring them to detain all irregular migrants, reflect and promote oppressive norms. Oppression is located in the structures, values, and practices, not in individuals.

A consequence of this conceptualization is that detecting oppression requires utilizing a macroscopic analysis—one must see how everything is connected (history, politics, laws, social norms, etc.) to determine whether a particular instance is a random, unfortunate event, or part of a larger system of oppression. Again, we can return to Frye’s metaphor of the birdcage. If one examines why a birdcage encloses the bird using a microscopic and individual approach, the answer is elusive. This is because the person only sees one wire or one wire at a time, and the presence of one wire does not keep a bird from flying away. In other words, the investigator will fail to see how the wires are connected. But the bird is in the cage precisely because of how all of the wires are connected. So the only way to identify oppression is to step back, to get a macroscopic view of the various relationships between the wires, the environment, etc.

OPPRESSION IN THE DEATHS OF IMMIGRANTS IN US DETENTION

Now that we have some background on US detention policy and the nature of oppression, we can return to the four cases that I introduced at the beginning of the essay. Clearly, I do not have time to detail how all four died as a result of oppression in US immigration policy here, but I will suggest some ways that we should see their treatment and resulting deaths as examples of oppression. First, their treatment and deaths resulted from structural problems, not random bad acts or incompetent individuals. US policies,

practices, and norms around immigration and health care led to these deaths. The Enforcement and Removal Office (the agency in charge of inspecting facilities) does not impose meaningful consequences for failing inspection standards.³⁴ “As a result, inhumane conditions, including egregious violations of medical care standards, prevail across an immigration detention system composed of more than 200 detention facilities.”³⁵ According to four major human rights organizations, “Three failings stand out: (1) unreasonable delays in providing care, (2) poor practitioner and nursing care, and (3) botched emergency responses.”³⁶ Even the government’s own reports in 2016, 2017, and 2018 provide additional documentation of deficient medical care in ICE facilities.³⁷ And despite

new growth in immigration detention, the Trump administration has requested less money for Department of Homeland Security (DHS) oversight of detention to assure that conditions of confinement are safe, indicating that it plans to abandon basic standards developed over the past decade intended to protect the health, safety, and human rights of those held in immigration detention centers. These proposals would place more human beings than ever before into an abusive and wasteful system that already suffers from substandard medical care.³⁸

This is a systemic problem—it is a cage, not individual, unconnected wires.

Second, the system is set up in such a way as to place migrants in a double-bind. New detention policies have turned the decision to migrate and/or seek asylum itself into a double-bind—stay in one’s country and face violence, poverty, and economic deprivation, or come to the US and risk violence, detention, and ill health. No matter what any individual migrants does, she is vulnerable to negative consequences.

Third, these policies place migrants in double-binds because of their social group membership and nationality. Not all migrants, even unauthorized ones, are detained. To the contrary, most detentions are still based on the discretion of ICE agents, because there are far too few beds to detain all of those who enter without authorization.³⁹ ICE officers overwhelmingly use their discretion to target migrants from Mexico and Central America, and, within this group, LGBTQ migrants.⁴⁰ There is no detention, for example, for those crossing the Northern Border without permission or proper documentation. Moreover, the Trump Administration has expressly said it wants to stop immigration from Central America specifically, and has threatened Central American governments to cut off aid if they do not help, indicating that these policies are, again, targeted toward specific groups and not all migrants. Beyond this, the Border Patrol has been directed by the Department of Homeland Security to target those coming from Central America.⁴¹ In fact, the now infamous zero-tolerance policy that has led children to be separated from their parents specifically directs all federal prosecutors to detain and arrest all who enter the United States “illegally” on the *Southern Border*⁴² and send them to federal prison or detention centers while they

wait for their cases to be heard. These individuals are in detention because of their national membership, not their actions.

CONCLUSION

Raul Morales, Moises Tino, Roxsana Hernández, Raquel Calderon de Hidalgo, and the dozens like them who died at the hands of our government came here in search of a better life—a life free of sexual violence, gang violence, starving wages, and constant uncertainty. Jakelin Caal Maquin and Felipe Alonzo-Gomez were simply following their parents. Instead of helping them, or treating them with even minimal decency, the US government detained them in such a cruel and inhumane manner that they passed from this world. I hope that by reframing their experiences as expressions of oppression, we can fight to make sure that those who follow them at least get the chance they deserve to come out of the process alive and well.

ENDNOTES

1. Miriam Jordan, "8-Year-Old Migrant Child from Guatemala Dies in US Custody," *New York Times*, December 25, 2018, <https://www.nytimes.com/2018/12/25/us/guatemalan-boy-dies-border-patrol.html>; Camilo Montoya-Galvez, "Migrant Child Dies in Border Patrol Custody; Second in a Month," *CBS News*, December 26, 2018, <https://www.cbsnews.com/news/migrant-child-dies-in-border-patrol-custody-today-2018-12-25/>.
2. Human Rights Watch, "Systemic Indifference: Dangerous and Substandard Medical Care in US Immigration Detention," *HRW.org*, May 8, 2017, <https://www.hrw.org/report/2017/05/08/systemic-indifference/dangerous-substandard-medical-care-us-immigration-detention>.
3. Carla Green, "Transgender Honduran Woman's Death in US 'Ice-Box' Detention Prompts Outcry," *The Guardian*, May 31, 2018, <https://www.theguardian.com/us-news/2018/may/31/roxana-hernandez-transgender-honduran-woman-dies-us-ice-box>.
4. Nicole Chavez, "Transgender Woman in Migrant Caravan Dies in ICE Custody," *CNN*, May 31, 2018, <https://www.cnn.com/2018/05/31/health/transgender-migrant-dies-ice-custody/index.html>.
5. Sandra Garcia, "Independent Autopsy of Transgender Asylum Seeker Who Died in ICE Custody Shows Signs of Abuse," *New York Times*, November 27, 2018, <https://www.nytimes.com/2018/11/27/us/trans-woman-roxsana-hernandez-ice-autopsy.html>.
6. Human Rights Watch, "US: Poor Medical Care, Deaths, in Immigrant Detention," *HRW.org*, June 20, 2018, <https://www.hrw.org/news/2018/06/20/us-poor-medical-care-deaths-immigrant-detention>.
7. Daniel Gonzalez, "Another Death at Eloy Migrant-Detention Center," *AZCentral.com*, November 28, 2016, <https://www.greenbaypressgazette.com/story/news/politics/border-issues/2016/11/28/another-death-elyo-migrant-detention-center/94574478/>.
8. Daniel Gonzalez, "Autopsy: Eloy Immigration Detainee Died of Blood Clots in Her Lung," *AZCentral.com*, January 2, 2017, <https://www.azcentral.com/story/news/local/arizona/2017/01/02/autopsy-elyo-immigration-detainee-died-blood-clots-her-lung/95966572/>; Michel Marizco, "Autopsy: Guatemalan Detainee Who Died in Arizona Died from Blood Clot," *KJZZ.com*, December 20, 2016, <http://kjzz.org/content/10529/autopsy-guatemalan-detainee-who-died-arizona-died-blood-clot>.
9. Dan Moshenbourg, "Raquel Calderon de Hildago Did Not Die. We Killed Her." *Women In and Beyond the Global*, December 6, 2016, <http://www.womeninandbeyond.org/?p=20723>.
10. Shannan Van Sant, "Teenager Is Latest Migrant Child to Die in US Custody," *National Public Radio*, May 20, 2019, <https://www.npr.org/2019/05/20/725117838/teenager-is-latest-migrant-child-to-die-in-u-s-custody>; Dara Lind, "The Crisis of Children Dying at the Border, Explained," *Vox*, May 22, 2019, <https://www.vox.com/2019/5/22/18632936/child-died-border-toddler-patrol-three-five>.
11. Irwin Redlener, "A 16-Year-Old Boy Died in CBP Custody. Blame Immigration Policy," *Washington Post*, May 23, 2019, https://www.washingtonpost.com/opinions/dont-blame-border-patrol-for-child-deaths-at-the-border-blame-immigration-policy/2019/05/23/f9477a06-7d76-11e9-8bb7-0fc796cf2ec0_story.html.
12. AILA, "Deaths at Adult Detention Centers," *AILA.org*, accessed July 2, 2019, <https://www.aila.org/infonet/deaths-at-adult-detention-centers>.
13. Human Rights Watch, "US: Poor Medical Care, Deaths, in Immigrant Detention," June 20, 2018, <https://www.hrw.org/news/2018/06/20/us-poor-medical-care-deaths-immigrant-detention>.
14. *Ibid*; Meera Senthilingam, "Half of Recent Immigrant Detainee Deaths Due To Inadequate Medical Care, Report Finds," *CNN*, June 20, 2018, <https://www.cnn.com/2018/06/20/health/immigrant-detainee-deaths-medical-care-bn/index.html>; Patricia Sullivan, "Report: Half of Recent Immigrant Detainee Deaths Due to Inadequate Medical Care," *The Washington Post*, June 20, 2018, https://www.washingtonpost.com/local/immigration/report-half-of-recent-immigrant-detainee-deaths-due-to-inadequate-medical-care/2018/06/19/fe313670-73df-11e8-b4b7-308400242c2e_story.html.
15. National Immigrant Justice Center, "Policy Brief: ICE's Failed Monitoring of Immigration Detention Contracts," *ImmigrantJustice.org*, October 1, 2016, <https://www.immigrantjustice.org/research-items/policy-brief-ices-failed-monitoring-immigration-detention-contracts>.
16. Patricia Macías-Rojas, *From Deportation to Prison: The Politics of Immigration Enforcement in Post-Civil Rights America* (New York: New York University Press, 2016), 17.
17. *Ibid.*, 53.
18. *Ibid.*, 59.
19. *Ibid.*, 59.
20. *Ibid.*, 61.
21. *Ibid.*, 63.
22. *Ibid.*, 69.
23. Lisa Riordan Seville et al., "22 Immigrants Died in ICE Detention Centers During the Past 2 Years," *NBCNews.com*, January 6, 2019, <https://www.nbcnews.com/politics/immigration/22-immigrants-died-ice-detention-centers-during-past-2-years-n954781>.
24. Stephanie Silverman, "The Difference That Detention Makes: Reconceptualizing the Boundaries of the Normative Debate on Immigration Control," in *The Ethics and Politics of Immigration: Core Issues and Emerging Trends*, ed. Alex Sager (London and NY: Rowman & Littlefield International, 2016), 108.
25. *Ibid.*
26. OAS, "IACHR Expresses Deep Concern over the Deaths of Two Migrant Children in the Custody of Immigration Authorities in the United States," *OAS.org*, January 7, 2019, http://www.oas.org/en/iachr/media_center/PReleases/2019/001.asp.
27. The Office of the High Commissioner of the United Nations, "The Economic, Social, and Cultural Rights of Migrants in an Irregular Situation," *United Nations* (New York and Geneva, 2014), https://www.ohchr.org/Documents/Publications/HR-PUB-14-1_en.pdf.
28. US Immigration and Customs Enforcement, "ICE Health Service Corps," *ICE.gov*, last updated February 26, 2019, <https://www.ice.gov/ice-health-service-corps>.
29. Human Rights Watch, "US: Poor Medical Care, Deaths, in Immigrant Detention," *HRW.org*, June 20, 2018, <https://www.hrw.org/news/2018/06/20/us-poor-medical-care-deaths-immigrant-detention>.
30. United Nations, *Universal Declaration of Human Rights*, 1948.
31. Joseph Carens, *Ethics of Immigration* (New York: Oxford University Press, 2013).

32. Marilyn Frye, "Oppression," *The Politics of Reality* (Trumansberg, NY: Crossing Press, 1983), 2.
33. *Ibid.*, 3.
34. National Immigrant Justice Center, "ICE's Failed Monitoring of Immigration Detention Contracts," *National Immigrant Justice Center Policy Brief*, September 2016, <https://www.immigrantjustice.org/sites/default/files/content-type/research-item/documents/2016-11/Inspections%20Policy%20Brief%20FINAL2%202016%2010%2003.pdf>, 1.
35. *Ibid.*
36. Human Rights Watch, ACLU, National Immigrant Justice Center, and Detention Watch Network, "Code Red: The Fatal Consequences of Dangerously Substandard Medical Care in Immigration Detention," *Human Rights Watch*, June 20, 2018, https://www.aclu.org/sites/default/files/field_document/codereportdeathsicedetention.pdf, 45.
37. *Ibid.*, 13; Riordan Seville et al., "22 Immigrants Died in ICE Detention Centers During the Past 2 Years."
38. Human Rights Watch, ACLU, National Immigrant Justice Center, and Detention Watch Network, "Code Red," 2.
39. Silverman, "The Difference That Detention Makes."
40. Sharita Gruberg, "ICE Officers Overwhelmingly Use Their Discretion to Detain LGBT Immigrants," Center for American Progress, October 16, 2016, <https://www.americanprogress.org/issues/lgbt/reports/2016/10/26/291115/ice-officers-overwhelmingly-use-their-discretion-to-detain-lgbt-immigrants/>.
41. Patrisia Macías-Rojas, *From Deportation to Prison: The Politics of Immigration Enforcement in Post-Civil Rights America* (NY: New York University Press, 2016).
42. Lydia Wheeler and Raphael Bernal, "Sessions Orders 'Zero Tolerance' Policy at Southwest Border," *The Hill*, April 2, 2018, <http://thehill.com/regulation/administration/381991-sessions-orders-zero-tolerance-policy-at-southwest-border>.

tactics utilized to preserve historically shaped structures and systems upholding racial stratification. In this "post-racial era," it is necessary to investigate those places masquerading as "race-neutral" that perpetuate racism and racial injustice. Nowhere is this clearer than in the "immigration debate."

The immigration debate is immersed in a history riddled with racist practices and projects of racialization. In the "post-racial era," the immigration debate has dropped overt racism used to define citizenship and has instead proposed "race-neutral" policies to legally justify the categorizations of "citizen" and "migrant." Overtly racial tropes have been replaced by tropes of lawfulness, concealing the racism of anti-immigration discourse behind the veil of terms like "illegal immigrants."⁵ Post-1965 immigration policies have not only participated in engendering the drastic increase of dangerous clandestine methods of immigration from the Global South; they have also continued fortifying the racial border outlining citizenship.⁶ "Race-neutral" or "non-racist" immigration policy reforms have preserved the racial categorization used to facilitate the inclusion of legitimate members of this political community (citizens) and the exclusion of those who are deemed not worthy of entry ("illegal immigrants"). As Grant Silva succinctly remarks, "The immigration debate cannot be excised from discussion of racism, nor can racism be excised from the immigration debate. Tropes of lawfulness and 'waiting in line' are but red herrings."⁷

In accordance with Bell's thesis on the permanence of racism, I argue that "non-racist" immigration policy reform does little more than contribute to the "myth of postraciality." When racism and racial prejudice remain unaddressed and unchallenged, as I argue they are in what passes for the immigration debate in the US today, "preventive" strategies towards justice, such as those offered by liberal political philosophers like John Rawls, fail to adequately respond to the needs of racialized migrants. The insistence on overlooking race and racism while responding to societal injustices, like those produced by this nation's immigration system, further perpetuates racial injustice by invalidating lived experiences of racialized minorities and diverting attention away from racism.⁸

This essay first explains how whiteness remains a requirement for full-status citizenship despite the formal outlaw of the "whiteness clause." Overtly racist immigration legislation—stretching from the 1790 Naturalization Act until the 1952 Immigration and Nationality Act—enacted and endorsed whiteness as a requirement for citizenship.⁹ Nevertheless, following the 1952 and 1965 immigration reforms, which mark the legal abolition of both the "whiteness clause" and the national origins quota, whiteness remains a formative identity category in US society. It is a standard for inclusion used for determining which groups are worthy of receiving the rights and protections guaranteed by citizenship. Following Bell's thesis on the permanence of racism and Marx's criticism of the limits of "political emancipation" in his essay, "On the Jewish Question," this section offers a criticism of "non-racist" efforts seeking to eliminate racial injustice strictly through the legal outlawing of racism in immigration policy. "Non-racist" immigration legislation

A Negative Path Towards Anti-Racist Immigration Policy

Alan Chavoya
NORTHWESTERN UNIVERSITY

Despite claims that the United States entered a "post-racial era"¹ and that racism is a thing of the past, racism has permeated US society in nuanced manners that have resulted in further racial stratification.² As David Theo Goldberg argues, "What the claim about postraciality as the end of race suggests, rather, is simply that a certain way of thinking about race, and implicitly of racist expression, has been giving way to novel understandings, orders, and arrangements of racial designation and racist expression."³ While there have undoubtedly been noticeable legal shifts away from overt racism, these legal shifts, as critical race theorist Derrick Bell expressed, are merely "temporary 'peaks of progress'" and not the end of racism.⁴ Bell's "racial realism" and his thesis on the permanence of racism is particularly illuminating for understanding the limits of legal reform for combating racism, especially the legal conundrum we face in this "post-racial era": policy and legal reform are evidently limited in combating racism, yet policy and law are needed to respond to racial injustices.

Rather than liberating society from racism, post-racialism and post-racial legal reform have functioned as obfuscating

remains insufficient for successfully responding to the racial injustices suffered by migrants of color.

The preceding critique of the inefficacy of “race-neutral” immigration policy allows for a return towards the question of justice. I ask: What would be *justice* for racialized migrants? My goal is to criticize how the typical path taken by political philosophers when entertaining questions of justice—namely, the positive path as proposed by Rawls—begins from a position of a well-ordered society inhabited by equal citizens and therefore lends itself to a “post racial” narrative. As an alternative, I argue that the “negative path,” one that begins from the experiences of injustice and inequality, as proposed by Luis Villoro, Naomi Zack, and Charles Mills, among others, is better suited to respond to the needs of racialized migrants.¹⁰ The former’s insensitivity to the present conditions of racial injustice in the immigration debate marks a dead end, while the latter indicates a much more viable and racially conscious path for responding to the needs of migrants suffering from racial injustice.

By way of a conclusion, I highlight the need for explicitly *anti-racist* immigration policy. I am not attempting to offer a solution for the racism plaguing the immigration debate. While policy and law are necessary, they are not sufficient conditions for eliminating racism. I am instead suggesting that efforts at the level of law and policy aimed at responding to the needs of racialized migrants must account for their lived experiences by being anti-racist. While I agree with Bell’s skepticism of legal reform, I interpret his criticism as applicable to “non-racist” policy. In this sense, anti-racist policy offers a viable path towards actual racial justice that will not result in mere “peaks of progress.”

PART I: ON THE LIMITS OF NON-RACIST IMMIGRATION POLICY

Marx’s essay, “On the Jewish Question,” provides a powerful critique of political emancipation within a community that elucidates the limited efficacy of non-racist immigration policy. Regarding the secularization of religion, Marx states, “The most flexible form of the opposition between Christian and Jew is the religious opposition. How is an opposition to be done away with? By making it impossible. How does one make a religious opposition impossible? By abolishing religion.”¹¹ By “making opposition impossible” and “abolishing religion,” Marx is not suggesting a superficial approach, which he defines as “political emancipation.” Marx’s scathing critique of political emancipation is best summarized when he states, “The limitations of political emancipation are immediately evident in the fact that a state can *liberate* itself from a limitation without man himself being truly free of it, and the state can be a free state without man himself being a free man.”¹² Thinking about how racism in immigration policy has prevailed despite its formal overrule, Marx’s critique of political emancipation can become a critique of non-racist immigration policy as merely a type of this limited political emancipation.

Legally, the US has “liberated itself” from racism by making its overt forms unconstitutional, but the political members

of the US, who are the ones legislating, upholding, and deciding rulings on immigration policy, have yet to be liberated from racism. Just as the state Marx is writing about emancipates itself from religion while its overwhelming majority does not cease to be religious, so too has the US immigration system emancipated itself from racism. It is in this way that non-racist immigration policy has assisted whiteness, which, in a “post-racial-era,” is allowed to remain publicly innocent while preserving the legacies of racism that have and will continue to construct it. The more “migrant” is defined in terms of illegality, the more “citizen” becomes an innocent category.

Critical race theorists Ian Haney-López and Cheryl Harris are particularly insightful in their understanding of how legal reform and policy have both arbitrated the racialization of different categories and concretized racial stratification in the US. In regards to the racial formation of whiteness, Harris argues, “The law’s construction of whiteness defined and affirmed critical aspects of identity (who is white); of privilege (what benefits accrue to that status); and, of property (what *legal* entitlements arise from that status).”¹³ Thus, law has functioned as a tool that defines racial categories and dictates the privilege or disadvantage of different racial identities. Immigration law, in particular, has displayed an affinity towards these projects of racialization in the US, crafting a border between lawful, legitimate members of the political community categorized as “citizen,” and those who are excluded and without rights: “illegal aliens” or “impossible subjects.”¹⁴

From the genesis of immigration legislation in the US, whiteness and citizenship have been legally synonymous. The 1790 Naturalization Act is among the earliest instantiations of immigration policy to legally mandate whiteness as an official requirement for naturalization, which José Jorge Mendoza notes, “created an entire subclass of people (i.e., non-white persons) who were permanently ineligible for US citizenship, but not necessarily denied residency.”¹⁵ For the better part of the next two centuries, whiteness operated legally as a determinant for whom privileges and protections granted by citizenship could be extended. For the duration of this period, whiteness and policy dynamically shifted as new groups of migrants who did not neatly fit the black-white racial paradigm posed a challenge to the previously established racial categories in citizenship law.¹⁶ As new groups of migrants entered the racial discourse, different policies were enacted to more rigidly define this border between citizens and non-citizens.¹⁷ Through elucidating investigations into the history of immigration policy and its overt relationship with whiteness from 1790 to 1965,¹⁸ two observations can be made: 1) whiteness is a dynamic racial identity whose meaning is modified through ever-changing social practices,¹⁹ and 2) immigration law is one of these social practices that has continuously responded to the needs of racially privileged groups in the US.²⁰

In the wake of the civil rights movement, the 1952 and 1965 Immigration and Nationality Acts, two pieces of non-racist policy, have shaped the current Immigration and Nationality Act as they ended overt forms of racism and ethnocentrism in immigration policy. Along with

the abolition of racially discriminatory policies, the 1965 immigration reform introduced a yearly numerical cap of 20,000 migrants-per-year that was *equally* applied to every country. Both Ngai and Mendoza agree that the enactment of this cap disproportionately affected Mexican migrants, who, in the years prior to the cap, constituted over 400,000 of temporary migrants in the US.²¹ In turn, the increase of unauthorized migrants, particularly Mexicans and Central Americans, necessitated the policing and surveillance of the US-Mexico border and the neighborhoods that appeared to harbor these migrants, i.e., Mexican neighborhoods, since according to Fox News, Guatemala, El Salvador, and Honduras are Mexican countries. These typically forceful policing tactics have become normalized within the immigration system and have reinforced the border in a manner that has, according to Carlos Alberto Sánchez, intensified the racial definition of the categories “citizen” and “illegal,” the former representing whiteness, inclusion, and the right to have rights, while the latter represents racialized migrants who are excluded from having rights.²²

Current work in crimmigration law exemplifies the justification for these tactics by demonstrating how criminal law and enforcement has converged with immigration law and enforcement.²³ This convergence has resulted in the further criminalization of migrants of color and has provided a clearer path for rampant deportations, ICE raids, and indefinite detention.

Racially neutral immigration policy not only contributed to the increased policing of Mexicans or those who “look Mexican” in the US, but it also marked a shift in which “illegal immigrants” became targets of racist anti-immigration discourse and tactics. The current version of the Immigration and Nationality Act provides racist sentiment a kind of camouflage in the form of tropes of lawfulness, which still dominate much of the present-day immigration debate. In this “post-racial era,” anti-illegalist sentiment has largely replaced the overtly racist sentiment against migrants, but it has nevertheless allowed the racism permeating the immigration system to prevail. This phenomenon is exemplified in current dog-whistle politics, which criminalizes migrants and veils racist sentiment behind cries for national security, despite empirical data showing that migrants are less likely to commit crimes than US-born citizens.²⁴

Another consequence of opting for “non-racist” immigration policies in the “post-racial era” is the ability granted to racially privileged groups to dissociate themselves from their histories of racism, that is, from their facticity or thrownness. Regardless of immigration reforms, immigration law is not displaying any linear “progress” towards something like a just immigration system. By purchasing into the linearity of progress, history is understood as something that can be left behind; something that, at best, is considered to minimally affect present circumstances.

“Post-racialism” offers an understanding of historically racist practices against migrants, making racism a thing of the past that no longer exists in the immigration system. “Non-racist” immigration policy assumes this position and only serves to reinforce the illusion that racism in the

US immigration system perished following the 1952 and 1965 reforms. This position, which can be categorized as a disinterest in or dismissal of history, is exemplified by Michael Walzer’s work, which has significantly impacted much of the philosophical discourse on the ethics of immigration. While discerning how *equal* members of a political economy distribute power to one another, Walzer calls for an investigation into how the group is constituted, a task he clarifies by saying, “I am concerned here not with the historical origins of the different groups, but with the decisions they make in the present about their present and future populations.”²⁵ For Walzer, all the community’s past wrongs do not factor into the consideration for how membership will be distributed. As Silva’s own analysis of Walzer points out, Walzer does take issue with racism and racist policies, but the dismissal of history limits how far we can challenge members of the political community in justifying their existence. When this dismissal occurs, “racial normativity is afforded opportunity to reside in the historic fabric of a political community’s existence.”²⁶ The legacies of colonialism and racism remain imperceptible when a political community’s decision to extend or limit membership status is only oriented towards the future and fails to consider how its past shapes its present and future. Non-racist immigration policy is, at best, superficially concerned with legacies of colonialism and racism. Its emphasis on formally overruling overt forms of racism in the immigration system is insufficient for eradicating the institutional racism that facilitates the conflation between whiteness and citizenship, and migrant and criminal. Through non-racist immigration law, the US has only become “politically emancipated” from racism. The racism in the immigration system remains and more effectively targets migrants of color.

PART II: TOWARDS ANTI-RACIST IMMIGRATION POLICY

In what follows, I explore two possible paths towards justice for racialized migrants: the positive and the negative path. After critiquing the former and arguing for the latter, this section concludes by suggesting anti-racist immigration policy as an alternative to non-racist policy. The return to the level of policy is necessary because, as Cherríe Moraga and Gloria Anzaldúa advise us, “Theory alone *can not* wipe out racism. We do not experience racism, whether directed at ourselves or others, theoretically.”²⁷

Luis Villoro’s “justice” chapter in *Tres retos de la sociedad por venir: justicia, democracia, pluralidad* offers one of the clearest and briefest articulations for perceiving the difference between what will be referred to as the positive and the negative paths towards justice.²⁸ Concerning the point of departure for the positive path, Villoro states, “The theories most in vogue tend to proceed from the ideas of a rational consensus between equal subjects who relate to one another on terms that reproduced the characteristics of a well-ordered democracy.”²⁹ For societies embarking on the negative path, “Instead of proceeding from the consensus for establishing justice, [they] proceed from its absence; rather than moving from the determination of universal principals of justice towards its realization in a specific society, [they] proceed from the perception of

real injustice in order to project what could remedy it.³⁰ While it is possible for the positive path and the negative path to reach the same conclusion, more often than not, they fail to do so because, as Villoro has laid out, they are responding to different needs. The positive approach desires the preservation of equality, or the prevention of unfair inequalities, among the members of its “well-ordered democracy,” and the negative approach to justice seeks to remedy the injustices plaguing a particular society. Without much debate, it is fairly obvious that the negative approach is more viable for actualizing justice for racialized migrants, for this approach is actually sensible to their experiences.

Through critical readings of Rawls’s ideal theory of justice, one of the most popular positive approaches, it becomes quite clear how such a theory of justice cannot graph onto the realities of migrants of color. Rawls’s ideal theory of justice assumes a “well-ordered” just society, making it immediately difficult to imagine how this theory can be bridged with real injustices. In critiquing Rawls’s omission of race, Charles Mills articulates the meaning of the ideally just society in Rawls’s theory of justice as “a society without any previous history of injustice.”³¹ Recalling Walzer’s disinterest in history, Mills’s critique illustrates how Rawls’s ideal theory of justice cannot make sense of the past legacies of racism and colonialism that have shaped the current racial injustices many migrants of color are facing. Rawls’s theory is incompatible with the lived experiences of racialized migrants. The futural focus on how those people behind the veil of ignorance *will* enter a just society signifies that Rawls’s theory is only concerned with ensuring that the presently “just” state of this society will be preserved in the future. The belief that there is any society that is actually just in the sense that Rawls assumes is an illusion engendered by those racially privileged groups whose facticity corresponds to over five centuries of colonialism.

The positive approach’s dependence on formal equality only extends to groups who are deemed worthy of consideration for access to full-membership status in the political community. Hence, passionate discourse around “equality for all” should always be interpreted as “equality for all lawful members of this particular political community.” Historically, the understanding of “equality for all” has remained true in the US. As the previous section demonstrated, only those who are able to become lawful US citizens are worthy of being treated equally. It is from this understanding of “equality” that immigration policymakers outlawed overt racism. According to Zack, “Good government, as well as bad government, has a dynamic material foundation, in the customs of a people, traditions, ongoing institutions, and whether or not officials on all levels obey just laws.”³² The ideality of equality and justice correspond to the material conditions that demand just laws. However, it should be added that only the conditions that matter to *citizens* are worthy of consideration. Thus, the needs of migrants of color are not accounted for.

Contrasting the positive approach, the negative approach to justice begins from concrete cases of injustice. If we are interested in understanding justice for migrants, the negative approach makes the racial injustices of migrants

of color the point of departure. Villoro’s analysis of the negative path towards justice is insightful for imagining migrant justice. Villoro highlights the relationship between exclusion, injustice, and justice. He indicates that it is only through the concrete experiences of a particular injustice that the constitution of the moral subjects of a particular political community can be reimagined to make the political community more inclusive. Hence, for Villoro, justice is not simply a perennial concept that must be theorized and contemplated before being actualized in any given community at any point in history.³³ Justice is instead a response to a concrete injustice that renders unjustly excluded groups new members of the political community worthy of full moral consideration. This expansion of full moral consideration resonates with Zack’s applicative justice, which is defined as the “extension of existing practices of justice to members of new groups.”³⁴ Both approaches not only allow for novel imaginings of what justice for racialized migrants will be, but they also provide blueprints for developing policy that can adequately respond to the injustices that migrants of color face. In the next section, I envision what this type of policy can be.

CONCLUSION: IMAGINING ANTI-RACIST IMMIGRATION POLICY

In responding to the facticity of the people who are excluded unjustly, the negative path of justice demands that any type of response, in this instance, policy, take into consideration the present situation as one that is historically informed and shaped. Thus, the history of unjustly excluded groups must also be accounted for by responses to their situations. What would this mean for migrants of color? In one sense, it would suggest that non-racist approaches to policy do not suffice for actualizing something like migrant justice, since non-racist immigration policy fails to account for the legacies of colonialism and racism that have shaped this political community by building a racial border around its members. Policy that does account for the facticity of this situation is a bit challenging to define, for any effort to do so must also account for the dynamic material conditions to which it responds. The type of immigration policy that could respond to the lived experiences of migrants of color is what I imagine *anti-racist* immigration policy would be.

Crimmigration law is a promising place where anti-racist immigration policy can be successfully executed in the manners previously described. Throughout the essay, I have pointed out how non-racist immigration policy has allowed racism to masquerade as anti-illegalist sentiments. This is particularly obvious in the current enforcement of disproportionate detention and deportation of migrants of color.³⁵ Both practices are defined as civil or administrative mechanisms designed to effectuate the government’s immigration enforcement goals. Hence, the right to due process that is afforded to people treated under the criminal system is not required for migrants who are indefinitely detained and deported—the majority of whom are racialized migrants from Latin America and the Caribbean.³⁶ Often, detained migrants are simply handed a piece of paper and deported without a trial. Enforcement tactics are not motivated by explicitly racist legislature. The amendments to the Immigration and Nationality Act of

1996, the creation of the Department of Homeland Security following the events of September 11, 2001, and the recent policy designed to end the so-called “catch-and-release” practices were justified by appealing to concerns for national security.

At the level of policy, two things should occur: 1) a rejection of the national security justification for practices like mass detention and deportation; and 2) an extension of rights and protections to irregular migrants to ensure a just process while detained. Towards the first point, the burden of proof would be shifted towards the system to demonstrate that it is not being racist and that migrants actually pose a threat to national security, rather than simply being allowed to use dog-whistles like rapists, or MMA fighters, when describing migrants. As the system presently exists, the burden of proof is on migrants to prove they are not bad ones. Instead, the type of policy I am suggesting turns that around and forces the justifications of these policies to show they are not racist. The burden of proof must also be significant enough so that a handful of cases are not used to generalize all migrants. If this shift occurs, the current administration will have greater difficulty justifying its actions against migrants since the empirical data indicates that over three quarters of detained migrants pose no threat to national security and that migrants are less likely to commit violent crimes than US-born citizens.³⁷

Regarding the second point—the extension of rights and protections—I think it is necessary to do two further things: 1) end mass detention and deportation as the primary tactics for enforcing immigration-related purposes; and 2) eradicate indefinite detention. These two moves could help significantly reduce the population inside detention facilities and would alleviate the caseload for immigration judges. I call for these approaches because many ICE, CBP agents, etc., contend that they are understaffed and underfunded. I consider these excuses unacceptable for the conditions inside detention centers and the manners in which migrants are apprehended. However, if there are too many migrants, then reducing the population would mean they have to come up with other ways to justify their actions. In this sense, anti-racist immigration policy is a tactic that prohibits racism from hiding behind cries for protecting the nation from criminals. Obviously, these are very crude estimations for what anti-racist immigration policy could be. However, as a basis for anti-racist immigration policy, I think it is enough to motivate the purpose of this type of policy as shedding light on the various ways racism is housed in the immigration system. From here, we can begin tearing down the racism that is at play in this nation’s immigration system.

ENDNOTES

1. The conceptualization of US society becoming “post-racial” or reaching a state of “post-racism” has made its way into racial discourse following the presidential election of Barack Obama. “Post-racialism” claims that this society is or is close to being outside of racial reference. Following the election of Obama, many have been quick to remark that such an election serves as evidence that race no longer matters in the US. This essay follows the similar skepticism of “post-racialism” that one finds in the works of Ian Haney-López, Grant Silva, David Theo Goldberg, Khalil Gibran Muhammad, Michelle Alexander, and others. See Ian Haney-López, *Dog Whistle Politics: How Coded Racial Appeals*

Have Reinvented Racism and Wrecked the Middle Class (New York: Oxford University Press, 2014); Haney-López, “Post-Racial Racism: Racial Stratification and Mass Incarceration in the Age of Obama,” *California Law Review* 98, no. 3 (June 2010): 1023–74; Haney-López, *White by Law: The Legal Construction of Race* (New York: New York University Press, 2006); Grant J. Silva, “Embodying a ‘New’ Color Line: Racism, Anti-Immigrant Sentiment and Racial Identities in the ‘Post-Racial’ Era,” *Knowledge Cultures* 3, no. 1 (2015): 65–90; David Theo Goldberg, *Are We All Postracial Yet?* (Cambridge, UK: Polity Press, 2015); Khalil Gibran Muhammad, “Where Did All the White Criminals Go?: Reconfiguring Race and Crime on the Road to Mass Incarceration,” *Souls* 13, no. 1 (2011): 72–90; and Michelle Alexander, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness* (New York: The New Press, 2010).

2. The concept of racial stratification in this paper is adopted from thinkers like Ian Haney-López and Eduardo Bonilla-Silva. Haney-López and Bonilla-Silva utilize sociological concepts about social stratification in order to understand how race facilitates the “strategic misallocation of resources.” See Haney-López, “Post-Racial Racism: Racial Stratification and Mass Incarceration in the Age of Obama,” and Eduardo Bonilla-Silva, *Racism without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America*, 4th ed. (Lanham: Rowman & Littlefield, 2014).
3. Goldberg, *Are We All Postracial Yet?*, 6.
4. Derrick Bell, “The Racism is Permanent Thesis: Courageous Revelation of Unconscious Denial of Racial Genocide,” *Capital University Law Review* 22, no. 3 (1993): 576.
5. Grant J. Silva, “Embodying a ‘New’ Color Line,” 68. As Silva points out, “The claim is often made that although one does not have a problem with immigration per se, the problem is ‘illegal’ immigration; the problem is with people not waiting in line.”
6. Here, the following two pieces of policy come to mind: the 1996 Anti-Terrorism and Effective Death Penalty Act and the Illegal Immigration Reform and Immigrant Responsibility Act. As Tanya María Golash-Boza tells us, “these laws eliminated judicial review of some deportation orders, required mandatory detention for many non-citizens, and introduced the potential for the use of secret evidence in certain cases.” See *Deported: Immigrant Policing, Disposable Labor, and Global Capitalism* (New York and London: New York University Press, 2015), 105. These two pieces of legislature, which do not overtly discriminate on the basis of race, mark a shift in policy and enforcement towards an overwhelming employment of detention and deportation. These practices have disproportionately affected migrants of color.
7. Silva, “Embodying a ‘New’ Color Line,” 67.
8. Think here about arguments made against people of color that what they perceive as racial injustice is really an issue about class, not race. Not only do these arguments wrongfully negate the lived experiences of people of color, they also preserve an erroneous understanding of race and class as mutually exclusive categories.
9. See, generally, Mae G. Ngai, *Impossible Subjects: Illegal Aliens and the Making of Modern America* (Princeton, NJ: Princeton University Press, 2004); José Jorge Mendoza, “Illegal: White Supremacy and Immigration Status,” in *The Ethics and Politics of Immigration: Core Issues and Emerging Trends*, ed. Alex Sager (London UK: Rowman & Littlefield International, 2016), 201–20; Ian Haney-López, *White by Law: The Legal Construction of Race*.
10. While Zack and Mills do not explicitly refer to their non-ideal theories of justice (“applicative justice” and “rectificatory justice”) as a “negative path,” their point of view is quite similar to Luis Villoro’s, who refers to his theory as a “negative path towards justice.” Therefore, I decide to refer to Zack’s, Mills’s, and Villoro’s approaches to justice as the negative path.
11. Karl Marx, “On the Jewish Question,” in *Selected Writings*, ed. David McLellan, (Oxford: Oxford University Press, 2000), 48.
12. Marx, “On the Jewish Question,” 51.
13. Cheryl I. Harris, “Whiteness as Property,” *Harvard Law Review* 106, no. 8 (1993): 1725.
14. Ngai, *Impossible Subjects*, 4-5. According to Ngai, “Immigration restriction produced the illegal alien as a new legal and political subject, whose inclusion within the nation was simultaneously

a social reality and a legal impossibility—a subject barred from citizenship and without rights. Moreover, the need of state authorities to identify and distinguish between citizens, lawfully resident immigrants, and illegal aliens posed enforcement, political, and constitutional problems for the modern state. The illegal alien is thus an 'impossible subject,' a person who cannot be and a problem that cannot be solved."

15. Mendoza, "Illegal: White Supremacy and Immigration Status," 208.
16. Ngai, *Impossible Subjects*, 38.
17. As both Ngai and Mendoza point out, the black-white racial binary could not account for Chinese, Japanese, and Indian immigrants, and upon questioning their ability to assimilate, different policies (Chinese Exclusion Act of 1882) and Supreme Court rulings [*Takao Ozawa v. US* (1922) and *US v. Bhagat Singh Thind* (1923)] rendered these groups unworthy of US citizenship.
18. See Ngai, *Impossible Subjects*; Mendoza, "Illegal"; and Haney-López, *White by Law*.
19. See Linda Martín Alcoff, *Visible Identities: Race, Gender, and the Self* (Oxford: Oxford University Press, 2006).
20. Haney-López, *White by Law*, 7.
21. Mendoza, "Illegal," 216; Ngai, *Impossible Subjects*, 260-61. The figure of 400,000 may represent an undercount for the number of temporary migrants from Mexico in the US at the time.
22. Carlos Alberto Sánchez, "'Illegal' Immigrants: Law, Fantasy, and Guts," *Philosophy in the Contemporary World* 21, no. 1 (2014): 101-103.
23. See Juliet Stumpf, "The Cimmigration Crisis: Immigrants, Crime, and Sovereign Power," *American University Law Review* 56, no. 2 (2006): 367-419; Jennifer Chacón, "Overcriminalizing Immigration," *The Journal of Criminal Law & Criminology* 102, no. 3 (2012): 613-52; César Cuauhtémoc García Hernández, *Cimmigration Law* (Chicago: American Bar Association Publishing, 2015); and Mendoza, "The Contradictions of Cimmigration," *APA Newsletter on Hispanic/Latino Issues in Philosophy* 17, no. 2 (2018): 6-9.
24. See Christopher N. Lasch, "Sanctuary Cities and Dog-Whistle Politics," *New England Journal on Criminal and Civil Confinement* 42, no. 2 (2016): 159-90.
25. Michael Walzer, *Spheres of Justice* (New York: Basic Books, 1983), 31.
26. Silva, "Embodying a 'New' Color Line," 78.
27. *This Bridge Called My Back: Writings by Radical Women of Color*, edited by Cherríe Moraga and Gloria Anzaldúa (New York: Kitchen Table: Women of Color Press, 1983), 61-62.
28. Luis Villoro, *Tres retos de la sociedad por venir: justicia, democracia, pluralidad* (México: Siglo XXI, 2009).
29. *Ibid.*, 12. Translation from Spanish to English by author.
30. *Ibid.*, 13.
31. Charles Mills, *Black Rights/White Wrongs: The Critique of Racial Liberalism*, (New York: Oxford University Press, 2017), 140.
32. Naomi Zack, "Starting from Injustice: Justice, Applicative Justice, and Injustice Theory," *Harvard Review of Philosophy* 24 (2017).
33. Villoro elucidates this point through his interpretation of the following three historical moments in which personal experience of unjust exclusion resulted in the development of theories of justice: 1) Bartolomé de Las Casas experienced the exclusion of the Indians in America, allowing him to postulate an equitable treatment of the colonized as a new moral subject; 2) John Locke experienced exclusion in the form of religious intolerance, enabling him to project a new moral subject that will include the virtue of tolerance; 3) regarding political power, eighteenth-century French revolutionaries experienced the exclusion of the Third State, propelling them to construct the universal subject of human rights as the new moral subject. Villoro, *Tres retos de la sociedad por venir*, 23-25.
34. Zack, "Starting from Injustice."
35. See Golash-Boza and Pierrette Hondagneu-Sotelo, "Latino Immigrant Men and the Deportation Crisis: A Gendered and Racial Removal Program," *Latino Studies* 11, no. 3 (2013): 271-92.

36. See Allison Cunneen, "Demanding Due Process: Time to Amend 8 U.S.C. §1226(C) and Limit Indefinite Detention of Criminal Immigrants," *Brooklyn Law Review* 83, no. 4 (2018): 1497.

37. See R. Sampson, V. Chew, G. Mitchell, and L. Bowring, "There Are Alternatives: A Handbook for Preventing Unnecessary Immigration Detention," (Melbourne: International Detention Coalition Report, 2015), 7-15.

Immigration and International Justice

Jorge M. Valadez

OUR LADY OF THE LAKE UNIVERSITY

Discussions about immigration often center on such issues as the economic liabilities or benefits created by immigrants, the potential threats to national security of poorly protected borders, the cultural impact of increased immigration, or the protection of immigrants' human rights. These discussions generally take for granted that states have the unconditional right to regulate the terms and conditions under which people can enter state territory. The question of what justifies the territorial rights of states is rarely raised. Yet, the justification of the territorial rights of states is of central importance for the moral legitimacy of state immigration policies, which set the terms for long-term admission and residence in particular countries. From a moral standpoint, it is important to inquire about what could justify the practice of states to routinely employ physical barriers and the threat of force to enforce control of their territorial boundaries. The power of exclusion is particularly important because preventing people from poor countries from accessing affluent states can doom them to a life of extreme poverty and deprivation or expose them to harm from rampant organized criminal activity or environmental devastation.

Immigration is an international issue because it involves the regulation of movement of people across national boundaries. Since this movement is regulated by the coercive power of governments, immigration deals with issues concerning moral obligations and relations between states, i.e., issues of international justice. In this essay I maintain that the justification of the territorial rights of states, and the international obligations to which they give rise, are of central importance for identifying immigration policies that address global injustices. My focus will be on immigration policies, which regulate permanent or long-term residence in a country, rather than policies determining asylum, that are presumably designed to protect people from transient or temporary harms, which raise different kinds of moral considerations.

This essay is divided into three parts. In the first part I examine the moral justification of the territorial rights of states within the context of international justice. In the second part I propose some principles of international justice that could be used to justify more open immigration policies strategically designed to address international inequalities and injustices. In the third section I suggest some immigration policy guidelines based on the view of international justice outlined in the second part of the paper.

CHALLENGING THE TERRITORIAL RIGHTS OF STATES

Central to the question of whether states can legitimately exclude people from entering their territory is the issue of the moral justification of their territorial rights. John Simmons identifies state territorial rights as including the following: (1) the right to exercise legal jurisdiction over individuals within state territories; (2) the right to regulate the movement of persons, non-human beings, and materials across state boundaries; (3) the right to control the land and natural resources not privately owned within state territorial boundaries; (4) the right to regulate and tax privately owned land; and (5) the right to prevent the disintegration of the state's territory by prohibiting such actions as territorial transfers to aliens or secession by substate groups.¹ My analysis will center on the second and third territorial rights because, even though interesting questions arise concerning the justification of the other three state territorial rights, the second and third rights are the most relevant for immigration. Hereafter, when I talk about state territorial rights, I will be referring to these two territorial rights.

Focusing on the second and third territorial rights, on what basis can a state justifiably claim that it has the exclusive right to control the land and natural resources that comprise its national territory? If we make the eminently reasonable assumption that every human being should have access to the earth's land and natural resources to satisfy their basic material needs and to flourish, what gives states the power to tell the rest of the world community that they have the exclusive collective right to use and control those sectors of the earth that make up their national territories? States can hardly claim that they have a morally unconditional right to exercise these territorial rights, since it is historically well established that practically all states acquired their territories through conquest, invasive settlement, broken treaties, partitioning between imperial powers, and other morally illegitimate means. Even on those rare occasions in which states might have acquired their territories peacefully and without displacing pre-existing communities, they still cannot claim to have acquired their territorial rights legitimately, since they did not get the consent of the world community to obtain exclusive control of the land and natural resources comprising their national territory.

A central issue here is how to respond to the realization that the territorial rights of states cannot be justified in a morally unconditional way. We could try to reconfigure the existing distribution of land and natural resources among the world's people using some principle of justice that would ensure a fair distribution. This alternative, however, is highly unrealistic and would be practically impossible to implement. People have deep connections to their national territories, and it is utopian to believe that they would be willing to relocate en masse to satisfy some abstract principle of justice. National economic interests in maintaining control of land and natural resources are so strong that we cannot start from scratch and carry out a process of global land and resource redistribution. Moreover, it is not at all clear what principle of justice we could use to carry out this territorial distribution to ensure that it produced a fair

outcome for all of the world's people. Neither would it be easy, or even possible, to determine in a just manner who gets to live in which of the earth's territories. Given the differential value of land and the earth's resources and the historical uncertainty concerning which groups or individuals first occupied which territories, endless disputes would likely arise regarding who can legitimately claim to occupy sectors of the earth's surface. Finally, no global political authority presently exists, or could realistically be developed in the foreseeable future, that could carry out such a redistribution in a way that ensured that it would be universally acceptable and just.

THE CONDITIONAL LEGITIMACY OF STATE TERRITORIAL RIGHTS

Since we cannot start from scratch regarding collective control by political communities of territories, an alternative is to provide a conditional basis for recognizing the illegitimately acquired territorial rights of states. That is, just because the territorial rights of states cannot be justified in a morally unconditional way does not mean that we cannot provide a *conditional justification* for them. If from a realistic standpoint we must live, at least for the foreseeable future, with territorialized states, we could make their legitimacy contingent on their observance of certain principles of international justice. There may be compelling reasons why we may want to recognize the territorial rights of states, *albeit in a way that places strong moral conditions on such recognition*. It is particularly important that these reasons be consistent with the basic egalitarian principle of showing equal moral concern for all people and not merely members of our own state. We would then show that we are not recognizing these territorial powers merely for pragmatic reasons that do not give sufficient moral consideration to the plight of others. Our case for the conditional legitimacy of the territorial rights of states must therefore involve both pragmatic and normative considerations.

Perhaps the best way to begin a conditional recognition of the territorial rights of states is to realize that it is a practical necessity for people to organize and coordinate their activities for the purpose of employing the world's land and natural resources to satisfy their material needs. That is, it is an unavoidable real-world problem for people to find some socially organized way to extract, refine, develop, and employ the earth's land and natural resources. Moreover, economic development occurs within sociopolitical and cultural contexts. Laws governing the ownership and use of property, the terms of legal contracts, the provision of credit and capital, informal norms governing economic transactions, and many other features of an economic system are embedded within particular political communities. States provide the needed stable socio-cultural and legal institutional frameworks within which short- and long-term economic planning and development can occur. And putting aside the practical reasons for recognizing states, perhaps the most important moral reason for the existence of political communities like states is that they enable individuals to treat each other justly and protect one another's basic rights by providing the legal and sociopolitical institutions that allow for the

articulation, contestation, and implementation of human rights. Basic human rights—such as those governing property ownership, health care, free speech, and social welfare—can, and are, in fact, interpreted differently in countries with different sociopolitical cultures. A crucial moral function of self-governing political communities like states is to provide the institutional structures to allow their members to disambiguate, vote on, and enforce their interpretations of basic rights in accordance with their own cultural and sociopolitical traditions.

Furthermore, if the institutional frameworks that make economic activity and political governance possible are to be normatively legitimate, they must be grounded on processes of self-governance that are ultimately accountable to the members of these political communities. The institutions and forms of sociopolitical organization of a society should be subject to revision and adaptation as people respond to the ongoing problems and issues that they face collectively as a political community. Self-governance can be seen as the process through which people in political communities legitimize their collective decisions, coordinate their activities to successfully adapt to changing circumstances, and determine the course of their future. Also involved in self-governance is the capacity of a political community to determine who falls within its jurisdictional reach, who can participate in its decision-making procedures, and to whom its leaders owe democratic accountability.

In the absence of a world government that provides the political, economic, and socio-cultural administrative structures to develop the earth's natural resources, self-governing territorialized political communities such as states seem to be viable alternatives. States could be seen as administrative units designed to provide the economic and sociopolitical structures needed for people to develop the earth's resources and collectively govern their lives.² This is not to say that states are the only or best form of political organization that could fulfill these functions. In the indefinite future alternative forms of political organization could emerge that supplant states. At this point in history, however, despite overstated claims regarding the "demise" of states, they still remain the principal forms of political organization that people employ to govern themselves. And in the foreseeable future there are no feasible plans to dismantle or radically transform states. In fact, important recent developments, such as greater concerns about security and the increase in the value of scarce natural resources, have strengthened state claims regarding ownership of their land and natural resources.³ These developments make it unlikely that territorialized states will disappear anytime soon.

However, in order to function effectively as administrative units to develop the earth's land and natural resources and exercise self-governance, territorialized political communities must be able to regulate entrance and membership. It is very difficult to see how political communities could undertake effective long-term economic planning—like determining a national savings rate to meet future economic needs and contingencies, for example—if they did not have accurate projections of

the size of their population. Rapid and dramatic increases in population would make economic calculations for meeting social needs inaccurate or useless. For instance, sudden increases in population would mandate the rapid development of physical and social infrastructures to meet the needs of the new residents. Even wealthy countries like the US would find it difficult to expand and rebuild their physical infrastructures—including roads, bridges, sewage, transportation, and water and electrical systems—to meet the increased demands. Similarly, strategic developments in education and other social needs would be very hard to reliably carry out without accurate knowledge of the makeup and size of the national population. The problems that would arise from sudden large influxes of immigrants are not mere theoretical possibilities, for there are empirical indications that the numbers of people from developing countries willing to relocate to affluent countries are very great.⁴ Depriving self-governing territorialized political communities of the right to regulate their borders and regulate membership would undermine their administrative capacity in serious ways.

These considerations show that it is important to balance the rights of national communities for self-determination with the rights of the members of the world community to have realistic opportunities to satisfy their basic needs and to flourish. State territorial rights are not absolute or morally unconditional, for they should be constrained by the recognition that they were illegitimately acquired through coercive or unilateral means. What moral obligations follow from the conditional recognition of state territorial rights? I propose that a morally principled way to grant conditional legitimacy to the territorial rights of states is to impose significant international moral responsibilities between states and the world community. States incur a profound moral debt to the world community when the latter recognize their territorial rights despite the morally illegitimate ways in which they were acquired. This moral debt creates a relation of moral reciprocity between states involving international justice obligations. More specifically, states are obligated to morally reciprocate by respecting certain provisions in three areas of international justice dealing with nondomination, compensation, and ecological integrity. These areas of justice jointly constitute a conception of international justice holding among states. In the next section I briefly discuss this conception of international justice and its implications for immigration policies.

IMMIGRATION POLICIES AND INTERNATIONAL JUSTICE

The conception of international justice that I outline here is grounded on the idea that states are bound to the world community, and therefore to one another, by a nondiscretionary relation of moral reciprocity that arises from the recognition by the world community of their territorial rights. The continued recognition of these territorial rights, however, should be seen as conditional on states respecting certain provisions that are crucial for international justice, namely, the nondomination, compensation, and ecological integrity principles, which enable states to treat one another justly in the international

arena. These principles are also designed to support the functions that territorial rights enable states to perform, namely, the development of the earth's land and natural resources within the context of self-governing political communities. *It makes sense to impose these principles for the continued legitimacy of state territorial rights because they support the reasons for recognizing these rights in the first place.*

The nondomination principle requires states to support fair conditions of participation for all other states in the global economy. Given that the economic prosperity of states in the contemporary world depends on their successful participation in the global economy, it is imperative that they be able to participate on fair terms. They must be able to participate in the international economic system without, for example, unfair conditions of trade, finance, or intellectual property ownership imposed by more powerful states, who typically try to employ their greater economic and political strengths to secure unfair advantages. Powerful states, for example, limit developing countries' access to their markets and heavily subsidize some of their domestic industries, making it practically impossible for developing countries to compete against them in the global market.⁵ Even though poor countries may in some cases also subsidize some of their industries, these modest subsidies pale in comparison to the large subsidies of affluent countries, which use their greater economic resources to provide their industries with significant advantages. Since the rationale for recognizing the territorial rights of states is to enable them to develop a portion of a common global resource base, it is just to require that all states support fair conditions for economic development and participation in the global economy. The best way to ensure that this provision is met on an ongoing basis is through the creation of a just global system of trade, production, finance, and intellectual property.

According to the compensation principle, states have an obligation to rectify certain inequalities and resource deficiencies that prevent underdeveloped states from participating successfully in the international economic system. The creation of the existing system of states sometimes involved unjust colonial processes of conquest, forced labor, and resource extraction. These unjust processes have played an important role in the present incapacity of some states and political communities to develop economically and to participate successfully in the global economy. The responsibility falls mainly on those states that benefited from historical exploitative practices. Yet, even in cases in which systematic exploitation was not involved, certain states emerged, as a result of contingent historical processes, with an impoverished resource base. These states still merit economic aid because of their participation in an administratively beneficial global partitioning of land and natural resources in which they emerged with material endowments of relatively lesser value. In either case, the world community has a moral obligation to provide economic and technological aid to impoverished states that, as the result of systematic exploitation or contingent processes, are unable to flourish materially and participate on fair terms in the global economic system. To ensure that such aid is employed

effectively, its provision could be tied to the implementation of noncorruption measures by developing countries.

The ecological integrity principle imposes on states the obligation to refrain from environmentally destructive practices that degrade the common resource base on which we all depend for our survival and well-being. States have the responsibility, according to this principle, to safeguard the ecological integrity of the territories that the world community has entrusted them with. Since territorial rights in effect grant states' proprietary control over those parts of a common biosphere that comprise their territories, these powers could be seen as a form of ecological stewardship. Moreover, states should abstain from practices that have a negative environmental impact on territories outside of their own. Many ecological problems transcend national boundaries and thus the economic practices of states may affect other political communities as well. It is imperative that all states do their part in promoting the ecological sustainability of our common global resource base, particularly if they have been primarily responsible for the environmental degradation of our planet.

Even though all three principles are important for international justice, the nondomination and the compensation principles are particularly important for fair immigration policies. To the extent that it can be established that economically and politically powerful states have undermined the capacity of weaker states to flourish economically and participate effectively in the global economy, policies should be implemented to address these past injustices. Even though, given limitations of space, I cannot make a detailed case for this claim here, it is reasonable to maintain that powerful states have in fact systematically hindered the capacity of poor countries to develop economically.⁶ As we observed earlier, wealthy countries have closed off some of their markets to competition from developing countries and have heavily subsidized some of their industries, to the detriment of the economic progress of the developing world. Moreover, the expropriation of enormous amounts of valuable resources such as gold and silver by European countries from Latin America played a significant role in the former's economic and technological advantages, which they maintain in contemporary times.⁷ Even those countries that did not experience the brunt of colonial domination merit aid, for they may have ended up with resource-poor territories in the process of state formation. Within the context of the theory of global responsibilities that I outlined here, *more liberal immigration policies could play an important role in the economic and social development of poor states.*

Immigration policies that address global inequalities should focus on two priorities: (1) a more liberal approach to immigration in which a greater number of low-skilled workers are admitted and placed on the road to citizenship and (2) temporary migration of low-skilled workers from developing countries, which usually have high unemployment rates. Regarding the first of these priorities, some economically developed countries, particularly in Europe, need the influx of immigrant laborers who are allowed to attain citizenship status. Demographic projections indicate that many affluent countries, due to

declining birthrates and aging populations, will need a significant influx of additional laborers to fill the jobs that will be created in their low-skilled economic sectors.⁸ Even though at present there is considerable political resistance in wealthy countries to liberalizing immigration policies for low-skilled workers, the economic needs of wealthy countries should make these immigration policies more politically feasible.

The remittances sent by immigrant workers to their home countries are a major source of foreign capital for poor and developing countries. In countries such as Mexico, for example, remittances provide a source of foreign currency inflow that is surpassed only by petroleum. In 2017, Mexican laborers working abroad sent home more than \$28.77 billion dollars in remittances, which is the highest amount ever recorded.⁹ Globally, the World Bank estimates that once records are finalized, remittances to low- and middle-income countries will reach \$528 billion dollars in 2018.¹⁰ These figures show that remittances should be considered as a major component of development strategies for the developing world. In addition, those immigrants that return home provide important financial and social capital to their countries. Their knowledge, contacts, and experiences can provide much needed entrepreneurial impetus to their home economies and help improve their country's social and political institutions.

An advantage of temporary labor migration is that it does not undermine self-governance, since states would be able to regulate migrant flows more effectively if they realistically took into account the economic forces that drive the global movement of people. Further, by combining the legalization of foreign workers with strict enforcement of legal penalties for employers hiring undocumented workers, states could more adequately deal with the security problems created by not knowing whom those individuals are who are living within their borders illegally. Temporary migration policies, however, should include provisions that safeguard the human rights of laborers. Temporary worker programs are notorious for the exploitation of workers, and provisions should be put in place to prevent such abuse. Temporary workers have the right to work in the country of destination without fear of being exploited by their employers and others who may take advantage of their vulnerability. Given the long history of the exploitation of temporary workers, we need strong international organizations that can oversee and regulate temporary worker programs and that can impose enforceable penalties on states that violate worker rights.

Finally, we should note that the immigration policies of regulated openness I advocate are more in line with the basic egalitarian principle of expressing equal moral concern for all of the world's people than an open borders position. The realization that we owe a moral responsibility to *all* of the people in poor countries and not merely to those able and willing to immigrate is of fundamental importance for understanding why an open borders policy ultimately fails to respect this egalitarian principle. While the open borders position would help those with the resources and capacities to immigrate, it would not help the most vulnerable people in developing countries who

are left behind, such as the disabled, the ill, the elderly, the very young, and the extremely poor who do not have enough resources to migrate. Further, as observed earlier, equality of opportunity to flourish should be understood within the context of the real-world constraints imposed by the need for self-governing political communities, including choice countries of destination, to effectively exercise their administrative functions. It is plausible to maintain that open border policies would undermine the successful performance of these administrative functions. It is also important to be mindful of the political feasibility of our proposals, and it is surely the case that open border policies are highly unfeasible from a political standpoint. International justice advocates cannot simply ignore the concrete problems that would arise, particularly for choice destination countries, if they embrace unrestricted policies of immigration. Ironically, open border policies would create an open market for immigration slots that would most likely be filled not by the world's poorest and neediest people, but by those from countries of origin who are relatively better off, more highly educated, and more resourceful. Such policies would likely lead to a new order of global inequality, in which wealthy countries benefit from the influx of some of the most talented and resourceful people from developing countries, while poor countries would likely be worse off as a result of losing some of their most capable and educated citizens. We need regulated borders in order to prioritize helping the world's neediest and most vulnerable countries.¹¹

Programs involving more fixed-term work visas would ultimately help to create more economically viable communities in developing countries, particularly if temporary worker programs systematically reinforce the connections between temporary workers and their home communities. Some countries in Latin America, for example, are already implementing such programs, in which immigrant donations to the civic improvement of their communities are matched by state and municipal funds.¹² Returning immigrants could use their capital, knowledge, and connections, as noted earlier, to contribute to the economic and social development of their countries and communities. In short, my position promotes the egalitarian principle of expressing moral concern for all people more than the open borders position, because it works towards a vision of global justice in which *all* human beings are able to flourish in self-governing political communities without having to uproot themselves from their families, friends, and communities in order to achieve a decent material existence.

SUMMARY OF PROBLEMS WITH OPEN BORDERS
 Because the open borders position continues to be embraced by progressive thinkers with the belief that it reflects and promotes the interests and welfare of the world's most disadvantaged people, it is important to be clear about the numerous problems with it. These flaws should be considered as serious by progressive thinkers because they may very well undermine the welfare of precisely the group of people whose welfare they are most concerned to promote.

1. As mentioned in this essay, it is likely that the immigration slots available under open borders (since countries, at some point, would have to place some limits on immigration) would be filled not by the disabled, old, very young, ill, and poor of the world but by those with more resources, connections, education, health, and so forth.
2. Countries would not be able under open borders to strategically design immigration policies that specially favor the world's neediest people. Policies designed to help the most vulnerable and needy would have to involve regulated, not open, borders.
3. The Brain Drain phenomenon, which would likely be exacerbated by eliminating immigration restrictions, negatively affects the social, economic, and political development of some of the poorest countries in the world by facilitating the exodus of some of their most educated, capable, and entrepreneurial citizens.¹³
4. The homelands of Indigenous groups, who have fought long and hard to attain control of their traditional homelands, would be threatened by the uncontrolled influx and eventual control of outsiders. Since Indigenous groups often inhabit ecologically rich lands, there would be strong incentives for outsiders to attain control of resources in their territories. Outsiders, upon arrival, could claim that for the sake of justice and equality they should have equal rights of political representation as members of the Indigenous groups (after all, a principal rationale for open borders is equality).
5. Under open borders, choice countries of destination, even if relatively wealthy, would have trouble developing rapidly the necessary physical infrastructures, such as affordable housing, water purification systems, roads, bridges, dams, and sewage and electrical systems. They would also have difficulty providing the needed social infrastructural needs, such as education, health care, social security, legal services, and police protection for the rapidly increasing population under open borders.
6. Under open borders, countries would not be able to address their past foreign policy decisions that now endanger the citizens of other countries who have helped them with their military and other initiatives. People in Iraq, for example, who cooperated with US personnel are now endangered for their prior cooperation by the groups now in control in Iraq. The US would be unable under open borders to provide preferential treatment for these Iraqis, since under open borders some groups could not be given preference over others. Regardless of what one may think about the original motivations of the relevant US foreign policies, the country owes a moral debt to these Iraqis.

7. A new world order would likely be created under open borders, one with greater inequalities and injustices, as powerful and wealthy countries would benefit from the aptitudes, resources, and initiatives provided by the more talented people from other countries with relatively more resources and education.

Even though the open borders position is motivated by justified concerns about global justice and equality, on close analysis this position would likely not achieve its commendable objectives. We need more carefully structured immigration policies that take into account the complexities of the actual world.

ENDNOTES

1. This definition of state territorial powers is provided by John Simmons in "On the Territorial Rights of States," *Philosophical Issues* 11, no. 1 (2001): 300–26, 306.
2. J. Coleman and S. K. Harding, "Citizenship, the Demands of Justice, and the Moral Relevance of Political Borders," *Justice in Immigration*, ed. W. F. Schwartz (Cambridge: Cambridge University Press, 1995), 18–62.
3. J. Valadez, "The Continuing Significance of Ethnocultural Identity," in *Identities, Affiliations, and Allegiances*, eds. S. Benhabib, I. Shapiro, and D. Petranovic (Cambridge: Cambridge University Press, 2007), 321–24.
4. L. Pritchett, *Let Their People Come: Breaking the Gridlock on Global Labor Mobility* (Washington, DC: Center for Global Development, 2006).
5. See, for example, Thomas Pogge, *World Poverty and Human Rights: Cosmopolitan Responsibilities and Reforms* (Cambridge: Polity Press, 2002), and Branko Milanovic, *Worlds Apart: Measuring International and Global Inequality* (Princeton: Princeton University Press, 2005).
6. Pogge, *World Poverty and Human Rights*.
7. Jorge Valadez, "Pre-Columbian and Modern Philosophical Perspectives in Latin America," in *From Africa to Zen: An Invitation to World Philosophy*, eds. R. C. Solomon and K. M. Higgins (Lanham: Rowman and Littlefield Publishers, Inc., 2003), 86.
8. Pritchett, *Let Their People Come*, 6.
9. The Borgen Project, "Sending Money Home: Record High Remittances to Mexico in 2017," *BorgenProject.org*, May 15, 2018, <https://borgenproject.org/remittances-to-mexico-2017/>, accessed on 3/1/2019.
10. The World Bank, "Accelerated Remittances Growth to Low- and Middle-Income Countries in 2018," *WorldBank.org*, December 8, 2018, <https://www.worldbank.org/en/news/press-release/2018/12/08/accelerated-remittances-growth-to-low-and-middle-income-countries-in-2018>, accessed on March 1, 2019.
11. The reasoning here is that by providing access to the labor markets of wealthy countries, the second part of my proposal would help workers from the poorest countries provide financial resources to those back home through remittances.
12. D. Kapur and J. McHale, "Should a Cosmopolitan Worry about the 'Brain Drain?'" *Ethics and International Affairs* 20, no. 3 (2006): 305–20.
13. *Ibid.*

Are Our Racial Concepts Necessarily Essentialist Due to Our Cognitive Nature?

Eric Bayruns Garcia

THE GRADUATE CENTER, CUNY

The Corona neighborhood of New York City has been a predominantly Dominican neighborhood since the late 1970s and early 1980s. One day in the summer of the 2000, I was on my way into Manhattan on the elevated number 7 metro line, and I witnessed two Dominican women arguing quite loudly outside of the turnstile to enter the metro. The women's heated argument devolved from personal insults to just one thing, which they said to each other over and over. Their argument devolved into them yelling at each other that the other was blacker. Both women by US standards would be taken to be black. The argument occurred in Spanish and both likely came to the US recently, or at least as adults, as could be told by their Dominican accents specific to the cultural region of the Dominican Republic called Cibao.

The takeaway from this anecdote is that these two Dominican women hurled what they thought would be one of the most hurtful insults that they could sling at their interlocutor, namely, that their interlocutor was very black. Verbatim, one woman yelled, "¡Tu eres mas negra!" and the other woman responded with "¡No, tu eres mas negra!" This translates to the first woman saying, "You are blacker than me" and the other responding "No, you are blacker than me." By my lights, the Dominican idea of blackness contains essentialist negative content, where the content includes things such as inferior moral disposition and inferior aesthetic qualities. Here, the Dominican notion of blackness at least partly explains why the argument was so heated, because in Dominican culture attributing blackness to someone implies (1) that they are *necessarily* lacking in aesthetic value and (2) that they are *necessarily* morally deficient or inferior. And if someone implies that another person is necessarily aesthetically lacking and morally deficient, then this plausibly will cut deep as an insult. I argue that this anecdote is incompatible with recent work from Ron Mallon and Daniel Kelly on the nature and content of subjects' racial concepts and representations. Specifically, it is incompatible with Mallon and Kelly's claim regarding racial representations' content in mixed-race cultures like those from Latin America.

In the *Oxford Handbook of Philosophy of Social Science*, Ron Mallon and Daniel Kelly argue that racial representations are non-trivially fixed by "innate, domain-specific, species-typical mechanisms."¹ They claim that many social scientists and race theorists do not properly countenance these psychological mechanisms in their explanations of racial *difference* because many social scientists and race theorists have an anti-psychological and anti-individualist bias. This bias prevents race theorists from properly understanding the role that human psychology plays in creating unjust racial outcomes because these outcomes are a result of innate human dispositions to *categorize* people by race.

Mallon and Kelly distinguish between two kinds of social constructionism about race. The first kind is *parallel constructionism*. Parallel constructionism is the view that *both* racial difference *and* racial representations are the product of joint human choice and decision. The second kind is *hybrid constructionism*. Hybrid constructionism is the view that joint human decision does not completely explain racial representations' content but rather that a more complete explanation must refer to innate psychological mechanisms which constrain racial representation's content. Mallon and Kelly endorse hybrid constructionism, and they claim that most race theorists are parallel constructionists where paradigm examples of parallel constructionists are Linda Martín Alcoff,² Sally Haslanger,³ Charles Mills,⁴ and Paul C. Taylor.⁵

Mallon and Kelly claim that hybrid constructionism predicts, at least, that (1) racial representations are stable over time and (2) that racial representations should vary more in mixed-race cultures than in cultures where there is less racial mixing. I argue that hybrid constructionism's predictions do not obtain and thus hybrid constructionism requires further evidence. I argue that the historical record is inconsistent with hybrid constructionism, and I suggest that humans may not be innately disposed to categorize people by race even though we are likely disposed to categorize people into in- and out-groups. So, in this paper, I hope to show that there is an evidence set that is inconsistent with hybrid constructionism.

Good prescriptions are based on good descriptions. So good anti-racist prescriptions should be based on good, or accurate, descriptions of why people's racial representations are racist or have essentialist content. Mallon and Kelly's view of the social construction of race, hybrid constructionism, is inconsistent with much of the philosophy of race literature. If people's racial representations are constrained by innate-domain-specific mechanisms such that they have essentialist content, then anti-racist prescriptions should reflect this. If parallel constructionism is true, then it plausibly makes sense to hold people morally accountable for holding, say, that whites are superior to non-whites, because racial representations are primarily the result of human *choice* and *decision*. On the other hand, if hybrid constructionism is true, it plausibly makes less sense to hold people morally accountable for holding the belief that whites are superior to non-whites, because racial representations are primarily the result of an innate-domain-specific-species-typical psychological mechanism. Or, at the very least, if hybrid constructionism is true, then the way that we hold people accountable for holding that whites are superior to non-whites will importantly differ from the way we would hold them accountable if parallel constructionism was true.

The paper proceeds as follows. In section one I explain Mallon and Kelly's view that some of our racial representations' content is due to innate human psychological mechanisms rather than solely, or largely, human choice. In section two I explain a study which suggests that we may not innately categorize humans by race. In section three I present evidence from antiquity that is inconsistent with hybrid constructionism's prediction that racial representations are

stable over time. In section four I present evidence that is inconsistent with hybrid constructionism's other prediction that racial representations should vary more in mixed-race cultures because our innate mechanisms adapted to track "species-like" populations rather than mixed-race populations which are not "species-like."

SECTION I

Mallon and Kelly suggest that there is a race puzzle. If race is not biologically real, then it is unclear how race can serve as a prediction basis for the social sciences. The idea here is that the social sciences require something real with causal regularity on which to base their predictions. The answer given by race theorists and social scientists is that categorizing groups of people can involve causal regularity because human social practices are done in accordance with categorizations. In other words, (P1) our decisions can affect our categories, (P2) our categories can affect our social practices, and (P3) our social practices have causal properties, so (C) our race categories can enable us to make predictions about people so categorized.

If we categorize someone as a professor, then *both* the person categorized as a professor acts in certain ways because she is so categorized, *and* the rest of society acts in certain ways in relation to this person qua professor because she is so categorized. That this person acts in certain ways is what Mallon and Kelly call a social role where a social role is a set of expectations which the person who inhabits this role either consciously or unconsciously attempts to meet. Ian Hacking similarly points out that particular *kinds* of people can be constructed where kinds of people act in accordance with the idea of a particular kind because they have been categorized, either by others or by themselves, as a particular kind of person.⁶ Therefore, we can make predictions about this professor because she has been so categorized.

Mallon and Kelly call this the *standard answer* to what they call the race puzzle. This answer partly explains why, say, white American economic outcomes, on average, differ from black American economic outcomes. By Mallon and Kelly's lights, the standard answer involves that *both* racial difference in the world *and* racial representations are socially constructed. On the standard answer, enough humans choose to live with, not remedy, and perpetuate racial injustice such that racial injustice obtains. And, on the standard answer, their racial representations are also a result of human choice. Here, there is what Mallon and Kelly describe as social construction at *two parallel levels*. One level of construction obtains at the level of actual racial-outcome difference in the world. A second level of construction obtains at the level of mental representations of, say, white and black people. They call this kind of social construction, at two *simultaneous* levels, parallel constructionism.

Mallon and Kelly reject parallel constructionism because "recent work in evolutionary and cognitive psychology has suggested that important features of racial representations are explained by appeal to mental mechanisms that are species-typical, domain-specific and innate."⁷ Rather, they endorse what they call hybrid constructionism, where

hybrid constructionism is the view that these psychological mechanisms constrain our racial representations such that they have some content rather than other content. The kind of content that these mechanisms cause our racial representations to have is content of a racially essentialist kind.

Mallon and Kelly suggest that racial essentializing is a feature of our psychology, which was naturally selected or improved humans' fitness in an evolutionary sense. A racial representation is essentialist if it represents a group of people as *necessarily* having certain features, characteristics, or properties. So, on a hybrid constructionist view, our racial representations are constrained by an evolved disposition to represent other groups in essentialist ways. This view differs from parallel constructionism because by Mallon and Kelly's lights hybrid constructionism provides a *more* complete explanation of racial representations' content because it appeals to human psychology, but parallel constructionism provides a *less* complete explanation of racial representations' content because its appeal to human choice and decision cannot as completely explain our racial representation's content. Mallon and Kelly suggest that some social psychology evidence supports hybrid constructionism. They cite experiments by Susan Gelman and Henry Wellman where children attribute cow behaviors to a baby cow even though, the child participants are told, it was raised by pigs.⁸ They take this to show that humans will associate properties typically associated with a kind even when environmental and developmental situations are changed. More importantly, they take this to show that *humans* may similarly associate black typical properties with a black child even though the child was raised by, say, white parents in a white community. Or, at very least, they take this to show that even young children represent others in essentialist ways.

SECTION II

In this section I briefly explain a study which suggests that humans do *not* have "innate, domain-specific, species-typical mechanisms" such that humans naturally categorize people into groups that share some racially determined essence. I grant that we may have mechanisms that dispose us to represent people in an in-group and out-group fashion according to some shared feature of a set of people. The study does not suggest that we do not have *any* such mechanisms, but rather it suggests that we do not have a mechanism of this kind dedicated to *racial* categorization.

The study I explain below is a psychology study in which Shutts et al. test whether three- and four-year-old children use social categories like gender or race to make first-personal and third-personal inferences about others' properties and social relationships.⁹ The study's participants were white male and female children from middle-class urban or suburban centers in New England and the Midwest.

The study consisted of seven experiments. In the first experiment, *gender* rather than race was a "more potent guide to social preferences" when three-year-old children were asked to pick with whom they would rather be

friends.¹⁰ Here, seventy-four percent (74%) of children indicated that they would rather be friends with children of the same sex when presented with pictures of a male child and a female child. Fifty-four percent (54%) of children indicated that they would rather be friends with children of the same race when presented with pictures of a black child and a white child.

In the second experiment, when participants were asked to select an activity that they prefer like "having a pool party" or "having a birthday party," and then were asked to indicate whether children represented in photos of either (1) a black or a white child or (2) a male or female child also preferred the same activity, seventy percent (70%) of participants chose photos of children with the same gender, whereas, fifty-three percent (53%) of participants selected photos of children with the same race. Here, for three-year-old children, gender rather than race was the salient social category.

Neither the third nor the fourth experiment showed any statistically significant use of both gender and race when three-year-old participants made *third-personal* inferences about either social *activity* preferences or social *relationship* preferences.

Experiments five, six, and seven all tested *four-year-old* children's ability to make *third-personal* inferences about others' social activity preferences and friendship preferences. In the fifth experiment seventy percent (70%) of four-year-old participants inferred that children prefer friends of the same gender. Sixty-eight percent (68%) of children in this experiment inferred that other children prefer to be friends with their own race. Here, the older set of participants, four-year-old children, used both and gender and race to infer others' friendship preferences.

Experiments six and seven tested to see if four-year-old children used race or gender to infer that people of the same gender or race prefer the same social activities. Neither race nor gender was found as a robust inference basis.

I now consider what this study suggests for race as a category that humans innately use to categorize others. The study seems to suggest that race is *not* a category which three-year-old children use to first-personally infer whom they would like to be friends with and with whom they share activity preferences, whereas gender *does* serve as a basis upon which three-year-old children make similar first-personal inferences. Race also did not serve as an inference basis for four-year-old children when inferring about social activities.

One explanation of the data is "that children are predisposed to consider gender information when evaluating people but do not possess dedicated *mechanisms* for evaluating others in accord with their race."¹¹ That is, here, gender seems *more likely* the innate category. This seems at odds with Mallon and Kelly's hypothesis that we have innate, domain-specific, species-typical mechanisms which cause our racial representations to be essentialist in nature. Shutts et al. note that even on an evolutionary-psychology

picture gender representations seem more likely to have mechanisms devoted to it than race because gender distinctions have been available much longer than race distinctions.

Mallon and Kelly took Gelman and Wellman's 1991 study to show that children represent species in essentialist ways.¹² They also took this as evidence in favor of their hypothesis that children categorize people by race in an essentialist way. Shutts et al. note that "children's performance in Experiments 6 and 7 contrasts with previous findings that 4-year-old children use gender and race to guide category-based inferences about others' biological (Gelman et al. 1986) and psychological properties."¹³ What seems to explain this inconsistency is that Gelman used *noun labels* like "girl" and "boy" to pick out targets which participants were supposed to match. Shutts et al.'s study *did not use noun labels* but rather experimenters used pictures and pointing gestures to pick out match targets. Shutts et al. point out that Waxman¹⁴ (2010) tested whether the use of noun labels in picking out match targets would make a difference, and in Waxman's study "children did not engage in race-based property induction at all when the target was not described with a noun label."¹⁵

Waxman's study suggests that the use of noun labels is what caused children to infer that cows would still exhibit cow behavior even though they were raised by pigs. If the use of noun labels is what caused children to categorize cows in an essentialist way, then Mallon and Kelly's inferential leap to the conclusion that *we, humans*, innately infer in a similar essentialist way seems unwarranted.

SECTION III

In this section I present historical evidence that is inconsistent with Mallon and Kelly's claim that racial representations are stable over time. Mallon and Kelly claim that racial representations should, on their view, be roughly the same throughout the historical record because we as humans are naturally disposed to categorize people in essentialist ways. Put another way, this mechanism determines that our racial representations have essentialist content.

Race theorists largely agree that race as an idea or category was born sometime after Christopher Columbus reached the shores of the island of Hispaniola in 1492.¹⁶ Bartolomé de las Casas, in the fifteenth century, came up with the initial ordinal ranking of peoples because Europeans had to establish what *kind* of people, first, the Native Americans were and, then, what kind black Africans were.¹⁷ It was quickly established that Natives and blacks were human enough for salvation, but not human enough to enter reciprocal relations with Europeans. But this historical account of the idea of race's genesis is inconsistent with Mallon and Kelly's claim that racial representations should not vary. On their account, essentialist race notions must have been present long before Columbus reached the shores of the Taino (Native inhabitants) island of Quisqueya (the Native word still used today to refer to Hispaniola).

Greek and Roman antiquity seems like a good place to look for essentialist racial representations, which are consistent

with hybrid constructionism because the Greek and Roman world was well aware and very familiar with sub-Saharan and eastern African peoples who would qualify as phenotypically black today. It is well documented that there were sub-Saharan blacks living throughout the Greek and Roman world.¹⁸

Race theorists and Greek and Roman scholars seem to largely agree that the modern notion of race, where race picks out some kind of biologically determined essence, did not exist in antiquity. Frank Snowden claims that even though the ancients (1) accepted slavery, (2) exhibited inter-cultural ethnocentrism, (3) valued notions of beauty derived from their own cultures over others, and (4) distinguished people from other places as barbarians, the ancients did not have anything “comparable to the virulent color prejudice of modern times.”¹⁹ According to Snowden, most scholars who have looked at the evidence share the view that “the ancients did not fall into the error of biological racism; black skin color was not a sign of inferiority.”²⁰

Now, Snowden does note that the Greeks, Romans, and even Egyptians largely viewed their “aesthetic canons” as superior to others.²¹ But he also notes in earlier work that there are numerous examples in Greek and Roman literature of Greeks and Romans expressing admiration for Ethiopian aesthetic canons. The ancient Greeks and Romans referred to dark-skinned Africans as Ethiopians. Preferring one’s own aesthetic canon can be explained by cultural influences rather than by appeal to innate mechanisms. Moreover, “on the whole, the number of expressed preferences for blackness and whiteness in the classical literature is approximately equal.”²² Thus, this aesthetic-preference point does not seem consistent with hybrid constructionism. Or, at the very least, this aesthetic-preference point does not seem to be evidence for hybrid constructionism.

Now, Benjamin Isaac suggests that Greeks and Romans were *proto-racists*.²³ He argues that the roots of modern racism can be found in Greek and Roman antiquity. However, he admits that “Greek and Roman antiquity did not know of the sort of racism that Western civilization developed in the nineteenth and twentieth centuries, since they had no concept of biological determinism.”²⁴ On his view, Greek and Roman proto-racism was based in group difference, which was caused by different environments. That is, Greek and Romans were at worst *environmental determinists* because they attributed the differences that existed between groups not to biology or skin color, but rather to the ways that environments affect our physical and mental constitution.

Isaac likens their proto-racism to ethnic prejudice. Isaac says, “If, we read that people are stupid and courageous because they live in a cold climate, then it can be argued that this is a form of proto-racism, since there is the implicit assumption that these people are stupid through physical factors beyond their control.”²⁵ Here, even if we grant that the ancient Greeks and Romans were proto-racists, their racial representations’ content was not determined, or constrained, by an innate psychological mechanism, but rather their racial representations seemed free to

track non-biological reasons for differences between genetic populations. Thus, hybrid constructionism is inconsistent with, at least, this historical evidence of racial representations in Greek and Roman antiquity.

SECTION IV

I now present sociological evidence that is inconsistent with hybrid constructionism’s second prediction that racial representations should vary more in mixed-race cultures because our innate mechanisms adapted to track “species-like” populations rather than mixed-race populations which are not “species-like.” Mallon and Kelly take Latin America to be a case where racial representations vary because of population mixing. Latin America differs from the US because historically the US population has been and is less racially mixed. Put another way, Latin American populations are not sufficiently species-like for our innate mechanisms to track. As a result, Latin American racial representations do not have essentialist features which US racial representations have.

The first bit of evidence which is inconsistent with hybrid constructionism’s second prediction is my own experience as mixed-race person or light-skinned Dominican. I have seen and experienced firsthand that Dominican racial representations have essentialist content. For example, when I was a child, I was praised precisely because I was light-skinned, and lighter skin tone was thought to be inherently better than darker skin tone. I also witnessed inter-familial maltreatment of family members because of their skin color. The motivation behind my family members’ maltreatment and my good treatment are racial representations which contain essentialist content. The content is roughly that blackness is *inherently* bad and whiteness is *inherently* good. The inherent value comes from necessary properties people who are either black or white have. In the Dominican Republic, blackness is associated with Haitians, and Haitians are unfortunately thought to be thieves, uneducated, and generally immoral. Whiteness is associated with Spain, where Spain is associated with refinement, good pedigree, and the Church. Linda Martín Alcoff similarly points out that “in the Dominican Republic, ‘black’ is defined as Haitian, and dark-skinned Dominicans do not self-identify as black but as dark Indians or mestizos.”²⁶ This seems inconsistent with Mallon and Kelly’s claim that racial representations will differ in societies which consist in mixed-race populations. Latin populations are mixed race and the Dominican population is a paradigm example of a mixed-race population, but racial essentialism is *prominent* in the Dominican population’s racial representations.

Evidence from sociology shows that in Latin American countries like Venezuela, Brazil, and the Dominican Republic there is a correlation between one’s socioeconomic status and the likelihood that one will either identify as white or identify closer on the racial spectrum to white irrespective of one’s actual skin-tone.²⁷ Edward Telles and Tianna Paschel suggest that “high-socioeconomic status tends to be associated with whiteness, or at least with non-blackness.”²⁸ This suggests that in Latin America (1) either white identity or an identity on the color spectrum closer to white is preferred to black identity, and (2) there is some kind of

desire to be whiter in people who *recognize themselves* as insufficiently white. This is further evidence of essentialist racial representations at work in Latin America where Mallon and Kelly claim there should be none.

However, one could object that the sociological evidence does not show that Latin American racial representations contain essentialist content because Latin Americans value white identity above others not because they think white identity is inherently or intrinsically superior to black identity but rather because they think it is economically advantageous. That is, the average Latin American person values white identity not because being identified as white indicates that they have more intrinsic worth than others, but rather because being identified as white provides more economic and social opportunities.

The objection fails because even if Latin Americans are partly motivated by economic considerations to identify as white, Latin Americans are also likely motivated by non-economic considerations to identify as white. In her study on Latin American censuses from 1850 to 1950, Mara Loveman concludes that "tacit beliefs about whiteness contained in US and Latin American censuses are very much alike."²⁹ That is, Loveman concludes that racial representations in Latin America are similar to racial representations in the US. And racial representations in the US, by Mallon and Kelly's lights, contain essentialist content due to our innate disposition to categorize by race. Thus, there is a tension *between* Loveman's claim that Latin American and US beliefs about white identity are similar *and* Mallon and Kelly's claim that Latin American racial representations should differ from US racial representations in terms of their respective essentialist content.

Loveman points out that among Latin American censuses that collected race data, the white category always preceded other categories like mestizo, mulato, or indigenous. A mulato is someone of both black and white ancestry. A mestizo is someone of both indigenous and white ancestry. Loveman says, "That this presentational choice seemed 'obvious' speaks to the successful *naturalization* of the idea that the category 'white' belongs at the top of any racial hierarchy."³⁰ That is, that the people who created the censuses uniformly assumed that the white category should be listed before other identity categories even in countries with non-white majorities suggests that they took for granted that whites were the superior race. This defeats the objection that only economic considerations motivate Latin Americans to value white identity over others because census creators do not obviously have an economic motivation to list the white category before other categories on census questionnaires.

The only Latin American census from 1850 to 1950 that did not list the white category first was the 1921 census. This was the first Mexican census after the Mexican Revolution of 1910. According to Loveman, census creators listed the mestizo category first as a deliberate political act. She says, "The inversion of the 'natural' order of racial categories was clearly a political act by those who produced the official statistics."³¹ The inversion of racial categories suggests that the dominant racial category was white. Moreover, if the

white category was not imbued with negative features, then the inversion would likely not have taken place because it would not have risen to the level of something that should be done. A plausible explanation of the inversion is that whiteness was viewed by census creators as invoking things like the superiority of whiteness over mestizo identity such that they felt compelled to invert the order of the white and the mestizo categories. If this explanation is plausible, then Mallon and Kelly's prediction, that racial representations should vary in mixed-race cultures like Latin America, should seem less plausible.

CONCLUSION

I hope to have made plausible the view that two of hybrid constructionism's predictions do not obtain. The first prediction is that racial representations should be stable across history, particularly in populations with a relatively high degree of genetic reproductive closure. That is, across history, racial representations should exhibit essentialist content in societies where populations are relatively genetically homogenous because humans adapted innate psychological mechanisms which cause us to represent species-like populations as having necessary features. I have presented expert opinions that held that there is no evidence of *essentialist* racial representations in Greek and Roman antiquity. I take Greek and Roman society to consist in populations that had sufficient reproductive closure. Thus, there seems to be an inconsistency because according to hybrid constructionism, there should be essentialist racial representations in Greek and Roman antiquity, but there is no evidence of them despite evidence of a *large* array of racial representations which are *not* essentialist.

I also hope to have plausibly suggested that hybrid constructionism's second prediction does not obtain. I presented my own firsthand testimony of racial representations with essentialist content in the Dominican Republic and in a New York City Dominican enclave. I also have presented sociological findings and observations which seem inconsistent with this second prediction.

Finally, I reviewed a study by Shutts et al. which casts doubt on the psychology study which Mallon and Kelly partly based their inference that we adapted an innate mechanism which causes our racial representations to have essentialist content. Shutts et al.'s study suggests that race is not used as an inference basis for three-year-old children in both first- and third-personal inferences, and it suggests that four-year-old children do not use race to infer third parties' properties. This research seems to cast some doubt on Mallon and Kelly's hypothesis that we have innate mechanisms which cause our racial representations to have essentialist content.

ENDNOTES

1. Ron Mallon and Daniel Kelly, "Making Race Out of Nothing: Psychologically Constrained Social Roles," in *Oxford Handbook of the Philosophy of Social Science* (New York: Oxford University Press, 2012), 515.
2. Linda Martín Alcoff, *The Future of Whiteness* (Cambridge: Polity Press, 2015).
3. Sally Haslanger, "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" *Nous*, 34, no. 1 (2000): 31–55.

4. Charles Mills, "White Ignorance," in *Race and Epistemologies of Ignorance*, eds. Shannon Sullivan and Nancy Tuana (Albany: State University of New York Press, 2007).
5. Paul C. Taylor, "Appiah's Uncompleted Argument," *Social Theory and Practice* 26, no. 1 (2000): 103–28.
6. Ian Hacking, *The Social Construction of What?* (Cambridge: Harvard University Press, 1999).
7. Mallon and Kelly, "Making Race Out of Nothing," 512.
8. S. A. Gelman and H. M. Wellman, "Insides and Essences: Early Understanding of the Non-Obvious," *Cognition* 38 (1991): 213–34.
9. Shutts et al., "Children's Use of Social Categories in Thinking about People and Social Relationships," *Journal of Cognition and Development* 14, no. 1 (2013): 35–62.
10. *Ibid.*, 42.
11. *Ibid.*, 56, my emphasis.
12. Gelman and Wellman, "Insides and Essences."
13. Shutts et al., "Children's Use of Social Categories in Thinking about People and Social Relationships," 58.
14. S. Waxman, "Names Will Never Hurt Me? Naming and the Development of Racial and Gender Categories in Preschool-aged Children," *European Journal of Social Psychology* 40 (2010): 593–610.
15. Shutts et al., "Children's Use of Social Categories in Thinking about People and Social Relationships," 59.
16. See Michael Omi and Howard Winant, *Racial Formation in the United States: From the 1960s to the 1990s* (New York: Routledge, 1994); Anibal Quijano and Michael Ennis, "Coloniality of Power, Eurocentrism and Latin America," *Nepantla: Views from the South* 1, no. 3 (2000): 533–80; and Paul C. Taylor, *Race: A Philosophical Introduction* (Cambridge: Polity Press, 2013).
17. Walter Mignolo, *The Idea of Latin America* (Malden: Blackwell, 2005).
18. See Benjamin Isaac, *The Invention of Racism in Classical Antiquity* (New York: Princeton University Press, 2013); Jeremy McInerney, *A Companion to Ethnicity in the Ancient Mediterranean* (Malden: Wiley Blackwell, 2014); Frank Snowden, Jr., *Blacks in Antiquity* (Cambridge: Harvard Press, 1970); and Frank Snowden, Jr., *Before Color Prejudice: The Ancient View of Blacks* (Cambridge: Harvard Press, 1983).
19. Frank Snowden, Jr., *Before Color Prejudice*, 63.
20. *Ibid.*
21. *Ibid.*
22. Frank Snowden, Jr., *Blacks in Antiquity*, 179.
23. Benjamin Isaac, *The Invention of Racism in Classical Antiquity*.
24. *Ibid.*, 37, my emphasis.
25. *Ibid.*, 38.
26. Linda Martín Alcoff, "Mestizo Identity," in *The Idea of Race*, eds. Robert Bernasconi and Tommy L. Lott (Indianapolis/Cambridge: Hackett, 2000), 156.
27. See Edward Telles and René Flores, "Not Just Color: Whiteness, Nation, and Status in Latin America," *Hispanic American Historical Review* 93, no. 3 (2013); and Edward Telles and Tianna Paschel, "Who Is Black, White, or Mixed Race? How Skin Color, Status, and Nation Shape Racial Classification in Latin America," *American Journal of Sociology* 120, no. 3 (November 2014): 864–907.
28. Edward Telles and Tianna Paschel, "Who Is Black, White, or Mixed Race?," 871.
29. Mara Loveman, "Whiteness in Latin America: Measurement and Meaning in National Censuses (1850–1950)," *Journal de la société des américanistes* 95, no. 2 (2009): 207–34.
30. *Ibid.*, my emphasis.
31. *Ibid.*, 11.

Epistemic Humility Now!

Damían Bravo Zamora

UNAM & UNIVERSIDAD ANÁHUAC, MÉXICO

I. INTRODUCTION

"Don't be so overly dramatic about it, Chuck. What you're saying is a falsehood, and they're giving, [. . .] our Press Secretary, gave alternative facts to that."¹

Genuine listeners are likely to have a hard time trying to understand this statement. If a falsehood is contrary to the truth, if you accuse someone of saying a falsehood and you endorse a claim that's opposite to the alleged falsehood, then it would seem that what you are defending is not merely an alternative claim, let alone an "alternative fact"—whatever that means—but the truth. I won't try to investigate what might have been meant by this particular utterance at this particular moment. I am more intrigued by the relation between this relativistic statement and the proposal, by the same group of people, to "build a big, beautiful, powerful wall" that would separate a rich world-region and a poor one. And, in fact, my aim is the much more general one of investigating whether there is any relation between relativism, on the one hand, and exclusivism, on the other. I am also interested in the question whether there is a plausible alternative to relativism.

For those familiar with the history of Western philosophy, a few familiar bells are probably ringing right now. There are no facts, we've been repeatedly told by philosophers, but only interpretations. Stated absolutely ("there are no. . ."), of course, relativism is self-stultifying. Do relativists believe in what they say? Do they believe in its *truth*? If they do, we have only to point out to them that it is inconsistent to believe in the truth of what one says if what one says is that there are no truths. If they don't believe in what they say, then they are neither taking *us* nor *themselves* very seriously, and we may decide at this point either to keep on playing games with them or to do something else with our time, but it would certainly be foolish to keep on taking them seriously. We know all that since Plato. What has not been explored enough, I think, is the question whether relativism can avail itself, and wants to avail itself, of any thought or argument in order to counter those discourses that unabashedly claim the moral, intellectual, and cultural superiority of one people over another. Relativism and exclusivism have come hand in hand before. One of the stellar proponents of relativism, Friedrich Nietzsche, had no intention whatsoever to disguise his firm belief in the superiority of one part of humanity over the other part, or his desire to keep the "superior" and "strong" part of humanity separated from the "inferior" and "weak" one. Belief in the equal dignity of all human beings can only occur, according to Nietzsche, when one has adopted the viewpoint of the slaves. Thus, if one adopts the egalitarian standpoint, one has become, whether consciously or unconsciously, a victim of the slaves' rebellion in morality: one has adopted the morality of resentment that stems from the slaves' condition and their gradual revenge.

These worm-eaten physiological casualties are all men of *resentment*, a whole, vibrating realm of subterranean revenge, inexhaustible and insatiable, in its eruptions against the happy, and likewise in masquerades of revenge and pretexts for revenge: when will they actually achieve their ultimate, finest, most sublime, triumph of revenge? Doubtless if they succeed in *showing* their own misery, *on to the conscience* of the happy: so that the latter eventually start to be ashamed of their happiness and perhaps say to one another: "It's a disgrace to be happy! *There is too much misery!*". . . But there could be no greater or more disastrous misunderstanding than for the happy, the successful, those powerful in body and soul to begin to doubt their *right to happiness* in this way. Away with this "world turned upside down"! Away with this disgraceful mollicoddling of feeling! That the sick should *not* make the healthy sick—and this would be that kind of mollicoddling—ought to be the chief concern on earth:—but for that, it is essential that the healthy should remain *separated* from the sick, should even be spared the sight of the sick so that they do not confuse themselves with the sick.²

A modern-day Nietzsche would perhaps add, "And what better way is there to separate the healthy from the sick, to spare the healthy the very sight of the sick and to prevent any admixture between these two sets of human beings, than to build a big, beautiful, powerful wall between them?"

II. RELATIVISM AND EXCLUSIVISM

My current concern, however, as I said before, is not with Nietzsche's exclusivism as such, but with the relation between exclusivism and relativism. The latter can sometimes look like a reasonable dose of uncertainty that stems from the realization that we are creatures with finite intellects. However, as we will shortly see, Nietzsche seamlessly slips from this reasonable dose of skepticism to the self-stultifying claim that *there are only interpretations*. It is only the latter, nonsensical claim which, in my opinion, is inextricably linked to Nietzsche's exclusivism. For only when you think that there are *only* interpretations, while you also think that interpreting is *essentially* "forcing, adjusting, shortening, omitting, filling-out, inventing, falsifying,"³ only then, I suggest, is your view both self-stultifying and thoroughly defenseless against exclusivism. Only then is the way paved for you to feel entitled to decide, at some given point, that you've had enough of listening to the others' perspectives. Up to that given point, you might have been clever and magnanimous enough to force yourself to view the world from the others' perspectives. But, let us not deceive ourselves, you've always "known" that not all perspectives are equally valid. How could you possibly think that, on what *grounds* could you possibly argue that all perspectives are equally valid, when you believe that there is *nothing* in the world *but* interpretations, while, amongst those interpretations, you can always very clearly distinguish those that appertain to the bossy, healthy, and strong, from those that appertain to the servile, sick, and weak? Quite unsurprisingly, logical, scientific discourse, with its proclamation of epistemic democracy

and its commitment to the attainment of an ever-wider, trans-cultural intersubjectivity, is to Nietzsche's eyes an unequivocal sign of the decadence of Western civilization:

Finally, as knowers, let us not be ungrateful towards such resolute reversals of familiar perspectives and valuations with which the mind has raged against itself for far too long, apparently to wicked and useless effect: to see differently, and to *want* to see differently to that degree, is no small a discipline and preparation of the intellect for its future "objectivity"—the latter understood not as "contemplation [*Anschauung*] without interest" (which is, as such, a non-concept and an absurdity), but as *having in our power* the ability to engage and disengage our 'pros' and 'cons': we can use the *difference* in perspectives and affective interpretations for knowledge. From now on, my philosophical colleagues, let us be more wary of the dangerous old conceptual fairy-tale which has set up a "pure, will-less, painless, timeless, subject of knowledge," let us be wary of the tentacles of such contradictory concepts as "pure reason," "absolute spirituality," "knowledge as such": — here we are asked to think an eye which cannot be thought at all, an eye turned in no direction at all, an eye where the active and interpretative powers are to be suppressed, absent, but through which seeing still becomes a seeing-something, so it is an absurdity and a non-concept of eye that is demanded. There is *only* a perspectival seeing, *only* a perspectival knowing; the more affects we are able to put into words about a thing, the *more* eyes, various eyes we are able to use for the same thing, the more complete will be our "concept" of the thing, our "objectivity." But to eliminate the will completely and turn off all the emotions without exception, assuming we could: well? would that not mean to *castrate* the intellect?⁴

As we can see, there is a small step between depriving oneself of the possibility of referring to the absolute multiplicity which is the universe—and that is exactly the possibility that one deprives oneself of when one says that there is *only* a perspectival seeing—and feeling entitled to disregard alternative perspectives deemed to be "inferior" to others. Nothing prevents slipping from one position to the other, since, if one believes that there are only interpretations and no facts, and also that some interpretations are "superior" to others, then there is no reason why one should give space to the "inferior" interpretations "for far too long."

III. EPISTEMIC HUMILITY AND THE ABILITY TO REFER TO THE UNIVERSE

Here is how pushing self-reflection and criticism a little further could have taken Nietzsche towards a radically different kind of philosophy, a more inclusive one. Not that he cared about it—he couldn't care less, as we saw.⁵ But there are relativists, like Foucault, Derrida, and Rorty, who are not exclusivists, and whose philosophies constantly denounce hidden, unsuspected, contingent forms of

oppression (Foucault⁶), speak in favor of inescapably unrealizable emancipatory promises and unconditional hospitality towards the radical other (Derrida⁷), and promote transcultural solidarity (Rorty⁸). Since they leave Nietzsche's relativism untouched, their expression of concern for the oppressed amounts to little more than a statement of good wishes. And we may, indeed, at some point have nothing else to say against exclusivism than that "we simply do not behave that way," and nothing else to do except to defend ourselves against its violence. But not at *this* point. There *is* something to say about Nietzsche's relativism; we are not obliged to buy it gullibly, nor to repeat it dogmatically. We are in a position not only to denounce untethered relativism as incoherent, but also to offer an alternative. And that gives us the possibility to reject exclusivism at this theoretical level.

In order to identify something as our interpretation of something else, we need to assume a lot of things. What do we mean by *our* interpretation? Who are we in this thought? Is it us, humans? Then we are assuming the existence of a universe which we humans inhabit, which we interpret through concepts, theories, and discourses. This idea of the universe need not be one of a *reified* or *hypostatized* universe: by "universe" we may simply refer to the all-encompassing multiplicity of what there is; this does not imply that the multiplicity itself constitutes an independent individual object in its own right, apart from its members, a totality in the sense of a Cantorian set ("a many thought of as a one"). Classes, as Russell once realized, may be thought of as logical fictions.⁹ Nor does this idea imply that the objects which constitute the absolute multiplicity which we are calling "the universe" are of any particular kind. In other words, we need not yet be committed to any particular conception with regards to all the kinds of objects that populate and can populate the universe. But we assume, at least, that we humans are part of that multiplicity. And, unless we are arrogant to the point of absurdity, no self-deceiving tendency should dissuade us from the thought that we are a very small, indeed preposterously minuscule, part of the universe. The better half of Nietzsche was undoubtedly familiar with this idea. He once wrote, surely as a criticism of the Hegelian pretension to be the very incarnation of Absolute Spirit (which is *All There Is*, and which attains its ultimate purpose, absolute knowing, only in the very act of philosophizing through nothing other than the Hegelian concepts themselves), that "if we could communicate with the mosquito, then we would learn that it floats through the air with the same self-importance, feeling within itself the flying center of the world."¹⁰

The crucial point, however, is that if we both take the thought of our finitude seriously and at the same time want to avoid self-stultification, then we have to recognize that we are committed to the truth and objectivity of at least these three thoughts: 1) There is a universe; 2) there is a way the universe is; and 3) we are embedded in that universe, i.e., we are part of it—a ridiculously small part of it at that. So, *pace* Nietzsche, at least *these* thoughts cannot be discredited as mere interpretations, if interpretations necessarily involve "forcing, adjusting, shortening, omitting, filling-out, inventing, falsifying." A contemporary philosopher, Thomas Nagel, points in the right direction when he says:

There are some types of thoughts that we cannot avoid simply *having*—that it is strictly impossible to consider merely from the outside, because they enter inevitably and directly into any process of considering ourselves from the outside, allowing us to construct the conception of a world in which, as a matter of objective fact, we and our subjective impressions are contained. . . . We discover objective reason by discovering that we run up against certain limits when we inquire whether our beliefs, values, and so forth are subjective, culturally relative, or otherwise essentially perspectival. Certain forms of thought inevitably occur *straight* in the consideration of such hypotheses—revealing themselves to be objective in content.¹¹

In other words, insofar as we want to describe some parts of our thought as subjective and perspectival, we are assuming that other parts aren't perspectival in exactly the same way. These latter parts are the thoughts that we simply must *have*, and assume to be objective, in order to intelligibly state the possibility that certain parts of our thought are merely perspectival (are mere interpretations in Nietzsche's sense). To repeat, these are thoughts like: There is a universe, there is a way the universe is, we are embedded in the universe, other beings (be they stones, numbers, gods, mosquitos, or subatomic particles) are equally embedded in the universe, etc.

Nagel's critique of relativism-subjectivism, on account of its self-stultifying nature, can only be fully appreciated if complemented with Nagel's own "realist" position, defended in his book *The View from Nowhere*. Here is a passage which, I think, correctly captures Nagel's ideas:

Creatures who recognize their limited nature and their containment in the world must recognize both that reality may extend beyond our conceptual reach and that there may be concepts that we could not understand. The condition is met by a general concept of reality under which one's actual conception, as well as all possible extensions of that conception, falls as an instance. This concept seems to me adequately explained through the idea of a hierarchical set of conceptions, extending from those much more limited than one's own but contained in it to those larger than one's own but containing it—of which some are reachable by discoveries one might make but others, larger still, are not. (The hierarchy could also include parallel conceptions, not intersecting our own but joined with it only in a larger one).¹²

The universe is larger than us—far larger. There is no reason to deny that some of our concepts can refer to objects that exist independently of our cognitive capacities, and that some of our statements and theories describe parts of reality which are independent from us and our cognitive capacities. With our concepts, we can refer to black holes three million times larger than our planet, located 53 million light-years away from us; we can refer to light, therefore, which has been traveling at nearly 300,000 kilometers per

second for the past 53 million years—all this is thoroughly compatible with the fact that we are small and contingent creatures, which appeared in a small planet in a rather small galaxy only three million years ago.¹³ At the same time, nothing assures us that with our concepts we will be able to understand, describe, or even name absolutely all objects in the universe. We can certainly have a concept like *universe* or *everything*, and through it we can refer, generally, to all there is. That is to say, we “have the general concept of *everything*, which includes both the things we can name or describe and those we can’t.”¹⁴ But the only reasonable inference seems to be to the claim that the set of things we can name or describe is incredibly smaller than the set of things we can’t name or describe. Let us call this claim the *epistemic humility* thesis. “Humility,” at this stage, has nothing to do with morality or ethics. It is much more of a prudential, methodological injunction to anyone engaging in future theoretical enterprises: Don’t be an arrogant mosquito!

IV. EPISTEMIC HUMILITY AND OPENNESS TO ALTERNATIVE WORLDVIEWS

But we are not far away from the moral realm, either. I do not mean to say that morality can be rationally derived from a set of non-moral claims. What I do mean to say is that the epistemic humility thesis gives us what untethered relativism cannot afford: a *reason* to oppose exclusivism at the epistemic level. And even though this opposition is carried out on prudential grounds, it has important consequences, which *may* be seen as a precondition—a necessary but by no means sufficient condition—for an attitude of openness to alternative worldviews.

We said before that relativism is defenseless against discursive proposals of oppressive practices. Since there are no facts but only interpretations, but amongst interpretations some can be seen to belong to the powerful, while others can be seen to belong to the oppressed, then all we have is “superior” and “inferior” interpretations clashing against one another. There is no non-arbitrary point of view from which one could advocate in favor of either the viewpoint of the oppressed or that of the oppressor. More importantly, there is no non-arbitrary point of view from which one could say that *all viewpoints are equally valid*. For relativism, this latter claim can only be seen as an expression of sympathy for the oppressed and a desire for their perspective to be taken into account—but then, again, this sympathy may be completely absent, as it was in Nietzsche’s case. Epistemic humility, however, does afford a reason, hence a non-arbitrary thing to say, regarding the principle that all perspectives are equally valid. The reason is that “from the point of view of the universe,” there is no reason why any perspective should have more value than any other. All earthlings, including humans, are not unlike minuscule mosquitoes when considered in relation to the whole universe. And just as there is in principle no reason why any particular mosquito’s perspective is superior to any other particular mosquito’s perspective, there is also in principle no reason why any particular human being’s perspective should be superior to any other particular human being’s perspective. The very idea of objectivity, of an independently existing universe

common to all perspectives and the ultimate object that those perspectives attempt to describe, presupposes the equality-in-principle of at least of those beings who have the same set of cognitive capacities.

My current claim is that even though there are dangers involved in these ideas of *objectivity* and *equality-in-principle-of-all-perspectives-of-beings-with-the-same-cognitive-capacities*—as there are dangers in any general idea—they are nonetheless indispensable if epistemic humility and openness to alternative worldviews are ever to play a role in our cultural and political lives. To see why these ideas are dangerous, we need only think of the way in which science may be and has been violently imposed upon peoples and individuals with different worldviews and practices. To see why these ideas are nonetheless indispensable for a more inclusive coexistence between different worldviews, we need to take into account the following factors: 1) Violent imposition is not a necessary feature of science or of any other enterprise of reason that aims at objectivity—indeed, as many philosophers from diverse traditions seem to agree, it is contrary to the spirit of rational argumentation to even attempt to impose a belief through violent means¹⁵; 2) to deny the existence of an independent, objective universe is a self-stultifying position—and there is *nothing* more radically opposed to the possibility of reaching an understanding between alternative views than a self-stultifying position: only non-self-stultifying positions may be presented to our interlocutors with the hope of attaining a common, shared view of a common, shared world; and 3) to say that the aim of a rational enterprise is the theoretical one of articulating an objective worldview accessible to all beings with common cognitive capacities is *not* to be confused with the proposal that any belief in particular is unrevisable—it is only to say that if any revision is proposed (if a viewpoint is eventually discovered to be merely subjective after all and an alternative view is suggested), it should be considered and evaluated with the same aim of articulating a more objective viewpoint. And even though the result of the revision will be a viewpoint, there is no reason to think that such a viewpoint will be personal or subjective, if by that we mean, as Nietzsche does, falsifying. Trivially, a thought is a thought, a theory is a theory, and a worldview is a worldview. But at least as long as some of our concepts and theories can be used to refer to an independently existing reality, there is no reason to think that any proposed revision will speak of a reality that only exists because we do. To go back to Nagel’s *The Last Word*:

The *aim* of situating everything in a non-first-person framework—a conception of how things are—is one to which there is no alternative. But that does not tell us what specific types of thought belong to this finally impersonal domain. . . . The *aim* of universal validity is compatible with the willingness always to consider alternatives and counterarguments—but they must be considered as candidates for objectively valid alternatives and arguments.¹⁶

The aim of locating our views in a *universally valid* framework, the recognition that there are some extremely

thin thoughts we *can't* get outside of (like the thought that we are embedded in the universe, that there is a way the universe is, that it is larger than ourselves, etc.): these are the kinds of ideas and aims that a relativist finds pretentious and dangerous. I have conceded that they are dangerous, but I believe that the charge of pretentiousness stems from a misunderstanding. For only the aim of situating our views in a universally valid framework, and only the idea that we are embedded in a universe indescribably larger than ourselves, make it logically possible to think that some of *our own* views may be merely personal, parochial, arbitrary, contingent, perspectival, falsifying. And unless we are able to do this, we will never be able to *respect* views alternative to our own as *alternative* views of the *shared* universe: we could never be in a position to realize that we were wrong about anything. It is hard to find an epistemic attitude more absurd, and more arrogant, than this.

V. CONCLUSIONS

If what I have been arguing is correct, untethered relativism should be rejected, not only because it is self-stultifying, but because it is theoretically harmless against exclusivism. Even presupposing that a relativist wants to reject exclusivism, it has nothing to offer in this confrontation, except feelings of sympathy and solidarity for those oppressed. Those feelings are important, but it is also important to know that there is an alternative to relativism. Epistemic humility has the advantage that it is not a self-stultifying position, and also that only under its motivating ideas and aims does the claim that we may be wrong about the shared universe make any sense at all. Of course, exclusivists may not even be ready to rationally discuss their ideas, having realized that entering into a rational discussion of worldview implies recognizing that they might be wrong. If exclusivists are explicit about this refusal, then, indeed, there is nothing more to do at the rational level. Only *then* should we give up and recognize that our spade has turned.

ENDNOTES

1. Words by US Counselor to the President Kellyanne Conway to journalist Chuck Todd during an interview on January 22, 2017. Todd had questioned Conway about the press secretary's false statements about attendance numbers to US President Donald Trump's inauguration ceremony.
2. Friedrich Nietzsche, *On the Genealogy of Morality*, Third essay, Section 14, p. 91. All emphases are in the original text.
3. *Ibid.*, Third essay, Section 24, p. 112.
4. *Ibid.*, Third essay, Section 12, p. 87. Emphases in the original.
5. But in case any doubts linger on, see the following passage: "The amount of 'progress' can actually be *measured* according to how much has had to be sacrificed to it; man's sacrifice *en bloc* to the prosperity of one *stronger* species of man—that would be progress . . ." *On the Genealogy of Morality*, Second essay, Section 12, p. 52.
6. See Michel Foucault, "Nietzsche, Genealogy, History," 76–100.
7. See Jacques Derrida, *Specters of Marx. The State of the Debt, the Work of Mourning, and the New International*, 94.
8. Richard Rorty, *Contingency, Irony, and Solidarity*, 5.
9. For defenses of a non-hypostatized idea of the universe, see Peter Simons, "The Universe," 236–50, and Damián Bravo Zamora, "On the Soberer Conclusions that May Be Drawn from Kantian and Cantorian Antinomies," 221–38.

10. "On Truth and Lie in an Extra-moral Sense," *The Portable Nietzsche*, 42.
11. Thomas Nagel, *The Last Word*, 20, 23–24. As we will shortly see, this criticism of relativism should be conceptually distinguished from the claim that there are unrevisable truths. For discussion of Nagel's position, see Timothy Williamson's *The Philosophy of Philosophy*, 260, and James Levine, "Logic and Solipsism," 238.
12. Thomas Nagel, *The View from Nowhere*, 98.
13. Cf. Pallab Ghosh, "First Ever Black Hole Image Released," <https://www.bbc.com/news/science-environment-47873592>.
14. Nagel, *The View from Nowhere*, 98.
15. Friar Bartolomé de las Casas (XVI century) famously criticized the violent indoctrination of the original peoples of the Americas during Colonization, because in his opinion the best way to bring a people to the "true doctrine" (by which he meant the Christian faith) was through reason and persuasion, which is essentially opposed to violence. (*Del único modo de atraer a todos los pueblos a la verdadera religión* [México: Fondo de Cultura Económica, 1942], 303). In the words of Enrique Dussel: "[For las Casas] the *only* way to bring the members of a foreign culture into a doctrine which is unknown to them is, by making use of the art of persuasion—through "a persuasive way, by means of reasons that appeal to their understanding and which are *softly attractive* in relation to their wills"—, to count on the free will of the listener so that, through no coercion, he can rationally accept the arguments offered. Fear, punishment, and the use of war and weapons are evidently the remotest means for such a rational acceptance of reasons" (from "Meditaciones anticartesianas," 305). Thomas Nagel has a related view, insofar as he thinks that reason is subject to its own standard, hence not to any external (e.g., psychological) standard, like fear of coercion: "When we juxtapose simple logical or mathematical thoughts with any other thoughts whatever, they remain subject only to their own standards and cannot be made the object of an external, purely psychological evaluation" (*The Last Word*, 58). Perhaps this idea is better known in Western philosophy through the influence of Immanuel Kant, who in his *Groundwork of the Metaphysics of Morals* argues in the following way: "Now I assert that to every rational being having a will we must necessarily lend the idea of freedom also, under which alone he acts. For in such a being we think of a reason that is practical, that is, has causality with respect to its objects. Now, one cannot possibly think of a reason that would consciously receive direction from any other quarter with respect to its judgments, since the subject would then attribute the determination of his judgment not to his reason but to an impulse" (Ak. 4:448). Other thinkers (like Karl-Otto Apel and Jürgen Habermas) could be cited as being of a piece in this regard, and it would undoubtedly prove rewarding to inquire into the differences and similarities in their views, but this comparative exercise has to be left for a future work.
16. Nagel, *The Last Word*, 68 and 69. Emphases added.

REFERENCES

- Bravo Zamora, Damián. "On the Soberer Conclusions that May Be Drawn from Kantian and Cantorian Antinomies." In *International Philosophical Inquiry* 56, no. 2 (2016): 221–38.
- Derrida, Jacques. *Specters of Marx. The State of the Debt, the Work of Mourning, and the New International*, translated by Peggy Kamuf. New York and London: Routledge Classics, 2006. (Originally published in 1993.)
- Dussel, Enrique. "Meditaciones anticartesianas: sobre el origen del anti-discurso filosófico de la Modernidad." In B. De Sousa Santos and M. P. Meneses *Epistemologías del Sur*, 283–330. Madrid: Akal, 2014.
- Foucault, Michel. "Nietzsche, Genealogy, History," translated by Daniel F. Bouchard and Sherry Simon. In *The Foucault Reader*, edited by P. Rabinow. New York: Pantheon Books, 1984. (Originally written in 1971.)
- Ghosh, Pallab. "First Ever Black Hole Image Released." April 10, 2019. *BBC News*. <https://www.bbc.com/news/science-environment-47873592>.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*, translated by Mary J. Gregor. Cambridge: University Press, 1997. (Originally published in 1785.)
- Las Casas, Bartolomé de, (1975) *Del único modo de atraer a todos los pueblos a la verdadera religión*. México: Fondo de Cultura Económica. (Originally written in 1537).

Levine, James. "Logic and Solipsism." In *Wittgenstein's Tractatus: History & Interpretation*, edited by P. Sullivan and M. Potter, 170–238. Oxford: University Press, 2013.

Nagel, Thomas. *The View from Nowhere*. Oxford: University Press, 1986.

Nagel, Thomas. *The Last Word*. Oxford: University Press, 1997.

Nietzsche, Friedrich. "On Truth and Lie in an Extra-moral Sense." In *The Portable Nietzsche*, edited and translated by Walter Kaufmann, 42–46. USA: Penguin Books, 1976. (Originally written in 1873.)

Nietzsche, Friedrich. *On the Genealogy of Morality*. Translated by Carol Deithe. Cambridge: University Press, 2006. (Originally published in 1887.)

Rorty, Richard. *Contingency, Irony, and Solidarity*. New York: Cambridge University Press, 1989.

Simons, Peter. "The Universe." In *Ratio (new series)*, XVI, no. 3 (2003–2004): 236–50.

Williamson, Timothy. *The Philosophy of Philosophy*. Oxford: University Press, 2007.

ARTICLES

Approaching Racial Embodiment, Aesthetics, and Liberation in José Carlos Mariátegui's Seven Essays

Omar Rivera

SOUTHWESTERN UNIVERSITY

In his wide-ranging socialist writings, Mariátegui re-contextualizes political crises, texts, and works of art from and for a Latin American positionality.¹ He analyzes them through alternative and marginalized historical and geopolitical frames in order to shed light on the complexity of his revolutionary juncture. He also puts into dialogue philosophers, artists, and political leaders from different eras, locations, and intellectual lineages. This is certainly a risky interpretive practice. It emphasizes unexpected aspects of theories which do not fit within accepted understandings, and aligns thinkers that are usually taken to be at odds with one another. It seems that, having shown that Latin America does not follow the economic, intellectual, and political developmental chronology that is apparent from Europe, Mariátegui cultivates a critical perspective in which affinities and tensions between various historical, epistemic, and social positions can come to be redrawn.² This eclectic theorizing responds to the demands of the urgent, transitory, non-systematic, and dynamic character of Mariátegui's situated philosophical and political reflections. Mariátegui's originality owes much to this approach. Without it, his Indoamerican socialism would not have been articulated.³ In this paper I follow this approach as I interpret Mariátegui's own texts in relation to decolonizing critiques of the concept and experience of race.

My discussion does not cover the way in which racist assumptions determine Mariátegui's socialism, including his writings on aesthetics. This is an issue that has been convincingly investigated and argued.⁴ Instead, I draw from a notion of racial embodiment from W.E.B. DuBois, Frantz Fanon, and Linda Martín Alcoff, and bring it to bear on Mariátegui's *Seven Essays* and its Peruvian context.⁵

This allows me to revisit his conception of race in relation to indigenista literature and explore embodied registers of the possibility of liberation that may otherwise remain implicit in his texts.

In my view, racial embodiment has two axes. First, it is the experience of one's body subject to an entrapping, racist gaze that projects meanings upon it (by "body" I mean posture, physical occupation of space and time, memory, affects, habits, and pre-reflective senses of self and belonging). I am referring here to senses of being behind in time and outside of space, to feelings of invisibility and disidentification, to being overwhelmed by guilt and nostalgia, among other embodied experiences. Second, racial embodiment involves a physical resistance to racist gazes in the development of affective detachments from it that sustain the formation of alternative and affirmative senses of self. Double consciousness exemplifies this, and I emphasize its embodied and affective dimension from within an oppressed positionality. Here lies the possibility of joy and other festive emotions, self-determination, and connectedness to places and lineages through marginalized cultural artifacts and narratives, among other examples. These two axes of racial embodiment, namely, entrapment and resistance, operate simultaneously and render fragmented oppressed selves. This fragmentation, however, can be partially offset by physical processes that consolidate the release from racist gazes, and the formation of resistant selves that enable the possibility of liberatory praxis. In this sense, racial embodiment can be approached in terms of its liberatory potency. In the discussion that follows, I reveal this potency through a study of Mariátegui's reflections on aesthetics and liberation.

THE PROBLEM OF RACE IN MARIÁTEGUI'S LIBERATORY AESTHETICS

In the seminal text in Latin American philosophy, *Seven Essays for the Interpretation of Peruvian Reality*, Mariátegui reveals persisting colonial structures underpinning economic, social, and political forms in a nascent modern Peru, and identifies Andean indigenous populations as embodying critical perspectives and praxical potencies that seek to overcome these structures. In order to demonstrate the existence of such an agent that would anchor and enact a struggle for liberation, Mariátegui falls into the trap of constructing a fixed representation of a specific indigenous racialized identity.⁶ For this reason Mariátegui is drawn to "indigenismo," an aesthetic trend that strives to capture representations of the "Indian" in order to show their pivotal role in the Peruvian national and political imaginary.⁷ This mixture of racial, liberatory, and aesthetic frames points to one of the greatest weaknesses in Mariátegui's revolutionary theory. He makes essentializing, representative claims that problematically romanticize indigenous Andeans, sometimes reducing them to stereotypical figures, and denigrate other racial and ethnic groups (like Chinese immigrants, people of African descent, and mestizos). Thus, he ends up enacting an entrenched racism in order to bolster his political program.⁸ Some of these disturbing claims appear in the seventh essay, "Literature on Trial," the section on "indigenismo" in particular. In fact, the knot between Mariátegui's racist views and his investment in

an aesthetic representation of Andean colonized subjects as revolutionary agents has been seen as grounds for the dismissal of the interpretation of Peruvian literature in the *Seven Essays* as well as his indigenous socialism more generally. I intend to complicate this dismissal in order to shed light on a different articulation of the problem of race in postcolonial contexts that I find in his texts.

By focusing on racial embodiment, in the discussion that follows I mobilize some of Mariátegui's own texts against the racist, representative strain that undermines his aesthetic and liberatory theories. In "Literature on Trial," for example, one not only finds evidence of Mariátegui's racist proclivities but also an understanding of race that unsettles them. Referencing Vilfredo Pareto, Mariátegui puts forth race as a constellation of physicalities that express "inclinations, interests, aptitudes for reasoning, observation, the state of knowledges."⁹ More importantly, these kinds of physicalities (that, in my view, also include affect and memory) are subject to transformation depending on "external factors," like "the actions of one society upon another," which include colonialism and its legacies.¹⁰ In other words, Mariátegui shows a historical, experiential, and physical understanding of racial determinations and finds them to be modulated by sociological and political factors, including oppression. This does not make him immune to racist views or save him from the representative aporias of liberation. Yet it reveals that he does not always hold an essentialist understanding of race and that he suggests a dynamic and contextual approach to racial embodiment, which complicates his adherence to the kind of racist representative aesthetics of liberation I described above.¹¹ This opens the possibility of reading Mariátegui's engagement with indigenismo beyond representative commitments that correspond to a colonialist racist gaze, and of exploring whether there is an implicit aesthetic mobilization of the liberatory potency of racial embodiment in the *Seven Essays*.

AN APPROACH TO RACIAL EMBODIMENT AND AESTHETICS IN MARIÁTEGUI'S WORK

Mariátegui's liberatory philosophy, especially in the *Seven Essays*, has two aspects: a "critical" one that involves a socialist analysis adapted to Latin American postcolonial conditions and a "resistive" one that explores the possibility of liberatory praxis informed by physicalities (affects, cultural habits and memories, perceptual orders, and other embodied, pre-reflective enactments) of oppressed, racialized, colonized peoples. In my retrospective reading, a guiding intuition of the second aspect appears to be that the study of racial embodiment reveals conditions for resistance. Racial embodiment appears in this respect as physical enactments both submitted to and resisting negating constructions of colonized identities. Resistance here comes to pass with memorial sensibilities that set into play excluded cultural lineages supporting alternative imaginaries and senses of self. As I noted in the beginning, this ambiguity means that racist projections can be divested of definitive sense, and resistant configurations of racialized identities can emerge articulating unforeseen possibilities of liberation.

I find that, like in Fanon's *Black Skin, White Masks*, racial embodiment in Mariátegui's writings can be understood on the basis of temporal determinations of racialized experience and affectivity.¹² In this sense, Mariátegui can be interpreted as exploring non-linear, non-episodic temporalities in relation to resistant physicalities. Such physicalities do not undermine dominant racist projections by embodying and inhabiting a non-oppressive, alternative, defined present. Instead, they let intentions gathered through resistant pre-reflective meanings (informing embodied cultures and senses of identity of the oppressed) germinate into praxical liberatory options. These intentions are effective as transformative renditions of an heterogeneous, indeterminate present, one that cannot be captured within episodic temporal logics. Such a "present," then, cannot be delimited solely by dominant meanings that exclude the racialized and oppressed. A concomitant experiential factor in this account of resistant racial embodiment is the lived memorial awareness that the embodied hold of racial domination converges with temporalizations of the present as a closed totality of meaning. This awareness attests to physical and sensuous temporalizations that are beyond the purview of such convergences and allow for the remembrance of silenced histories as informing diverse ways of being present.¹³

A close reader of Mariátegui, Aníbal Quijano, connects a modern concept of race with progressive temporalities that deem colonized, racialized peoples to be in an irrecoverable past, negating their coeval cultural and historical efficacy.¹⁴ Bringing him, Fanon, and Mariátegui together, I note that there is an intrinsic relation between the colonial/modern concept, embodiment and experience of race, and temporal sensibilities that articulate senses of self and culture in terms of totalized presents configuring episodic logics of past/present/future. The gaze that entraps racialized bodies through projected racist meanings works in concert with these temporalizations that articulate the colonized as in the past, which explains experiences of racial embodiment modulated by pastness, such as feelings of being always behind in time, nostalgia, and guilt. At the same time, the resistive dimension of racial embodiment implies temporalizations that are expressed memorially and affectively as comportments toward the past beyond its reduction to sequential logics and to the present without investments in its totalizing and exclusionary closure. In this sense, the liberatory potency of racial embodiment is concretely manifest as a modulation of memory in which linear temporalities recede in their definitive, colonizing force. These are resistive disruptions that enable porous worlds of meaning across power and temporal differentials to mix, overlap, and diverge in physical registers that are repressed by colonialism and its racist legacies.¹⁵ In my discussion of Mariátegui's indigenismo I focus on this aesthetic form's involvement in such disruptions that makes possible a recovery of silenced cultures and their histories.

Approaching indigenismo in this way implies renouncing objectivist and individualistic dispositions toward works of art and literature. Specifically, it implies emphasizing the affinities between aesthetic experiences and participation in rituals or festivals in postcolonial contexts. According

to Alberto Flores Galindo, festivals, the procession of El Señor de los Miragros in particular, crystallized Mariátegui's conception of the possibility of the revolutionary commitment of heterogeneous collectivities, inclusive of indigenous peoples. I suggest that it also defined his understanding of liberatory art as eliciting crowded, unruly, festive intimacies that relax social and political orders and transgress dominant delimitations of embodied, pre-reflective senses of self and communal belonging. In terms of the resistive temporality of racial embodiment, the time of festivals is not sequential and linear, but it engages the present as a return to itself that reconfigures the past in the present. The present of the festival is, then, open, versatile, indeterminate. It becomes embodied via affective and memorial enactments released from dominant gazes and that engender critical perspectives. I find that the relationship between this non-linear, open present and aesthetics in Mariátegui is most apparent in his attraction to surrealism as a joint political and artistic movement.¹⁶ My project here is to show that his engagement with "indigenismo" can also be understood on the basis of temporalities that disrupt linearity, and in a way that brings to the fore the liberatory potency of racial embodiment.

RACIAL EMBODIMENT AND "PICTURESQUE" AESTHETICS

In "Memory, Moment, and Tears," Pablo Oyarzún writes: "In an immediate temporal sense, trauma is something of the past that continues to happen in the present . . . it is something that never ends to belong to the past."¹⁷ He calls this insistence of the past "traumatic memory" and finds that it "persists for us . . . Latin Americans."¹⁸ In it the past is with the present but excluded from it, haunting it. It is trapped in a gap of time that cannot be integrated within the closure of the present, a closure that secures episodic, progressive, linear time.

However, "traumatic memory" is an unstable affect that can be submitted to the enforcement of a strict differentiation between past and present that confines the colonized in a former stage of a linear sequence. The modern, racist gaze can yield such enforcement. This kind of physical, temporal oppression is apparent in Fanon's phenomenological description of a corporeality of always being "late," one that blocks black bodies from action by internalizing the oppressor's projected racist meanings. Affectively, it is manifest as nostalgia for a petrified past and guilt for not joining the sequential flow of history. This constellation of sensibilities musters a reactive embodied entrenchment in progressive temporality and closed delimitations of the present, which facilitates essentializing constructions of racial identities as "past." An effect of this is representative, static, folkloric, and abstract renditions of colonized identities that have no relevance in the present and leave it untouched. These are fixed constructions that compromise the historical efficacy of excluded cultures defined by lineages of oppression, undermining their critical positionalities and liberatory potencies.

I extend Oyarzún's point to mean that the racial embodiment of trauma, unfolding as paralysis and guilt, spreads across different social spheres as a sensibility that results from

Latin American histories of colonization and a shared sense of an irrecoverable pre-Columbian past. It can yield aesthetic forms that aspire to constitute a national imaginary on the basis of a neat, episodic, historical progression that would leave the present intact and univocal in its closure, a present in which every past is resolved and absorbed, and in which every future is contained as a sequential projection of itself. An example of this is what Mariátegui calls "picturesque" aesthetics. It renders representations of racialized peoples striving to fully capture identities in ways that fit within dominant logics of the present. This leads to fetishizations, ornamentations, and other fabrications. This aesthetic form focuses on and values cultural cohesiveness and addresses political crises via homogenizing, progressive cultural, and national representations. An example of it is the literary trend of "surviving colonialism" that tries to configure a colonial identity but can't help but produce imitations of Spanish colonial literary forms. The emphasis here is on representations that seek to force processes of self-identification on the part of creoles, mestizos, and indigenous peoples that would settle into a solid identity. In "Literature on Trial" Mariátegui is critical of subsuming literature under such goals because it supports processes of identification that enable reactionary politics. According to him, such literary forms become divorced from revolutionary junctures, and occlude the relationship between aesthetics and liberation. In particular, Mariátegui suggests that "culturalist" investments do not engage the present in its social and political contradictions, and in its potency for transformation.

Perhaps one of Mariátegui's most insightful contributions in this regard is his thesis that "mestizaje," as it was taking form in his time, has an affinity with this kind of culturalist aesthetics. José Vasconcelos's *mestizaje*, for example, is driven by progressivism, futurity, and a messianic optimism. It has an affirmation of racial mixing that acquires a redemptive tone, one that is invested in a utopia that, Mariátegui argues, "ignores the present" and is, thus, ineffective in revolutionary terms.¹⁹ Uriel García's *mestizaje* is not utopian or messianic, but remains dangerously culturalist: he sees in *mestizaje* a syncretic cultural formation that can yield a totalized national identity. This commitment to cultural identity, in Mariátegui's view, disconnects García's work from the social, political, and economic conditions that affect the lives of the "mestizos" themselves.²⁰ He writes: "mestizaje must be analyzed not as an ethnic issue, but as a sociological issue," otherwise it becomes coopted to surreptitiously maintain a political stasis.²¹ For Mariátegui to analyze *mestizaje* as a "sociological" issue means to be attentive to contradictions within the present that attest to revolutionary potencies beyond progressivism. This is why Mariátegui concludes that an "ethnic" *mestizaje* is ultimately drawn to "evolve toward the social stage, or the kind of civilization of whites."²² *Mestizaje* in this sense turns into an "ethnic problem that is . . . completely fictitious and presupposed."²³ It is important to stress that "indigenismo" can appear to fit within a "picturesque," culturalist aesthetic form as well, since it can fall into abstract representations of "Indians." In fact, Valcárcel's *Tempestad en los Andes* argues for a historically continuous and consistent indigenous identity that would neatly compose a national imaginary, compromising the cultural complexity of indigenous and

mestizo lineages in modern Peru.²⁴ As I noted earlier, Mariátegui's own engagement with indigenismo can be seen as making similar claims, which emboldens some of his representative and racist views.

RACIAL EMBODIMENT AND "INDIGENISTA" AESTHETICS

Attention to racial embodiment in relation to "traumatic memory" and "picturesque aesthetics" shows a deep connection between the affective and memorial effects of colonization and specific aesthetic forms that emerge in colonial and post-colonial contexts. So far, I have suggested that representative and essentializing determinations of social and racial identities can be understood as repressing the disruption of linear temporality entailed in "traumatic memory," and as an affective re-entrenchment of an episodic historical logic in which the colonized fit in the past, and the colonizers in the present and future. In this section I develop an alternative unfolding of "traumatic memory" that does not fall back into linear temporalities and that I find suggested in Mariátegui's analyses of revolutionary memory and "indigenismo."

The following quote from "La Heterodoxia de la Tradición" speaks to this:

The ability to think about history and the ability to make it or create it come to be identified with one another. *Perhaps the revolutionary has an image of the past that is somewhat subjective, but it is animated and alive. . . . Revolutionaries incarnate the will of society of not becoming petrified within a stage. . . .*²⁵

This statement suggests a "liberatory memory" in which the past is indeterminate: "tradition is heterogeneous and contradictory in its components."²⁶ It yields an "image of the past" that cannot be resolved and left behind by a totalized present (or "petrified stage"), but is effective and "alive" impeding its closure. In my view, this particular memorial sensibility is an unfolding of "traumatic memory" that, rather than neutralizing the past, potentializes it in its excess of the present under the purview of linear temporality. The way in which the past haunts the present in traumatic memory is here modulated assuming a liberatory potential. This affective and memorial configuration can express a will that resists being "petrified" in a determinate historical "stage" enclosed by dominant social, political, and economic forms.

This sheds light on why, in his *Defensa del Marxismo*, Mariátegui does not posit class consciousness as a sufficient condition for revolutionary action. Class consciousness understands itself from within the present as a totality, as if it were a part of its structure, overdetermined by its economic logic. It, thus, easily falls prey to reformism and to progressive historical determinisms sustained by episodic, linear time. In revolutionary junctures, class consciousness has to be exceeded, shaken off, by a revolutionary will that affirms the present in its lack of closure, in its indeterminacy, as an untimely creation. It is moved by a liberatory memory that lets the past disrupt the present so as to set into play

silenced cultures and histories as sources for the articulation of resistant agencies. Resistance is here expressed in this memorial disruption of linear time in which the colonized ceases to be confined to the past.

In this respect, for Mariátegui, revolutionary will is sustained by sensibilities of a "morality of producers" that "does not emerge from an economic interest: it is formed in the class struggle, waged with heroic spirit, with passionate will."²⁷ This critique of class consciousness as a revolutionary factor can be applied to any identity category, including a racial, specifically indigenous, identity. This is why, even though Mariátegui identifies Andean indigenous peoples as a group having a pivotal role in his socialist revolution, he often resists understanding the Peruvian revolutionary juncture as exclusively an ethnic or racial issue. In particular, sustaining these kinds of junctures does not depend on affirming a racial identity comprehended within a definite past, present, or national imaginary.

At the same time, in my view, racial embodiment and the physicalities of oppression that it involves figure in both Mariátegui's analysis of the temporality of memorial and affective conditions for liberation and his interpretation of "indigenismo." In particular, for him this literary form is about indigenous nostalgia. He follows influential indigenistas in this respect. Valcárcel, for example, focuses on an indigenous nostalgia for a pre-Columbian past that he seeks to redeem through an indigenous resurrection. Enrique López Albújar sometimes presents indigenous peoples as nostalgically attached to a past irremediably foreign to modernization, radicalizing their uprootedness and leaving it unresolved (as in "Ushanan Jampi," for example). It seems to me that Mariátegui unfolds "indigenismo" in a different direction: to explore a modulation of the nostalgia of racialized, colonized peoples (which can be seen as an affective register of "traumatic memory") that is not anchored in a petrified past but uncovers a form of "liberatory memory." As he puts it, "literary indigenismo translates a mood, a state of consciousness of the new Peru."²⁸

Mariátegui asserts that the indigenistas "collaborate, consciously or not, with a political work of vindication—rather than restoration or resurrection."²⁹ "Vindication" here carries a sense of the past that has to be understood as eluding the futural temporality implied in "restoration" and "resurrection." It does not sediment colonized identities for the sake of redemptive political projects, but returns colonized peoples to the present in its fragmentation. Aesthetically, it eludes the lure of "picturesque" representations. "Restoration" and "resurrection," instead, connote possible resolutions of the past in the logics of a definite present and projected future, which allows for neat sequential and episodic temporalities. In this sense, they are reactionary political goals.

Mariátegui points to a "vindication" with aesthetic ramifications when he states:

What gives the Indian the right to prevail in the vision of the contemporary Peruvian is, above all, the conflict and contrast between his demographic

predominance and his social and economic servitude.³⁰

The aesthetic issue here is not that a representation of the “Indian” belongs to a vision that captures a Peruvian cultural and national identity. It is, rather, that an imaging of indigenous, colonized, and racialized peoples comes to prevail over, that is, overpower, a national and cultural vision. The “Indian,” Mariátegui writes, cannot be “valued and considered . . . as a national color or aspect, allocating him in the same plane as other ethnic elements of Peru.”³¹ This is not, then, an issue of the culturalism of “picturesque” aesthetics, but of its disruption as a form of vindication.

I suggest that the “indigenista” nostalgic, racial imaging of indigenous peoples that “prevails” over definite national imaginaries can be seen as modulating (or “translating”) a sensibility (or “mood”), eliciting a physical (specifically affective and memorial) dis-identification from linear, episodic temporality and the closure of the present it implies. This would be a nostalgia informed by “liberatory memory,” one that does not “petrify” a past that is “animated and alive” and that irrupts into the present releasing it to social and political forms informed by repressed pasts with creative potencies. In this sense, “traumatic memory” gives way to “liberatory memory” and an indigenista aesthetics of “vindication,” rather than a “picturesque aesthetics.” In its temporalization, an “aesthetic of vindication” enjoins sensibilities that sense the heterogeneity of the present fractured by a multiplicity of pasts (those of excluded Andean indigenous lineages in particular) without seeking to resolve it. Specifically, nostalgic sensibilities expressed aesthetically can come to determine willful, revolutionary dispositions within the political contradiction between the indigenous “demographic predominance” and their “social and economic servitude” as a critical juncture that exceeds totalizing renditions of the present.³²

“Indigenismo,” in this sense, allows for a temporalizing, physical (specifically affective and memorial) enactment in which racist projections (with tendencies to essentialize, de-historicize, and abstract oppressed identities) are destabilized by nostalgic attachments to colonized lineages and cultures as past but also as formative of a volatile, revolutionary, and heterogeneous present. This is a present in which excluded social and political imaginaries can become definitive and resistively meaningful. “Indigenismo,” then, sets into play an aesthetics that elicits processes undergone at the level of racial embodiment and supports articulations of liberatory praxis informed by pasts released from comprehensive logics of domination. In this respect, “indigenista” representations do not conform to a cohesive, “picturesque” aesthetic program that bolsters reactionary politics, but betray cultural irreducibilities, missed encounters, tensions, and power differentials that do not settle into a historical “episode” or determine a defined colonized identity. Through this disseminative force, “indigenismo” is an aesthetic that opens festive sites for the formation of crowds, of peoples, in contagious and transformative physicalities that affirm the inconclusiveness of the present as an opening toward liberation. I suggest, then, that ultimately Mariátegui’s “indigenismo” not only has to be understood as an aesthetics that does not dissociate

itself from revolutionary politics, but also as a transformative embodied process that can only be discerned within the fervor of peoples memorially attached to possibilities sheltered in the heterogeneity of their present (such as that of participants in celebrations like the procession *El Señor de los Milagros*). It is not surprising, then, that Mariátegui, in his famous polemic with Luis Alberto Sánchez, does not characterize “indigenismo” as an articulated program, but as a dynamic movement formed by the turbulent spirit of its revolutionary times and exceeding detached, lettered, and comprehensive attempts to fixate it.

CONCLUSION

I began by noting a critique of Mariátegui’s aesthetics, namely, that its representative investments in indigeneity can essentialize Andean peoples and cultures in order to support his revolutionary program. This would be a lettered manipulation reinforcing racist frameworks upon indigenous and other colonized peoples, and facilitated by “indigenismo.” In this sense, attention to Mariátegui’s aesthetics, especially in the context of the *Seven Essays*, can yield devastating critiques of his socialism both in its analytical and liberatory aspects. My intent in this paper is not to disprove these critiques or to show that Mariátegui is not racist. Rather, I offer a counter-reading by turning to thinkers that focus on racial embodiment and approaching Mariátegui’s writings through them. This enables me to put his rendition of “indigenismo” in dialogue with his concern with affectivities and aesthetics of resistance (found on his treatments of surrealism, for example). This intersection allows for an approach to “indigenista” aesthetics as a site in which physicalities, through destabilizing temporalizations (manifest in memory and its affective registers, like nostalgia), undo the nexus between dominant racist projections and totalized, reductive renditions of the present, and allow for the gathering of resistant meanings that inform liberatory praxis from otherwise excluded cultures and histories. I also put forth a related analysis of a reactionary “picturesque” aesthetics that shows in Mariátegui’s own texts resources to launch a critique of the kind of essentializing representational aesthetics that he falls into. In some ways, then, I read Mariátegui against himself, a tortuous hermeneutic that is, after all, part of his legacy.

ACKNOWLEDGMENTS

The research for this paper was supported by the American Council of Learned Societies, through its Frederick Burkhardt Residential Fellowship program. I also want to thank Cat Kelly, my research assistant, for their edits and comments.

ENDNOTES

1. See Mariátegui’s texts: *La Escena Contemporánea*, 15a ed. (Lima: Empresa Editora Amauta, 1988), and *Temas de Nuestra América*, 10a ed. (Lima: Empresa Editora Amauta, 1988), among others.
2. This is similar to Enrique Dussel’s “transmodern” approach. See his “World System and Transmodernity” in *Nepantla: Views From the South* 52 (2002): 221–44.
3. On the eclectic character of the *Seven Essays*, see David Sobrevilla, *El marxismo de Mariátegui y su aplicación a los 7 ensayos* (Lima: Universidad de Lima/Fondo de Desarrollo Editorial, 2005).
4. See in particular Nelson Manrique, *La Piel y La Pluma: Escritos sobre literatura, etnicidad y racismo* (Lima: SUR Casa de Estudios del Socialismo, 1999); and Renzo Llorente, “The Amauta’s

- Ambivalence: Mariátegui on Race," in *Forging People*, ed. Jorge J. E. Gracia (Notre Dame: University of Notre Dame Press, 2011).
5. See W. E. B. DuBois, "The Conservation of Races" and "Of Our Spiritual Strivings," in *W. E. B. Du Bois. A Reader*, ed. David Levering Lewis (New York: Holt Paperbacks, 1995); Frantz Fanon, *Black Skin, White Masks*, trans. Richard Philcox (New York: Grove Press, 2008), chapters 4 and 5; and Linda Martín Alcoff, *Visible Identities: Race, Gender, and the Self* (New York: Oxford University Press, 2006), chapter 4.
 6. For this critique see Jorge Coronado, *Andes Imagined* (Pittsburgh: Pittsburgh University Press, 2009).
 7. The two seminal texts that guide my discussion of Mariátegui's indigenismo are Antonio Cornejo Polar, "Indigenismo and Heterogeneous Literatures: Their Double Sociocultural Statue," *The Latin American Cultural Studies Reader*, ed. Ana del Sarto, Alicia Ríos, and Abril Trigo (Durham and London: Duke University Press, 2004), and Fernanda Beigel, "Mariátegui y las Antinomias del Indigenismo," *Utopía y Praxis Latinoamericana* 6, no. 13 (2001): 36–57.
 8. See, for example, Luis Veres, "El problema de la identidad nacional en la obra de José Carlos Mariátegui," *Amnis* [En ligne], 2 | 2002, mis en ligne le 30 juin 2002, consulté le 23 mai 2019. Available at <http://journals.openedition.org/amnis/172>; DOI: 10.4000/amnis.172.
 9. José Carlos Mariátegui, *Siete Ensayos de Interpretación de la Realidad Peruana* (Lima: Orbis Ventures, 2005), 332.
 10. *Ibid.*
 11. For a discussion of representation in Mariátegui's "indigenismo" see Omar Rivera, "Hermenéutica, representatividad y espacio en filosofías de la liberación social: una perspectiva latinoamericana," in *Sujeto, decolonización, transmodernidad: Debates filosóficos latinoamericanos*, ed. Mabel Moraña (Madrid: Iberoamericana, 2018).
 12. See Fanon, *Black Skin White Masks*, chapter 4.
 13. For a discussion of embodied temporalizations in this sense, see Omar Rivera, "Reading Alejandro Vallega Toward a Decolonial Aesthetics," *Comparative and Continental Philosophy* 9, no. 2 (2017): 162–73.
 14. See Aníbal Quijano, "Coloniality of Power, Eurocentrism, and Latin America," in *Coloniality at Large: Latin America and the Postcolonial Debate*, ed. Mabel Moraña, Enrique Dussel, and Carlos A. Jauregui (Durham: Duke University Press, 2008).
 15. I am referring here to Vallega's discussion of anachronic temporalities in Alejandro A. Vallega, *Latin American Philosophy from Identity to Radical Exteriority* (Bloomington: Indiana University Press, 2014).
 16. The main text on Mariátegui's relation to surrealism is, in my view, Fernanda Beigel, *El Itinerario y la Brújula. El vanguardismo estético-político de José Carlos Mariátegui* (Buenos Aires: Editorial Biblos, 2003). See also Omar Rivera, "Mariátegui's Avant-Garde and Surrealism as Discipline," *Symposium* 18, no. 1 (2014): 102–12.
 17. Pablo Oyarzún, "Memory, Moment, and Tears: A Speculative Approach to the Problem of Latin American Singularities," in *The New Centennial Review* 7, no. 3 (2007): 17.
 18. *Ibid.*
 19. Mariátegui, *Siete Ensayos*, 330.
 20. See José Uriel García, *El nuevo indio: Prólogo de Mario Vargas Llosa* (Lima: Fondo Editorial de la Universidad Inca Garcilaso de la Vega, 2015).
 21. Mariátegui, *Siete Ensayos*, 333.
 22. *Ibid.*
 23. *Ibid.*
 24. See Luis E. Valcárcel, *Tempestad en los Andes* (Lima: Editora Universitaria Latina S.A.C., 2006).
 25. "Heterodoxia de la Tradición," in *Invitación a la Vida Heroica. José Carlos Mariátegui. Textos Esenciales*, ed. Alberto Flores Galindo and Ricardo Portocarrero Grados (Lima: Fondo Editorial del Congreso del Perú, 2005), 408; italics mine.

26. *Ibid.*, 408.
27. José Carlos Mariátegui, *Defensa del Marxismo*, 14a ed. (Lima: Empresa Editora Amauta, 1988), 59.
28. Mariátegui, *Siete Ensayos*, 320.
29. *Ibid.*, 324.
30. *Ibid.*
31. *Ibid.*
32. *Ibid.*

América Tropical, *On the Force of Latino/a/x Thought*

Alejandro Vallega
UNIVERSITY OF OREGON

"AMÉRICA TROPICAL"

In 1932 the Mexican muralist David Alfaro Siqueiros painted an 80-foot mural in Los Angeles titled "América Tropical: Oprimida y Destruída por el Imperialismo." It would be his only one in the United States. After much controversy, officials of the city of Los Angeles had it whitewashed. Only with the rise of the Chicano mural movement in the 1960s did the mural again become an issue, and in 1988 the Getty Foundation begun working on its recovery. The original colors were never recovered; there were no color images of the original. What was so incendiary about Siqueiros's image? The mural depicted an indigenous Mexican crucified on a double cross beneath the American eagle, while two men with rifles aimed at the eagle from a nearby building. In his rendition of the Americas under North American capitalism one did not find Carmen Miranda with her fruit hat or "Ricky Ricardo" (Desi Arnaz) with his "babalooooo" and calling cheerfully, "Lucy! I am home!"¹ Instead, it was the harsh reality of Latino/Latina experience in North America and the analogous exploitation of peoples of color throughout the Americas that Siqueiros depicted with a crucifixion; the crucifixion an image that echoed the words of Guaman Poma de Ayala, when in 1600, in his *Chronicle*, he reminded the king of Spain that the colonizers had made a new sacrificed Christ out of the peoples of the Americas.² If one remains with the Latino/a/x experience and in light of that history of coloniality, without denying or diminishing the violence suffered under it, one sees in the affirmative creative sense that there is a present and potential shaping force underlying its life, lineages, histories, and thought. In this sense Siqueiro's image becomes a portal, a passage towards a new understanding of Latino/a/x presence in North America and, as a result, new paths open for understanding philosophy in light of the Latino/a/x articulate ways of being. In the following pages I begin to open this space as a path towards Latino/a/x philosophies that may be engaged as an elemental force for thinking philosophically our lives today.

Given the rise of Latino/a/x populations in the United States, their vote and economic presence have begun to raise questions about how to understand them and how to make their energy, creativity, and potential work for

North America. No longer is the issue that of recognizing a minority; the aim is to take advantage of and incorporate a new social, economic, political, and cultural force now unavoidably an element of the North American future. To this point, the traditional “orientalizing” of the Latino/a/x will not do, since it is clear that stereotypes, while providing a comfortable rubric for white America for the allocation of the non-white Americans, do not express the latter’s thought, visions, and aspirations. In short, a whole world—ways of being and of giving articulation to life—underlie the Latino/a/x experience, and it is this level of existence that would have to be engaged in order to begin to engage Latino/a/x minds and ways of life into a North American social political project. The ontological and existential level of engagement with the Latino/a/x experiences requires much more than the questioning and representation of a single ethnic or racial identity (the general and homogeneous image of the Latino/a/x held in the general North American imaginary). Moreover, a serious and substantial engagement would mean a transformation in North American consciousness. To say it in another, more direct manner, the issue is not for Latino/a/x lives and thought to become another tradition among North American traditions. It is not about getting a place at the table because the force of Latino/a/x existence lies elsewhere, in other modalities of being, other ways of being in the world and making sense of existing. We bring unthought and unimagined ways that will require the reinvention of the spaces and epistemic frames traditionally used to think and engage the world. This is why Gloria Anzaldúa, María Lugones, Chela Sandoval, and Linda Martín Alcoff, each in their own register and way, expose us to possibilities often unimaginable for the previous North American academy and intellectual tradition.

In my discussion I want to emphasize as well as dismantle the traditional racial determination of Latino/a/x identities in order to engage the possibilities for philosophical thought that underlie Latino/a/x lineages, histories, and experiences. This should also be sought at the level of intersectionality (race, gender, and labor form a tripartite system of oppression and exclusion, and only in dealing with the three may something like a liberatory or decolonial turn occur). However, for reasons of space, here I will keep mainly to race. I show in the following pages that given the distinctness and diversity of lineages, histories, memorials, and affective experiences that underlie Latino/a/x consciousness, thinking with it is not merely a matter of including another group into the already operative epistemic space, values, and logics that orient and frame contemporary North American and Westernizing philosophies. Ultimately, Latino/a/x distinctness figures diversifying possibilities which bear the opening to the transformation in the epistemic delimitations that orient the traditional homogeneous understanding of the disposition and kind of knowledge today seen as philosophy. In other words, to engage Latino/a/x thought would mean to begin thinking philosophy by rethinking it out of those concrete and distinct experiences that underlie the all too general and often obscuring term “Latino/a/x,” and even “latinx.” In what follows I focus on three fundamental issues: the displacement of the traditional racial term “Latino”; the exposure of the coloniality of power and knowledge that sustains the racial framing of Latino/a/x experience and

thought; and the possibility for rethinking philosophy that opens in light of the Latino/a/x influx of living thought beyond this system of oppression, exploitation, and exclusion.

LATINO/A/X DISTINCTNESS

Following a great part of the North American imaginary, the term Latino/a/x inscribes many peoples, histories, lineages, and ways of being under a single image. Latinos/as/xs are thought to come from “south of the border,” and with this pseudo geographical allocation appears the general image of a non-white, mixed-blood, Spanish-speaking person who has immigrated to the United States from South America, Central America, or Mexico. With this categorization Latinos/as/xs join the racial binary logic W. E. B. Du Bois captures with precision and foresight when he speaks in *The Souls of Black Folk* (1903) of “the color line,” i.e., the racism that separates white from colored races³ and which organizes the question of race throughout the modern world with its various forms of colonialism. Two other issues become salient here: As Edward Said shows in *Orientalism*, the races under the white gaze are subject to allocation through the masters’ imaginary.⁴ Thus, the Latino/a/x becomes the exotic, tropical, sexualized fruit, the illiterate manual labor, the warm-hearted ignorant child like “simpatico,” or the criminal problem: three images clearly portrayed throughout the Hollywood industry, thus throughout the world. The racist binary may be further engaged when one considers Frantz Fanon’s sharp psychiatric diagnosis of modernity. As the thinker from Martinique points out in *Black Skin White Masks*, the race division exposes a shared existential neurosis of the colonizer and the colonized, which takes place both at the psychological as well as at an epidermic level.⁵ In these terms the Latino/a/x only sees their future in becoming white, Anglo Saxon-like in mind and image, and the white believes themselves to be in charge, superior to colored people, and bearing the only possible destiny of humanity on their shoulders.

These analyses clearly situate Latino/a/x existence within a racist frame that expands throughout modernity in its westernizing form.⁶ But taking the image of a single binary division for understanding Latino/a/x reality obscures its even more complicated character. This becomes evident when one considers this experience in light of its Latin American lineages. As the Mexican philosopher Leopoldo Zea points out in “Negritude and Indigenism,” following such earlier claims as those of Azara in 1781,⁷ the issue of race in Latin America concerns not a distinct color line but mestizaje, a broad range of encounters, a palimpsest of racial differentiations and configurations.⁸ This term refers to a mixing that already begins in the sixteenth century, as documented by the chronicles of the period. Along with the destruction of the indigenous cultures Spanish rule, with its intention to include rather than exterminate, resulted in the mixture behind mestizaje. From this history appeared in Latin America mestizos, mulatos, criollos, castizos, cholos, and Zambos, to recall but a few names among the many inflections of the diversified and diversifying development of the Americas’ population.⁹ This clear lack of white and indigenous purity leads Jose Martí to say that there are no races in America.¹⁰ But the diversity of mestizaje does not exclude the recognition of distinct histories and lineages.

If Du Bois's clear-cut color line seems insufficient, it is because underlying it runs a radical diversifying reality with respect to lives, traditions, and ways of being.

Underlining the many "mestizo" configurations, the distinct formations of peoples and lives in the Americas does not only make impossible the reduction to a single type and sense of existence of the peoples that today begin to appear as a force in North America under the single name "Latino/a/x." More importantly, in that diversity one finds a profoundly rich background in light of which Latino/a/x thought and consciousness may be grounded. The register is not racial but concretely ontological: the issue is that of the distinct ways of being and determinations of identities one finds behind the Latino/a/x consciousness as indicated through the distinct diversification of lineages. In order to begin to engage this distinctness one may look into the roots of the racist prevalent system of power, and at its implications for our understanding of what counts for knowledge and philosophical thought. In this way a limited horizon becomes evident with respect to the epistemic projection that delimits the very possibility of engaging Latino/a/x experiences, identities, and thought.

LATINO/A/X DISTINCTNESS UNDER THE COLONIALITY OF POWER AND KNOWLEDGE

To speak of Latino/a/x experiences in their distinct articulate configurations of senses of being means to open a way towards many springs for the development and rethinking of philosophy today. This becomes evident when one considers the place of Latino/a/x philosophy in the very history of modern Western thought. This may seem an anachronistic claim, inasmuch as Latino/a/x philosophy seems something new. However, the relevance and place of Latino/a/x experience and thought in modernity becomes evident when one considers genealogically the lineages behind it.

Peruvian sociologist Aníbal Quijano makes evident in his work that modern Western thought centered around the *ego cogito* arises as the result of the long history of colonization of the Americas and the transatlantic trade, the development of a system of power and knowledge that is perpetuated to date and underlies capitalism, imperialism, colonialism, and today's globalizing free market economy. In his essay "Coloniality of Power, Eurocentrism, and Social Classification," Quijano exposes the origins of this system of oppression, exploitation, and exclusion, calling it the *coloniality of power and knowledge*.¹¹ This system's two basic poles are the development for the first time in world history of a single means of economic exchange and the control of production for the sake of the accumulation of wealth (the development of Capitalism) extending throughout the world, and the creation of a racial difference that accompanies the first.¹² The notion of a natural race difference between white European and the colored races sustains the economic calculative project of progress. This difference situates the white Europeans as naturally above the other races, while the other races become the natural brute labor to be put to use and exploited by the white Europeans and their "educated" descendants. This racial difference repeats the inequality wrought by force (not

reason) between conquerors and conquered, but this time in terms of the relationship between the *ego cogito* or the rational subject, and its other, the barbaric, the native, the uncivilized darker races.¹³ This binary racial division spread throughout the world and created a new human division, exemplified in the case of the Americas in the separation between peoples of color, those who are descendants of Europeans, and the Europeans themselves. This hierarchy is established through years of a systematic practice in which a differentiation is made by a series of allocations: each type/group/race is physically given a separate place in the city; their difference is also established by work functions and their wage assignments (or lack thereof, i.e., servant, slave, etc.). These differentiations result in a social placement, and in turn, the social placements and types of work produce an epistemic hierarchy. Depending on the required education, fitting type of work, and social standing, a certain natural potential and level of mental development and intelligence come to be assigned. As a result, peoples of African descent and indigenous descent, and many others, become the other of reason and of the project of white and mestizo modernity in Latin America. This appropriative and destructive configuration of "the other" is crucial to the configuration of Western identity in its European and later North American forms of domination. Given the new separation, the European mind may now distinguish itself from its "other," an "other" that has never been in a dialectic relation of power with the West or westernizing thought. Having constructed the other of Western rationalism, Europe and later North America and those who identify with them may recognize themselves by contrast: they see themselves as origin and inheritors of reason, and as angels of the project of freedom, equality, and justice that accompanies their version of the enlightenment which is a matter of calculative instrumental rationalism wedded to infinite capital production.

This brief genealogy exposes the foundation of the color line, and particularly, given its origins, this narrative touches the heart of the Latin American and Latino/a/x mind in North America. It is this racialized consciousness and degradation that is shared by Latin American and Latino/a/x thinkers in their quests for liberation. In the case of the latter, Mexicans, Puerto Ricans, Central Americans, and those Spanish-speaking peoples from other parts of the Americas or of such descent in North America take the place of the lower race with respect to the white Anglo Saxon and westernized ways of interpreting the world that still constitute the social, political, economic, and cultural centers of power.

Two other basic elements of this system are central to the understanding of philosophy and to the exclusion of Latino/a/x thought from the halls of the academy. Because the white European and later North American westernizing mind is thought of as the most advanced, its present state marks the apogee of human existence. At the same time, and as a result of the same ego-centrism, the only future for humanity may be found in the further development of its power and knowledge. This sets up a specific timeline, a historical unilinear way of looking at existence, in which present Western or westernizing rationalism determines the present and plays out the future, while all other cultures and

ways of being appear as parts of a past, or as backwards ways of understanding, at best with the potential to adapt to the single westernized vision of present and future (hence, in spite of the catastrophic economic situation in Europe and North America one still thinks of westernized countries as first-world nations, while relegating all others to being “emerging nations”).¹⁴ One sees the way this timeline operates explicitly among philosophy and the social sciences when one considers the set of binary terms that expose the timeline’s epistemically exclusionary character, as they determine the way objective/scientific knowledge looks at the phenomena and potential knowledge through the difference between Eastern-Western, primitive-civilized, magic/mythic-scientific, irrational-rational, traditional-modern, and so on.¹⁵ With the development of the coloniality of power and knowledge appears a sense of temporality that creates a certain disposition and through it provides the limits and horizons for all human knowledge. The experience of existence is situated by a temporality determined by the project of calculative production and manipulation prevalent in westernizing rationalism and its version of reason and the enlightenment.

In making explicit the way the system of the coloniality of power and knowledge situates Latino/a/x identities and thought, it is clear that the interpretation of Latino/a/x thought and experience under this system may only lead to exclusion by inclusion. Either Latino/a/x thought seeks to become closest to white westernizing thinking, or it may be interpreted as a backwards way of thinking and being that has been surpassed. In the first case the histories, lineages, and ways of being and of configuring the sense of experience distinctive of Latino/a/x lives must be abandoned. In the latter case some aspects of Latino/a/x thought and experience that fit the westernizing patterns may be kept, but only as secondary elements, adornments, and primitive insights that do not affect true knowledge and the furthering of human progress. In remaining with this epistemic framework nothing substantial may come from Latino/a/x existence and knowledge. The ways of being and articulate configurations of existence of Latino/a/x lives are denied agency and the possibility of any transformative originary impact.

**THINKING WITH LATINO/A/X DISTINCTNESS:
INVISIBLE MODERNITIES**

The narrow frame within which Latino/a/x thought and experience fall under the coloniality of power and knowledge may be undone, and Latino/a/x experiences reveal a rich and profound spring of possibilities for philosophy, when one takes seriously Quijano’s argument concerning the birth of modernity. In “Modernity, Identity, and Utopia in Latin America,”¹⁶ Quijano shows that just as in the case of the creation of the modern racial identities, modernity develops not only in Europe but also inseparably in Europe and the Americas.¹⁷ Undoubtedly with the Enlightenment one finds in Europe a shift from the religious to the secular.¹⁸ But the Peruvian sociologist adds:

If one considers the characteristic traits of the Enlightenment—the interest in the scientific investigations of the universe and the resulting

discoveries: the acceptance of the often radical intellectual risks implied in this behavior; the critique of existing social realities and the complete acceptance of the idea of change; the disposition to work reforms, against social prejudices, arbitrary power, despotism, and obscurantism—if these are the initial features of the movement of modernity, they are as documentable in colonial America as in Europe during the eighteenth century. . . . The intellectual and social movement of the Enlightenment . . . was produced and practiced simultaneously in Europe and America . . .¹⁹

In the Americas one preserves a difference within modern thought that becomes obscured whence calculative rationalism becomes fully equated with the Enlightenment. Together with calculative mathematical physics one finds the ideals of freedom, equality, and liberty, in the form of a humanist project. As Quijano explains, it is this humanism that is severed from the calculative operation of production of wealth in the name of progress in Latin America. And this separation makes explicit a possibility of a humanist modernity not yet taken up by westernizing rationalism and the project of capitalism. The humanism Quijano has in mind comes alive for him in the indigenous movements that shape anew the political horizon of the Americas, in a transformative encounter between indigenous and European, westernizing traditions. Furthermore, this encounter points to a fundamental and radical difference between the Latino/a/x experience and westernized thought.

In the same essay Quijano shows that, unlike westernized temporality, in Latin America temporality is not unilinear in its operation and development but simultaneous. What in westernizing thought appears as a past leftover, in Latin America occurs as various levels of practices which overlap to constitute realities, identities, and senses of being. This pyramidal experiencing of temporalities is lived concretely. Quijano offers the classic example of the overlapping of forms of exchange that together constitute one reality (bartering, serfdom, agriculture, industry, capitalism). In short: The temporality of Latin American experience unsettles the single history of progress established by the coloniality of power and knowledge. Furthermore, with the simultaneity of temporalities the ways of being and giving articulate configurations to existence thought meaningless return to become essential to the understanding of existence. No longer must philosophy follow the path of exclusion of the westernizing epistemic binary mentioned above.

Quijano’s analysis opens a wide path, perhaps unimaginable with respect to how existence, temporality, intersubjectivity, and cosmological thinking may be engaged in light of this simultaneous temporality at play in Latin American experience. For our purposes here, two points are crucial, and dramatically lead us back to Latino/a/x experience and thought, and do so by resituating the issue of their philosophies. First of all, from inside modernity, not as the other of westernizing rationalist capitalism, appear the excluded ways of being and thinking that remain to be engaged in order to take up

in full the question of modernity. From the excluded, the silenced, the humiliated minds and lives that constitute the underbelly of westernizing modernity arises a boundless flood of experiences, histories, lineages, and ways of giving determination to identities and of articulating ways of being. These remain to be thought, and those who carry these traditions in North America are the Latino/a/x peoples with their distinctive ways of being. They are the way to a philosophy that is coming, if philosophy will turn to its ownmost activity of giving critical articulation to life from life. Secondly, if this is the case, Latino/a/x philosophy poses a radical challenge to how one thinks philosophical thought and how one engages in it: that is, since the lineages and traditions that become central to this originary and radical modernity will not be found by only reading traditional texts, but in the literature, chronicles, and oral traditions of the Latino/a/x past, as well as in the way their lives in their practices, and projective imagination begin to configure other horizons and hence other possibilities for thought.²⁰

To return, in closing, to Siqueiro's mural. In light of our discussion, the figure of the Latino/a/x in North America appears not only as the site of oppression and suffering, but as a reminder and exhortation to engage a past and present that, along with that history of violence, brings forth what has remained for so long the blind spot of westernizing rationalist modernity, that is, the thought and lives of the oppressed, the silenced, excluded, and exploited: lives that bear unimaginable and fecund paths for philosophical thought, and for learning once again, the sense and depth of humanity and its originary expressions and spacings for freedom. Understood in this way, Siqueiro's image does not call for the inclusion of Latino/a/x thought into North America, but for the transformation of the spaces of knowledge in light of Latino/a/x experiences, our histories, and ways of being, a reality from which philosophical thought may occur anew today.

ENDNOTES

1. Ironically, "babaloo" is the father of the world in the Afro-Caribbean Lukumi, or Santería, tradition.
2. Felipe Guaman Poma de Ayala, *El Primer Nueva Corónica y Buen Gobierno*, trans. Jorge L. Urioste (Mexico, D.F.: Siglo Veintiuno, 2006).
3. W. E. B. Du Bois, *Writings: The Suppression of the African Slave-Trade / The Souls of Black Folk / Dusk and Dawn / Essays and Articles* (Library of America, 1987).
4. Edward Said, *Orientalism* (New York: Vintage Press, 1979).
5. "I believe that the fact of the juxtaposition of the white and black races has created a massive psycho-existential complex. I hope by analyzing it to destroy it." Frantz Fanon, *Black Skin White Masks* (New York: Grove Press, 2008), 12.
6. By "Westernizing," I intend the calculative rationalism wedded to capitalism and its timeline and model of progress that becomes identified with modernity.
7. In *Viaje de la América Meridional*, Félix de Azara writes, "Mixture improves the races and I think that these *mestici* have more ingenuity, sagacity and culture than the children of Spanish mothers and fathers." In Miguel Rojas Mix, *Los Cien Nombres de América* (Barcelona: Lumen, 1991), 302.
8. Leopoldo Zea, "Negritud e Indigenismo," *La Filosofía como Compromiso de Liberación* (Venezuela: Biblioteca Ayacucho, 1991), 304-305.

9. On African migrations' impact in Latin America: Dina V. Picotti, *La Presencia Africana en Nuestra Identidad* (Buenos Aires: Ediciones del Sol, 1998).
10. José Martí, "Our America," *Latin American Philosophy for the 21st Century*, ed. Jorge Gracia and Elizabeth Millán-Zaibert (New York: Prometheus Books, 2004), 251.
11. Aníbal Quijano "Coloniality of Power, Eurocentrism, and Social Classification," *Coloniality at Large: Latin America and the Postcolonial Debate*, ed. Mabel Moraña, Enrique Dussel, and Carlos A. Jáuregui (London: Duke University Press, 2008), 181.
12. *Ibid.*, 182.
13. "... a supposedly different biological structure that placed some in a natural situation of inferiority to the others." *Ibid.*, 182.
14. *Ibid.*, 190.
15. *Ibid.*
16. Aníbal Quijano, "Modernity, Identity and Utopia in Latin America," *Boundary* 20, no. 3 (1993).
17. *Ibid.*, 201-16.
18. *Ibid.*, 141.
19. *Ibid.*, 142-43.
20. As José Martí states, "Our Greece must take priority over the Greece which is not ours." Martí, "Our America," 248.

The Marxism of José Revueltas: A Struggle Against Orthodoxy

Sergio Rodrigo Lomelí Gamboa

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS, MEXICO

"Bitter is the encounter with evil, with its people,
with its space.

Evidently one was born for something else,

Something out of time and out of sense.

One would have wanted to love, sob, dance,

In another time and another planet

(even if it had been the same one).

But everything is forbidden, the sky, the earth.

They do not want us to be inhabitants.

We are suspected of being intruders in this planet.

They persecute us for it;

For going, for loving, for moving without being
told to and without chains.

They want to capture our voices,

To leave nothing left of our hands,

Of our kisses, or anything that our bodies love.

It is forbidden to be seen.

They persecute all joy.

They are dead and they kill us.

We are killed by the dead.

This is why we will live."

—José Revueltas, October 4, 1968

This poem was written just two days after the student massacres of 1968, after the Mexican government ordered to suppress a peaceful demonstration in the *Plaza de las Tres Culturas* in Tlatelolco (Mexico City). Five thousand soldiers and 200 tankettes surrounded the plaza and opened fire against a gathered crowd killing hundreds of students and civilians, and incarcerating 1,000 others.

As in Paris, and Prague in former Czechoslovakia, and the United States, there was a big social movement in Mexico during 1968. If one observes only its immediate causes, one could say that it grew out of something quite unpolitical, a brawl between students from the National Polytechnic Institute and students from a private school; though it quickly became radically political. But in a larger sense, the 1968 movement was the result of the accumulation of social unrest over the past decades, having to do with the government's repressive policies against labor unions, such as the repression against the rail workers and teachers (both in 1958).

The events leading up to the massacre in the *Plaza de las Tres Culturas* were part of the student movement, in which members of the whole university community took part. José Revueltas was one of the main intellectual figures associated with this event.

LIFE AND CONTEXT

José Revueltas was born in 1914 in Durango, Mexico, which is a state in the northern part of the country. It is not a border state, but lies just beneath Sonora, Chihuahua, and Coahuila, which share the largest portion of the border with the United States. He was born in a middle-class family of prominent artists: Silvestre Revueltas, his eldest brother, was a famous classical music composer; Fermín Revueltas was a painter, part of one of Mexico's avant-garde movements, Stridentism; and Rosaura Revueltas, his sister, was an actress, who had a role in Herbert Biberman's 1954 film "The Salt of the Earth."

The Revueltas family moved to Mexico City in 1920, where José Revueltas started going to a private school, *El colegio alemán*. After his father's death, he transferred to a public school. And then by 1925 he quit school and for the next four years taught himself at the National Library.

José Revueltas, his family, and his social and political environment are constituted greatly by the post-revolutionary context in Mexico. The alleged "Revolutionary Party" was in power, and would be for over seventy years, and though it upheld the pretention of being a *revolutionary* party, it had rather nationalist tendencies. The Revueltas family was very left wing in its political views, which was

not a rare thing at the time. But José was a radical, in the best possible way. He was a radical critic, not only of capitalism but also a radical critic of dogmatism and of the bureaucratic and totalitarian deformation of socialism.

He is better known for his extended literary production. He has several novels and short stories, and even cinematic scripts. His literary work is of late appreciation, but there is a growing agreement among scholars today that he is on par with Juan Rulfo as one of the greatest Mexican novelists.

He is also known for his political activism, which lasted throughout his whole life. He was first incarcerated at the age of fifteen for attending a rally at the Zócalo. That incarceration lasted six months. A few years later, in 1932, he was incarcerated again and was sent to a maximum-security prison at the Islas Marías, where he spent three months (It was after this experience that he wrote his first published novel, *Los muros de agua* [*Water Walls*], and then again in 1934 for organizing a strike of peasant workers in Nuevo Leon, another northern state. Finally, he was incarcerated in 1968 for his participation in the student movement; this time he was sent to "El palacio de Lecumberri," a prison in downtown Mexico City.

His theoretical and political production has had less attention, although recently there has been a renewed interest in exploring this area of his thought. He was a self-taught Marxist philosopher, and an extremely original one at that. The fact that his political thought has survived throughout decades but has not been prone to study has to do with the fact that he was permanently critical not just of capitalism and state policy, but of the communist party politics as well, which made him, and his theoretical production, a bit of an outcast. He was expelled from the Communist Party in Mexico, but made the effort to be reaccepted. His expulsion was related to the publication of a novel which deviated aesthetically from the accepted "socialist realism." One of the main characteristics of this artistic movement was the idea that the purpose of art was to promote socialist ideas, and that the party, and the proletariat, had to be portrayed, always, as highly positive characters or forces in society. Nonetheless, in his 1949 novel *Los días terrenales* (*The Earthly Days*), Revueltas explores the hypocritical contradiction in the morals of some party policies and militants who were willing to instrumentalize the life of even their own family, in order to fulfill their "historical duty." As an example of this, there is a pathetic scene where a militant of a socialist party is writing a report of a meeting while his baby daughter is dying in the next room. He decides not only not to do anything about it, but to prevent his female comrade and the mother of his child to do anything about it, because, to his eyes, that life is worth nothing compared to the historical mission of the proletariat. Evidently, party leadership did not take Revueltas's novel well and expelled him from the party, arguing that he showed existentialist and anti-Marxist deviations in his aesthetics.¹

The use of literature to explore some of his existential questions is well known. In one of his later novels, *Los errores* (*The Mistakes*), he deeply criticizes Soviet purges

and persecution of heterodoxy and dissent, and comes up with a very subtle term to name a condition or an emotion only known to those who have militated in an orthodox political organization: "party anguish." This sensation has to do with the anguish generated by the fear towards one's own party and its potential repression over one's own actions. Revueltas explores how different people react towards this condition, and how it is sometimes related to massive delations, ostracism, rebellion, and resistance, and inserts one of the main themes throughout Revueltas's thought: the struggle for autonomous consciousness.

In 1960, he renounced the party altogether, due to differences with the leadership, and founded an organization called the "Liga Leninista Espartaco" (E Spartacus-Leninist League). Towards 1963, he was expelled from this organization for being *too critical*. The members of the league argued that Revueltas's public opinions about the discrepancies between Marxist-Leninist parties around the world contravened the organization's interests. They thought that a member of a political organization did not have the individual right to think publicly regarding political issues.

Reflecting over public opinion regarding his own person, Revueltas said:

They have me for a heterodox Marxist, but in reality, they do not understand what I am: a product of Mexico, a monstrous country which we could symbolically represent as a being with the simultaneous shape of a horse, an eagle, a snake. Everything is contradiction amongst us.²

The fact remains that his thought is difficult to categorize because it changes through time and never settles with accepted forms. But this is difficult to grasp.

Often, from an external point of view, that is to say a point of view which is not formed or informed on Marx's take on society, one would think that Marxism is something quite homogenous. Therefore, there would be such a thing as a single theory or a unified body of theoretical theses common to all Marxian thought. But that is simply not true. There are multiple theories, sometimes even hostile amongst each other, which strangely fit the mega-general description of "Marxist theory": such is the case of Althusser and Luckács, or Sartre and Plejanov, or Mariátegui and Che Guevara, or Mao and the Soviets, or Luxemburg and Lenin. There is no such thing as "Marxism," in singular. Instead, we could speak, as it has been done, of a thousand Marxisms. Even so, it would be not very serious not to acknowledge the fact that there was the ideological and political pretension to homogenize and hegemonize Marxism as a single, unified theory. That pretension was held by the totalitarian and repressive Soviet State and its international organization, the Comintern.

One of the tasks of all philosophers who think that Marx's thought is still alive and profoundly relevant for our present day society is to criticize the official Soviet version of Marxism (and the official social-democratic version of Marxism), to denounce them as dogmatic doctrines rather

than critical deconstructive theories, and to show the possibilities of a Marxian thought not constrained by the ideological and political straitjacket of any orthodoxy.

Thus was the work of the best critical theory; that was the way Benjamin or Sartre approached Marxism. The intention of this text is to show, if only just a glimpse of it, in which sense José Revueltas did the same for Marxian thought in Mexico.

We must insist that José Revueltas is not an anomaly in this sense. There is not so much of a tradition of critical Marxism in Mexico, but there is a constellation of relevant critical Marxism produced in Mexico, which converges around the Faculty of Philosophy and Literature at the National Autonomous University of Mexico (UNAM). The constellation is formed by several philosophers such as Adolfo Sánchez Vázquez and Bolívar Echeverría, both critics of official versions of Marxism.

There is one last thing I would like to mention regarding Revueltas's life. There were two major political processes which marked breaking points in José Revueltas's thought. I've already mentioned the second one, the 1968 student movement, in which Revueltas took an active part and from which he developed the notion of "autogestion" to which I will come back later. But there was another one before that: the struggle of the railroad workers during 1958 and 1959. This was ten years before the 1968 movement. José Revueltas reflected profoundly on its outcome, and it enabled him to break away theoretically from the party, through the publication of a political work entitled "Ensayo de un proletariado sin cabeza" ("Essay of a Headless Proletariat") in 1962.

Through a superficial reading of this text, Revueltas may come off as an über-Leninist, criticizing all communist or socialist parties in Mexico for adhering to true Leninism, which is partly right. But there are, as well, different sorts of Leninisms. This has to do with the Stalin regime. After Lenin's death in 1924, Stalin took control over the party and the Soviet State. During this period, the name of the orthodox Marxist doctrine conducting the Third International Workers Association changed to "Marxist-Leninist." So there was a time during which if someone stated that he was a Marxist-Leninist, he meant that he adhered to the Third International, and probably militated in a "communist" party, and was in favor of Stalin's national and international policies. So, in some circles, stating that someone is a Leninist might imply that someone is in fact a Stalinist—which is not Revueltas's case.

During the Stalin era, Marxism was adopted as a positive scientific doctrine, which revealed alleged eternal positive truths. During this period, there were some parts of Marx's theoretical production that were not seen with good eyes and were not published by the Stalinist International press. In Mexico, this press was called "Editorial Progreso" ("Progress Press") and it published the standard canon for Marxist-Leninist thought. A text which was not published as part of the canon in this press was the "Economic and Philosophic Manuscripts of 1844," discovered in 1932. There, Marx writes prominently about alienated work.

Through the “Essay of a Headless Proletariat,” Revueltas tries to fuse together Lenin’s party theory to Marx’s theory of alienation contained in the “Economic and Philosophic Manuscripts.” He is, to my knowledge, one of the few Marxian theorists to have done so. This opens the arena of the philosophical problem engaged by Revueltas’s take on Marxism.

MARX’S THEORY OF ALIENATION

To understand Revueltas’s contribution to a theory regarding the philosophical problem of organization, one must have a basic comprehension of Marx’s theory of alienation. As I have said before, the *Economic and Philosophic Manuscripts* deal with this issue explicitly, though some specialists would argue that all of Marx’s theory has to do with this problem. Throughout this paper, I subscribe to this position, in contradiction to positions like those sustained by Althusser, which hold that Marx’s later work got rid of his early philosophical influences. To my eyes, Marx’s greatest work, *Capital*, is a book regarding alienation. The problem regarding alienation may be synthesized this way: relations of production developed in the capitalist society alienate people. This means, in its stronger sense, that people are no longer the active producers of their social environment, but rather passive reproducers of a social form which is being subjectively designed elsewhere. In its root, the problem of alienation describes a loss of subjectivity on behalf of human beings. They are no longer the subjects of history, but rather its objects, and sociality is being modeled by a logic or rationality that is alien to all human purposes, which is the value rationality of capital. A new pseudo subjectivity is born, which is called capital, or, in other words, value which has the ability to valorize itself, to make itself bigger. In one of the most telling bits of *Capital*, Marx puts it like this:

[T]he circulation of capital, suddenly presents itself as an independent substance, endowed with a motion of its own, passing through a life-process of its own, in which money and commodities are mere forms which it assumes and casts off in turn. Nay, more: instead of simply representing the relations of commodities, it enters now, so to say, into private relations with itself.³

Capital starts behaving as the new social subject. It is in its name that decisions are being taken in all aspects of human life: from the simple things such as where do we work or what do we do for a living; what do we study (this applies in some cases, in other not some much; for instance, studying philosophy is quite an anti-capitalist gesture in itself); what commodities are produced and where; but also the complex macro-economic problems: international policies, immigration policies, economic policies, etc.

Lukács wrote a groundbreaking book in 1923 called *History and Class Consciousness*. There he explained the narrow relationship between his take of Marx’s theory of alienation and the problem of political organization. From his point of view, the problem regarding organization is not merely a technicality, but a true philosophical problem. And it is from this point of view that I want to approach Revueltas’s comment and critique towards Lenin’s work regarding organization.

Lenin’s party theory tries to respond to the problem of alienation in its most basic form. For him, alienation implies a political problem specific to the working classes. As the proletariat is subject to alienating working conditions, their consciousness is limited. On its own, the proletariat can only reach what he calls a “trade-unionist” consciousness, which means that the proletariat will not struggle on its own for the radical transformation of society, but will struggle for better conditions of exploitation, such as a better salary, a shorter workday, more civil rights, etc. This being the situation, Lenin designs a party theory which, to his eyes, shortcuts the alienation problem in its political consequences. Revolutionary consciousness will come to the proletariat from outside the working class; it will come from the most radical intellectual sectors of the bourgeoisie, which will struggle alongside the proletariat for the emancipation of society. This thesis, by the way, was shared with Karl Kautsky, leader of the German social-democrats. So the party is formed by a compact cell of intellectuals who decide the party’s actions and politics, and a broad specter of militants and sympathizers who must execute the leadership’s commands. So, in Lenin’s party theory, there is a clear monopolization of the subjective activities regarding party life and organization.⁴

Lenin’s party theory is pretty standard for his time—it had conspiratorial and revolutionary objectives—but in its general form it was not unlike Max Weber’s depiction of modern parties throughout modern democracies:

The following is common to all parties: a nucleus of people have in their hands the active leadership, that is to say the formulation of slogans and the choosing of candidates; a group of people is adhered to them which fulfill a much more passive role and finally the rest of the members of the political association just play a role of objects . . . choosing one of the candidates and programs that the party presents to them.⁵

Weber theorizes over the consequences of this kind of political organization in much the same way as Marx had done. The outcome is a reification and bureaucratization of politics, where the ends are substituted by the means. Weber calls this problem the “iron cage.”

REVUELTAS’S CRITIQUE OF PARTY POLITICS, AND THE CONSTRUCTION OF THE THEORY OF “AUTOGESTION”

[A]utogestion means that a determined *something*, is managed and directed by its own decision towards the point where it has set itself to arrive.⁶

The main thing to understand is that for José Revueltas, party organization is a philosophical problem which has to do, mainly, with a cognitive problem. For him a Marxist organization deals with the issue of organizing class consciousness. It answers the question, “How do alienated people radically change the world?” So, for Revueltas, the question regarding organization is not a mere technical question, but rather an epistemological problem. It was always so, but by 1968–1971 his version of a Marxist

organization developed into its final form, which greatly deviates from the Leninist perspective.

There are several theoretical texts written either during the 1968 movement or during his imprisonment at Palacio de Lecumberri that deal with this subject. In a text edited with the title "Notas sobre la organización" ("Notes on Organization"), Revueltas reflects on the notion that a party has to deal with two main issues: on the abstract and theoretical level, it has to deal with something he calls "cognitive democracy," and on a more practical and concrete level, the party needs to deal with the organization of consciousness.

On the concrete, practical level, a party must act as one; there must be a unity in its action. But on the abstract, theoretical level, a party must allow for the free concurrence of opinions and tendencies, whether they be different or even in conflict with one another. Revueltas imagines a party where theoretical discussions are held unrestrictedly and in absolute freedom with the participation of every member of the party, where there is no subject which is considered taboo. This way, the subjective moments of party life are not monopolized by a single instance or committee. When the theoretical issue is resolved, then praxis overcomes the abstract level in a unified political action.⁷

His conceptions regarding organization are a result of his participation in the 1968 Movement where he built the notion of "academic autogestion." During the student strike, the university was alive with political, cultural, and theoretical activities. Revueltas saw in that experience the deepest emancipatory potentialities. He thought that in that experience you could see the most radical expression of academic freedom. He saw a critical and dialogical exercise of student activism: all students and solidary teachers building knowledge subjectively and democratically.

What is academic autogestion? It is to proceed with academic courses inside and outside curricula with the help of solidary teachers and students. It is to debate, to question, to refute, on round tables, seminars, assemblies, the issues and ideas of our time and our society [. . .] We need to create the most diverse forms of democratic organization for action, for dialogue, for controversy, ample, constant, tireless: committees, councils, symposia, encounters, dialogues with writers and intellectuals of all tendencies.⁸

Revueltas understood academic autogestion as the interruption of a university practice which only ratifies the *status quo*. Academic autogestion is student action that takes on the critical practice of producing knowledge, goes beyond curricula, and makes the totality of political and social problems an object for reflection and consciousness.

What is autogestion? A University and a superior education which are free, active, open to every national or international problem and willing to take action with them and regarding them, through study and analysis that flow into militant political activity.⁹

Revueltas understood that the core problem with alienation is the split dichotomy between subject and object. When that happens, subjectivity tends to be reified. That happens both in party life where the subjective leadership is monopolized by a single group, and in processes of traditional education, where all subjectivity relies on the teacher, and passive objectivity on the student. Revueltas saw in the 1968 student movement a subversion of this relationship, a conquest of subjectivity by students, an exercise of free criticism and unrestricted plurality, and he used that experience to theorize party life and to subvert traditional party theory.

This posits a problem for traditional Marxist organizations. The perspective of self-organization, or "autogestion" has been a subject of discussion between left-wing currents for a long time. Historically, it has been associated with political anarchism. More and more recent critical political movements around the world have tested more horizontal ways of organization: this has to do with the fact that traditional Marxist-Leninist organizations have been proved dangerous, alienating, and false.

What I think is extremely worthy of Revueltas's thought is that his theoretical commitment was never towards an author or towards a doctrine, but rather towards social reality and towards social movements, and I think it is reality that should confront our ideas and conceptions, and not the other way around. The willingness to transform one's ideas and to reassess our own conceptions is the basis of critical thinking, and I think Revueltas is a good teacher for that.

ENDNOTES

1. Adolfo Sánchez Vázquez, "La estética terrenal de José Revueltas," in *IncurSIONES literarias* (México: UNAM / FFyL, 2009).
2. José Revueltas, *Cuestionamientos e intenciones, Obras completas 18* (México, Era, 1978), 26. Arturo Anguiano Orozco, *José Revueltas, un rebelde melancólico* (México, Pensamiento Crítico, 2017), 25.
3. Marx, *Capital*, Chapter V. <https://www.marxists.org/archive/marx/works/1867-c1/ch04.htm>. Consulted April 24, 2019.
4. There are two texts in which Lenin expounds his party theory: *What Is to Be Done?* and *One Step Forward, Two Steps Back*. See Lenin, *Obras escogidas en tres tomos* (Progreso, Moscú, 1955).
5. Max Weber, *Economía y sociedad* (México, FCE, 2005), 231.
6. Revueltas, "Consideraciones sobre la autogestión académica," in *México 68: Juventud y revolución* (Mexico City: Era, 2016), 110.
7. See Revueltas, "Notas sobre la organización," in *Escritos Políticos*, Tomo III (México, Era, 1984).
8. Revueltas, "Nuestra revolución de mayo en México," en *México 68*, 38-39.
9. Revueltas, "Llamamiento a la toma de la universidad," en *México 68*, 60.

Thinking about Exile: Community, Violence, and Law

Arturo Aguirre Moreno

BENEMERITA UNIVERSIDAD AUTONOMA DE PUEBLA, PUEBLA, MEXICO

THRESHOLD

This is the era of the displaced and the refugee. In this era, it is urgent to reflect philosophically about displacement and exile as an existential situation, about the social stigma and the legal fragmentation of refugees as historical figures. In this text we approach exile from the point of view of philosophy.

Initially, we approach exile from a socio-political perspective whereby a person or group are forced to leave or remain outside of their country of origin due to well-established fears of being persecuted for reasons of race, religion, nationality, or public opinion. This exposes a relationship between violence and contemporary exile. It is enough to contemplate the world's horrors first seen in the Great War of 1914,¹ an event that marks the beginning of the most violent period (hostile, criminogenic, and mortal) in the history of human kind (approximately 160 million killed in the twentieth century as a result of war conflicts).² We see here the rise of instrumental rationality, applied over individuals and collectives, through very efficacious ways in order to manage fear, destruction, and the victims themselves.³

Considering such violence, the last six decades provide a testimony of the wide deterritorialization generated by the same interstated, intrastated, and subordinated wars, taking place for the sake of dictatorships, segregation, ethnic cleansing, genocide, civil wars, international interventions, coup d'états, military coups, and a rampant capitalism,⁴ all of which helps us outline the figure of the exile show up again in every latitude, for each historic paradigm of forced deterritorialization, and it helps settle the theoretical-affective closeness with another kinds phenomena. In this way, Castles and Miller state that "There are a few people, in industrial or developed countries, that have no personal experience about migration or its effects; this [is a] universal experience."⁵

The proximity between the phenomenon of forced deterritorialization and "resettlement" (*migratio*) promotes a tight linkage among contemporary mobilization studies, even though it is not necessarily that one phenomenon is reduced to the other, so that it is enough to witness this mobilization's frictions, rejections, links, and joints on urban spatiality, urged by exiles, expatriates, refugees, asylees, stateless, and travelers.⁶ About this, Edward Saïd says that "modern culture is in great measure an exile, emigrates and refugee's work [. . .] the difference between exiles from before and the ones of our time is (we can emphasize it) the scale: our time it is the refugee era, the one of the displaced one, of the massive immigration."⁷

What I seek here is to explore some structural factors about the exile phenomenon through "philosophical archaeology,"⁸ so long as this method tries to put in perspective the operational forces that survive between the crisis of exile and its consolidation in Western history. We seek to think about contemporary exiles and about extraterritorial exclusion, and, in this way, propose a conceptualization which will help to outline pertinent aspects of exile, in concerns of other current human extraterritoriality phenomena.

ONE AGAINST THE ONE: THE EXILE

The importance of being in community, to be with others in law and to live in common, will be understood in an negative way, by the exile, whom is left, by decree, "Without family, lawless and homeless."

The exile shows, in this way, the terrible, threatening idea that was cultivated by tradition with philosophical naturality: besides this property or community dominion, besides this cosmos, the scenario is harsh. The other, the senseless space, the far beyond (acosmia, káos) becomes an outside without destiny, without realization, without consecration, without work, without a place: "And he who cannot live in community, or who has no needs for its own sufficiency, is not a member of the city, it is a beast or a god."⁹ However, the Stagirite forgets the third possibility: that there's someone who cannot live in community, and is by imposition an exiled, someone who was and no longer is a member of the city.

As it was told, in front of the possibility of the dissolution of those who love war among their own (the exiled criminal), law should resist, strain, and totally submit the same undoable forces that threaten the order. This was understood by the Hellene with a great juridical clarity. This may have made exile's specialization and historical clarification to have specific gradations (filled with possible ethical, sociocultural, religious, or political causes), evident at the interdiction made by Plato as a theoretical legislator.

In *The Laws*, Plato observes the dangers of bloodshed inside the city-community, whereby what the philosopher does, about this, the widest crossovers between voluntary, involuntary, and "intermedial" homicides (by temper, thymos) performed by doctors, friends, from children to parents, from parents to children, between siblings, from master to slave, from slave to master, between slaves, from a freeman to foreigner and vice versa, between foreigners, between the citizens and the slaves.¹⁰

From this, Plato infers that the exile will work not only as a punishment (law's punitive execution), but also as a regulator of revenge and the contagious violence released by bloodshed within the city limits, which enables the Platonic justice (diké). In this regard, the evocations to purifications, the attendance to temples and ritual processes works as a device, not only for legal administration or at the free use of law forces, but, altogether, these evocations increase the punishment actions as a mediation with sociocultural representations being jeopardized.

Under this criteria, Plato determines that the use of the force starts when the crime's author ignores the containment "preludes" generated by the philosophical reason and the creator of law (reasonings coming from moderation, temperance, and bravery) and the reasoning (logón) (coming from rites and collective opinions) that affirm the fulfillment (revenge payment) which will be paid by the alleged murderer at the Hades.¹¹ In effect, Plato exposes this experience on the geometric mode, of which the Greek was the artificer: a metric techné of the force abuse and the correspondent punishment with excessive force of the law. Violence with violence is paid, even though inside the law this will have the qualitative trait of a revenge, not only legal but legit and impersonal, this that will be called justice (diké): is the city-one (with its institutions and juridical instruments) against the criminal.

This is the subject of energy and its equilibrium, of its virtue (areté) traveling from the epic's spirit and the Nemesis of the fifth century BCE tragedy.¹² For Plato, the city building and its maintenance inside virtue will depend on the relations of forces individual-community;¹³ but when this relation overflows, the rigor of the community force, turned into power, it is emphasized at the same time.

The Platonic speech submits, precisely, the law of exile to an obligation: "what is close to the greatest evil should have the greatest punishment";¹⁴ since there are other regulatory mechanisms for action, like those quoted by Plato: the view of social institutions, the victim's avenger or the fear (fobouménō) to Hades justice. The exile's punishment (phygén epibálein), with the criteria of *The Laws*, involves the search for narrowing and confining its author as a public enemy;¹⁵ all of this evidences i) the irreducibility of an irreversible criminal violence (a cause of death), ii) the corrosion and destructiveness of social links created over homicide (the community juridical derealization), iii) the possibility of becoming contagious (revenge: "deaths that should be purged through other deaths")¹⁶ as a correspondent violence, and iv) the arbitrariness (sociopolitical disequilibrium) that could be carried by a homicidal act.

Facing this homicidal violence—executed by that one who "loves war among their own"—Plato issues laws of exile that seek to isolate, to disarticulate, to set apart, and to turn the aggressor defenseless, making him part of a non-human, non-divine, non-animal dimension, but larval: behaviors and attributes of a being that is not alike or absolutely other, instead has been devalued by its own acts to a condition under which he could be killed.

It is possible to find, in this way, a compensatory relation between the act and the punishment. The community and the punished one appear before the application of the law through violence: the violence of that who dissolves the links, promotes and instills harm, in detriment of other or others' integrity; harm deliberately chosen and imparted by the agent, and that is undesired by that or those who suffer it;¹⁷ and here the force of the community law appears, which looks to fix the broken link by not only neutralizing the violence and its contagion but also the one

who created it. In this point, it is pertinent to go to Girard when he states that:

the damages that violence can trigger are so big, and the remedies are so random, that the importance falls over prevention. And the preventive field is fundamentally the religious field. The religious prevention can have a violent side. . . . The "clever" utilization of certain violence properties, specially its aptitude to move from one object to another one.¹⁸

To kill someone is not something that inside of the legal framework could match with revenge or compensation (as it was on Ancient Greek pre-law); even though the exile is not going to work as a punishment, neither in Plato nor in the Athenian law, solely on the juridical field, instead of that it works on the juridical geometry mixture, social sanction, and ritual process. This is legible in *The Laws*.

If Plato assumes the exile's intensities, this can't be reduced into a subjective gesture of the Athenian philosopher; instead, it will be an operative concept perceptible to Plato and that is beneath the Athenian society of the sixth and fifth BCE centuries. The latent force of the law, of which one should run and be afraid of (*fobéo*) works under a parameter where the coercion is legit and the applied violence is legal: and this follows the idea of the community virtues, considering this like a well-governed and well-educated city. Gernet is clear about it:

precisely in this family relationships order, [the curse] can be directed with a view to satisfaction and sanction, to the whole group. We have then have the background of a characteristic part of city's right, which guaranteed the respect of familiar moral with particular procedures in cases where an internal discipline was not enough; it happened, in example, when the homicide revenger was not showing up and, overall, in the case of "mistreats", about a certain relatives' category. The case makes us think about one of the Rome's "royal laws", many of them, at least, and it can be say without doubts, represent an authentic reminiscence of an old habit: in virtue of a ploratio of the father or the mother "mistreated" by the son, this was entrusted, under the sacer appellation, to the "relatives' gods" [Legis Regiae, IV, 1]. Disposition that is not much less juridical, since even the social sanction that goes with it has a religious nature. But, what does ploratio really mean? . . . the "consecration" of the guilty one equals to an out-of-law declaration. In order words: the ploratio fulfillment, magical act, can have an analogous effect to the one that will be accomplish after the juridical proceedings.¹⁹

It turns out, then, that the one without family, lawless and homeless stated by Nestor, sang by the Iliad poet, has been consolidated on the cursing writing of law. The testimony brought by the Platonic writing is relevant as a structural exile factor, compiled on *The Laws*: the city, the "all against one," is not fortuitous that the law about homicide and exile are right up next²⁰ to the exhibition made by the Old

Athenian to his interlocutor, Clinias, about the laws related to those who are disposed by the gods of the city (the wicked),²¹ to the traitors,²² and to all of those who damage the laws affecting the constituted order ("unfair facts, in general").²³

LAW SAYS IT SO: "KILL HIM WITH IMPUNITY"

If the Athenian constitution was based on the city-state organic idea,²⁴ this could give us a clue about the concentration of powers (juridical, political, and civil-religious), followed by minimum basis for individual rights; so the city, as a sociopolitical community, gets consolidated with almost unlimited and direct sovereignty, with the power to punish (juridical category, unsuspected until then).²⁵

So it seems a "natural" sequence what will be disposed by the Platonic law against the criminal, who was sentenced to exile "in the name of the whole city":

if someone with premeditation and injustice kills someone of his lineage by his own hand, he should first be set apart of the common activities so as not contaminate temples, the market, the seaport or any other public gathering, the same if there is or there is not someone to prohibit it: the law prohibits it, in effect, and it will be prohibiting it in the name of the whole city. . . . The guilty one will be punished with death and will not be buried in his victim's country, not only because of his ungodliness, but also because of his immodesty. If he runs away and refuses to be judged, the banishment will be perpetual (phygón). If one of this banishers (pheygéto) puts a step on the deceased's land, the first of their close ones who finds it, and even one of the citizens, kill him with impunity, or, tie him well and deliver him to the magistrates who judged him, to be killed. . . .²⁶

As it can be seen, the deterritorialization criteria as a sign of exile as punishment in all its intensity is just the beginning of the way the event will be happening. The force of the whole community's power, the name that signs this force, the name of the city, has in its act the power to transform everything that is related to it into law: it shapes power (violence as a damage in the form of law) and the individuals as citizens, but, besides that, it has the power to deform individuality by a force applied until the limit, not only the territorial limits but also the existential limits that makes the human being humane, as Aristotle says.

The exiled criminal carries the weight of the punishment that is tied to a process in which not only he is everyone's enemy, but anyone is also his enemy under "a simple and brutal principal":²⁷ he is the common enemy, a public and notorious wanted criminal.

Certainly, the juridical factor (the regime, power, and the force of the law) totally transformed the interpersonal and tribal notion as a wound to the city's commonwealth because it took care of—on the sociopolitical space qualification, that we recognized as polis, and the acts made on this space—attracting and involving the violent

events of individuals against individuals in a more abstract relationship.

The transformation, today we understand it, is a sophisticated sign: that one that goes from individual revenge to justice. To restore the community's functionality (broken by the unfair events) Greek thought introduced the law that fulfilled the promise which will be the direct threat to the aggressor who forgot to be afraid of the onto-juridical equation everybody-one. That is why a fellow citizen's ("brother," that of the same lineage, born from and in this political cosmos: a homeland, a phratría)²⁸ murder becomes a common problem and its shared solution becomes a civic duty: the city is the one cursing, banishing, pursuing, and executing; it is what exempts all juridical responsibility to whoever kills the pursued one, i.e., the exiled one.

EXILE

[a] person obligated to abandon or to stay out of his native country because of fears about being persecuted due to race, religion, nationality or public opinion; a person who considers his exile as temporary (even if it can last a long lifetime), with the hope of coming back to his homeland, whenever the circumstances allows it to, but it seems precluded or incapable as long as the factors that put him into the exile are still present.²⁹

In contrast to these considerations, we should specify that if the exiled was, before modernity, punished—in a direct and personal way—not only by ideological motifs but by juridical reasons, this was done because he was a disturber of the peace, as it can be seen since Homer and Plato.

It will be emphasized that in this horizon of comprehension of antiquity, communities are built by social networks, managed in a direct way and by direct representation, for which proximity and recognition were made by social bonds, formal and informal deals from some individuals to others. In this direction, Elnadi and Riffat say this about the exile:

On those remote times in which the community regulated in its minimum details the behavior of each one of its members, to exclude on of them was, practically a death sentence. Not only the group protection was denied, and he was left alone in front of the unknown, but also the link with his ancestors and the possibility of . . . building a home was forbidden. He had no more psychic support points to give him confidence. Lost for the community, was also lost for himself.³⁰

This geographic and demographic framework (space legal qualification and sociopolitical of the individuals) is "the-one" that Plato considers while thinking about the city-state, even though the exile will keep on operating also under similar circumstances within the Roman pre-law, early medieval law,³¹ and the Spanish pre-law.³² The exile is conceived, time after time, under this fundamental context in which the juridical history of the communities is looking to punish, by implementing a prevention mechanism about the violence spread and contagion, order alteration

or against any other fact that could endangered the community; all of this from the discursive relations created between the common order (philosophical-ontological), public crimes (juridical), the stigmatization (moral-religious), the rights (juridical) deprivation and persecution until death (existential threat).

With these measures it was possible to set a rational procedure structured for penal repression that goes from one side to the other. From there, to comprehend the community on its foundations implies, rigidly, to mark not only the horizontal and progressive line of a positive consciousness of itself that goes from the myth to the law, passing by the literary *paideia* to its consecration in philosophy and law.³³ The political community inherited by the West and reproduced in its fundamental characteristics time and time again along the centuries is forged also by the underlying violence and the exclusion principles which are made of violence,³⁴ that which changes its forms, gradates, and intensities, but that is latent, potentially active, to act against its own (those called by the community as “ours”: *polités*, *civies*, neighbor, or citizen) in all the legitimacy and legality; since in all the variants created by the West, the community is not losing (on the contrary, it seems to intensify) its constituent potential of repression and exclusion.³⁵

Like this, beyond the known geographic relegation, the most evident factor of exile, the accusation of a crime (*publicum*) weighs in the penal infraction, which at the same time authorizes, by the civic law, so that any citizen could apply against the aggressor the persecution and impugn death: “should be subjected to process for whom will like to pursue him in the name of the deceased one.”³⁶

The fearsome figure shows the metamorphosis: anyone can be a “hunter of exiles” (*phygadotherás*) as part of the right granted by the city; since the exiled is not a fellow-citizen (*phratér*) anymore, an equal and is now within the line of being in a barely defined zone (man-animal: larva) at being deprived of rights, security, and political recognition.

The “forced estrangement” of the enemy of the people has become the penal infraction of “all against one,” that arithmetically corresponds—and not without some strangeness—to that who with his violently deliberate action has become into the “one against all.”

This shows, certainly, the most evident characteristic and the most referred to in the history of the exilic condition: banishment. Because the exile is represented as banishment, our tradition deploys the concept to individual or a collective in movement, escaping, estranging away.³⁷ This has its profound significance.

THE CONCEPTUALIZATION OF EXILE

Insofar as the punishment that holds the exile becomes clearer, it is necessary to bring attention to deterritorialization as the most evident factor in exile, but not the only one. From here we are prepared to show the exile within the current context and to coordinate it with significations through time, with the purpose of giving an account of its meanings and the way they are interlaced between diverse eras and contexts.

Based on the information analyzed, what we have pointed out as exile arises in a determined historical moment and gets consolidated, at first as a term, then as the conceptual system of the community who signs and confirms the regulation of a moral, religious, philosophic, juridical, and political network.

In this way, we can conceptualize and suggest that exile means a set of criteria that shows regulated actions beginning with the juridical consolidation of the political community. Because this is a specialized mechanism, a group of discursive and operational instruments of territorial exclusion destined to revoke the individual’s full use of his citizenship rights, belonging and recognizing the sociopolitical links of the community. The displacement activated expel the communal disturber and his violent actions, living him without protection or security, therefore, as criminal—an individual violent toward the community and violated by it—is at the will of whoever wishes to kill him without impunity.

The relationship settled by the Aristotelian definition of man as a living animal capable of a political existence and also capable of articulation places the possibility of the word itself to what is politically agreed, what is fair and what is not, what is good, and how to persevere in it.³⁸ This is confirmed by Aristotle: “this thing’s community is what constitutes the home and the city.”³⁹ The impediment to this common word places the exile—who has a voice (*phoné*) but not a word (*logos*)—in a zone, not only of political recognition, but also of existential indifference: for which the Platonic mandate of not interacting with the infected being due to the violence he triggered against the community gets stronger.

According to the above, exile as punishment is not only a forced displacement, a criminal neutralization, or community strangeness; as if this is not enough, the exile is asked to live the criminal “abandonment” “a bando”⁴⁰—abandoned, forlorn, by the humane ordering that inscribes and sorts the ways of living in a determined space.

We recall that the Greek polis not only created a space but also invented and founded the close relationship between political reality and ontological reality in order and measure. That is why the deterritorialization in which existence gets formalized (“family-less”), the elimination of rights proclaimed in a common speech (“law-less”), and the roofless state (“home-less,” without protection) in which the exiled has been left in—“an out of” from the humane concurrence—will be the clear goal of the juridical mechanism. For this, forced exclusion was not only of the geometrical space—like the community neither was the individual’s arithmetical addition; instead, exile consisted and consists in the complexness of space as a vital horizon.⁴¹ The exile depends on the juridical tensions that regulate, sanction, limit, indicate, and control the situation (the here and now) with an absolute, continuous, and supposedly universal pretension (held on the idea of juridical-political and ontological proximity equality) to talk about civilized relation to prescriptive spaces, the we and the inside, against external spaces, the outside in which the savage, barbarian, and uncontrollable ruled by legitimate space prevails.⁴²

In this way the punished one is the derealization of every right and every familiarity. Deprived of rights, from "his zone" (that is not "his" because it was neither created nor elected), the community keeps the exiled in an extraordinary tension, because for him the community has become something that threatens and pursues him, that hunts him: "whoever finds him can kill him without impunity."

The Greek comprehended that exile was not only fulfilled by geographical marginalization, but also by its very possibility: the power of the law, of what is "commanded by the community," to decree and triggered the force of the law converted into violence in the search of the disarticulation and even the unpunished elimination, given that the exiled has no identity, is no longer a citizen, and has been deprived of his appropriations (cultural, political, and religious) and his properties (economical and social).

The criminal (exiled not because of being the exiled and a paradigmatic figure stops being a criminal to the eyes of the law) is the alterity of this possible violence for which the community, the harmed sociopolitical body, has not a designated place; it means it has not and is not giving destiny, space, or time because the spaces generated by the community are within its legal frameworks, as are the city walls.

CONCLUSIONS

Exile is a punishment that transfigures life as "death-in-life." The exiled life carries the taste for violence of an entire people (he is the killer of a "brother" in the law or deaf to what the community commands). He who is exiled is intermedium and an unrelated being and is also the sociopolitical-juridical sign of a profound alteration that was able to distinguish and select the accurate factors to get a contradictory relationship: death in life, that which does not achieve the corporal punishment or the absolute and explicit suppression of the criminal existence.

With this it is possible to devise criteria that evidences destructive actions, regulated by a displacement, expulsion, and a declared unprotectedness, capable of putting the individuals that were once a community inside a grey and undetermined zone.

The severity of exile during Antiquity has at a time "positive," "beneficial," "necessary," and "insuperable" criteria for the community so conceived within the promise-threat dynamic under which the collective regulations work. But, in general, exile is still very much operative in modern nation-states, this because the definition of space, the delimitation of borderlines and legal systems, opens places and generates coexistence spaces (voluntary and forced), but, at the same time, closes these places, thereby relegating and abandoning entire sectors of the population in sociopolitical death zones. This is possible due to the consolidation of identity into "citizenship" in the process of democratization. With this consolidation it was possible to guarantee, in a positive way, the union of civic identity, rights of belonging and participation: "nationality." But, on the other hand, the nation-state promoted and kept the power to deny citizenship and challenge rights and entry to some individuals to the juridical-political space.⁴³

Nowadays, and beyond the geopolitical dimension, there is the city's topopolitical power, which concentrates itself, wealth, and other dynamics that flow through an intermittency of rights, obligations, punishments, juridical-political process, everyday reconfiguration of sociocultural alterations, control, security, violence to privacy and intimacy, defenselessness, etc.⁴⁴ And for this very reason it has been necessary to create new structures of power that restrain space, diversity, encounter, and conflict.⁴⁵

In this current panorama we talk about the reach and meaning of exile. Perhaps, if it has been possible to explain with any clarity the notion of territorial displacement as an alteration factor for individual dispositions, the analysis should yield figures who had committed no "crime" at all, but are submitted to punishment similar to exile.

Exile, not only as deterritorialization, but as a disqualification of life, can contribute analytic criteria to those individualities and collectives that exist in states of warfare, familyless, lawlessness, homelessness, as they are displaced migrants, refugees, indigents, vagabonds, street children, unemployed (job-less), all of whom are abandoned to a social, economic, juridical, and political homeless situation.

ENDNOTES

1. See Javier Rodrigo, *Una historia de la violencia. Historiografías del terror en la Europa del siglo XX* (México: Anthropos-UAM, 2017).
2. See Christian Gerlach, *Sociedades extremadamente violentas. La violencia en masa en el siglo XX*, vol. 1 (México: Fondo de Cultura Económica, 2015); Eric Hobsbawm, *Guerra y paz en el siglo XXI* (Barcelona: Crítica, 2006).
3. Margaret Urban Walker, *Moral Repair Reconstructing Moral Relations after Wrongdoing* (New York: Cambridge University Press, 2006), 89–98.
4. See Alto Comisionado de las Naciones Unidas para los Refugiados, "La situación de los refugiados en el mundo: en busca de la solidaridad. Resumen," <http://www.acnur.org/fileadmin/scripts/doc.php?file=fileadmin/Documentos/Publicaciones/2012/8989> (consultada el 4 de enero de 2012).
5. Stephen Castles and Mark J. Miller, *La era de la migración. Movimientos internacionales de población en el mundo moderno* (México: Porrúa, 2004), 16.
6. Aribert Reimann, "Espacios del Exilio. La Experiencia transnacional en la Ciudad de México 1934–1960," *Actas del Congreso Internacional de Ahíla XVII* (Berlín: Freie Universität Berlín, 2016), 2950–70.
7. Eward Säid, *Reflexiones sobre el exilio* (Barcelona: Debate, 2005), 18.
8. Giorgio Agamben, *Signatura rerum. Sobre el método* (Barcelona: Anagrama, 2010), 148.
9. Aristotle, *Política* (México: UNAM, 1963), I, 1253a.
10. Plato, *Las leyes* (Madrid: CEPC, 1999), 864b-865a.
11. *Ibid.*, 869e.
12. Jacqueline de Romilly, *La Grecia antigua contra la violencia* (Madrid: Gredos, 2010), 29-60.
13. Nicole Loraux, *La ciudad dividida. El olvido en la memoria de Atenas* (Buenos Aires: Katz, 2008), 79–81.
14. Plato, *Las leyes* (Madrid: CEPC, 1999), 867b.
15. Philippe Gauthier, *Symbola. Les étranger et la justice dans les cités grecques* (Nancy: Universidad de Nancy II, 1972), 144-45.
16. Plato, *Las leyes* (Madrid: CEPC, 1999), 870c.

17. Vittorio Bufacchi, "Two Concepts of Violence," *Political Studies Review* 3 (2005): 199.
18. René Girard, *La violencia y lo sagrado* (Barcelona: Anagrama, 1983), 26-27.
19. Louis Gernet, *Antropología de la Grecia Antigua* (Madrid: Taurus, 1980), 206-207.
20. Plato, *Las leyes* (Madrid: CEPC, 1999), 867a.
21. *Ibid.*, 852a.
22. *Ibid.*, 857a.
23. *Ibid.*, 859b.
24. Eduardo Nicol, *La idea del hombre* (México: Herder, 2004), 111.
25. Louis Gernet, *Recherches sur le développement de la pensée juridique et morale en Grèce. Étude sémantique* (Paris: Abin Michel Ed., 2001), 433.
26. Plato, *Las leyes* (Madrid: CEPC, 1999), 871a-e.
27. Philippe Gauthier, *Symbola. Les étrangers et la justice dans les cités grecques* (Nancy: Universidad de Nancy II, 1972), 144.
28. The expression *áphrator* (that we translate as "familyless") in the Homeric formulation enunciated by Nestor has a resonance in here: the *phratría* was during the Homeric period a brotherhood or community that grouped a series of related families which were part of a tribe subdivision; in the Athenian *polis* each *phratría* refers to a political-religious community constituted by thirty families from the same tribe, see Émile Benveniste, *Vocabulario de las instituciones indoeuropeas* (Madrid: Taurus, 1983), 168-69.
29. Paul Tabori, *Anatomy of Exile. A Semantic and Historical Study* (Michigan: Harrap, 1972), 27.
30. Bahgat Elnadi, y Adel Riffat, "Luces y sombras del exilio," *Correo de la UNESCO* (París: UNESCO, 1998), 8-9.
31. Aldo Schiavone, *Ius. La invención del derecho en Occidente* (Buenos Aires: AH, 2009); Desiderio Cavalca, *Il bando nella prassi e nella dottrina giuridica medievale* (Milán: Giufre, 1978), 65-77.
32. Eduardo de Hinojosa, *El derecho germánico en el derecho español* (Madrid: CEH, 1915); David Nirenberg, *Comunidades de violencia. La persecución de las minorías en la Edad Media* (Barcelona: Península, 2001), 33-55.
33. Werner W. Jaeger, *Paideia: los ideales de la cultura griega* (México: Fondo de Cultura Económica, 1957), 19-47.
34. Michel Maffesoli, *La violencia totalitaria. Ensayo de antropología política* (Barcelona: Herder, 1983), 29.
35. Hannah Arendt, *Los orígenes del totalitarismo* (Madrid: Taurus, 1974), 559.
36. Plato, *Las leyes* (Madrid: CEPC, 1999), 871e.
37. Claudio Guillén, *El sol de los desterrados: literatura y exilio* (Barcelona: Quaderns Crema, 1995), 24.
38. Jacques Rancière, *El desacuerdo. Política y filosofía* (Buenos Aires: Nueva visión, 1996), 13-26.
39. Aristotle, *Política* (México: UNAM, 1963), 1252b.
40. Giorgio Agamben, *Homo Sacer I. El poder soberano y la nuda vida*, (Valencia, Pretextos, 2010), 135-42.
41. Doreen Massey, *For Space* (London: Sage, 2005), 55-59.
42. *Ibid.*, pp. 149-50.
43. Giorgio Agamben, *Homo Sacer I. El poder soberano y la nuda vida*, (Valencia, Pretextos, 2010), 203-18.
44. Judith Butler, *Marcos de guerra. Las vidas lloradas* (México: Paidós, 2010), 60-61.
45. Paul Virilio, *Ciudad pánico. El afuera comienza aquí* (Buenos Aires: El zorzal, 2006), 94-97.

Book Excerpt: Marx on Bolívar

Susana Nuccetelli

ST. CLOUD STATE UNIVERSITY

Some reactions to Bolívar's social and political philosophy have focused on aspects of his moral character. A prominent critic of this sort is Karl Marx, who in 1858 published a biographical note about El Libertador in *The New American Cyclopaedia*. After declaring that Bolívar had some traits of moral character unworthy of someone called "The Liberator," Marx goes on to ascribe to Bolívar moral vices such as cowardice, envy, brutality, disloyalty, boastfulness, despotism, and frivolity. Surely, these are bad traits of character for anyone to have, and especially for a military commander. But Marx here is engaging in a series of *ad hominem* (i.e., arguments against the person) that may be fallacious depending on the facts about Bolívar's actions and attitudes during the Wars of Independence. Evidently, only if based on accurate relevant evidence concerning those facts can criticism of this sort have any force. And even if it has force, it may leave Bolívar's philosophical thoughts untouched. On my view, we should be suspicious about Marx's arguments against the moral character of Bolívar because they count with an exceedingly weak evidential support: Marx's sources were narratives by a few European historians of the time, who probably had no firsthand knowledge of the facts of the war. Without further support, we have no reason to accept, for example, the charge that "the few successes of the corps [of Bolívar] were entirely owed to British officers, such as Col. Sands." Or that once the Spanish threat was reduced, Bolívar "no longer thought it necessary to keep up the appearance of generalship, but leaving the whole military task to Gen. Sucre, limited himself to triumphal entries, manifestos, and the proclamation of constitutions."

On the other hand, the unreliability of Marx's sources is by no means established. His critique has supporters (e.g., Draper 1968) who argue that, charitably construed, Marx objects to Bolívar's propensity to favor authoritarian power, something amply illustrated in the constitutions he wrote for Bolivia and Gran Colombia. True, Marx says that one of those documents qualifies only as a "Bolivian Code, an imitation of the Code Napoleon." Marx, of course, was not inclined to endorse the universal value of democracy, a form of polity that he predicted would become obsolete after the decline of capitalism, when society gets ripe for entering the next historical phase—namely, that of a proletarian revolution leading to socialism. But democracy was a progressive polity for him during the phase of capitalism. On the other hand, the phenomenon of restoration of empires during capitalism, to which he referred as "Bonapartism," was reactionary during that phase, Marx thought, for it implied a regression to absolute monarchies of a previous historical phase, even if they do not involve a return of power to a king.

Could Bolívar respond to the charge of Bonapartism? It is true that he declined the offer of the throne of Gran Colombia. But although he was not interested in becoming its king, he did entertain the possibility of being its president for

life, and also exerted dictatorial powers more than once. Furthermore, during the Wars of Independence he did cultivate a friendship with British imperialism, something that should have been suspicious to Marx—as equally must have been Bolívar’s enthusiastic praise for absolute monarchy.

In addition, numerous writings strongly suggest that Bolívar did not endorse democracy as usually understood as form of polity for Latin America. If this is what Marx really meant to say, I might agree. However, at no point does Marx provide textual evidence that he had read Bolívar. And he does not engage with Bolívar’s arguments, which is something I’ll do next.

There is abundant textual evidence of Bolívar’s sympathies for authoritarian governments. It suffices to look at classic texts such as the “Manifiesto of Cartagena,” “Jamaica Letter,” and the “Angostura Address”—or the constitutions he wrote for Bolivia and Great Colombia. Bolívar repeatedly offers, as a reason for preferring centralized governments, the hypothetical claim that liberal democracy would undermine aggregate happiness, safety for all, and political stability. For this hypothesis, he offers only the dogma that Latin Americans cannot have democracy because of their unique history—and, most important, their unique mixed racial and ethnic backgrounds. On his alternative for them, they should accept the leadership of a centralized government run by privileged Criollos, who were the only people electable in his system. Consider the people of Gran Colombia: had they accepted Bolívar’s constitution for them, they would have had a government with an elected president for life and an elected but hereditary senate dictating all their laws and regulations. In addition, what Bolívar called “Moral Power” would have been vigilant of their actions or beliefs, ready to censor them when considered “immoral.” After 1823, during a brief period as dictator, he in fact created such a branch of government, which in turn made illegal the teaching of Bentham’s writings (in what Andrade and Lugo-Ocando [2018] regard as a concession to the Catholic Church).

REFERENCES

- Andrade, G., and J. Lugo-Ocando. “The Angostura Address 200 Years Later: A Critical Reading.” *Iberoamericana–Nordic Journal of Latin American and Caribbean Studies* 47 (2018): 1.
- Bolívar, Simón. “The Cartagena Manifiesto.” 1812. Manifiesto Portal, <http://manifiestoindex.blogspot.com/2011/04/cartagena-manifiesto-1812-by-simon.html>.
- . “Jamaica Letter.” 1815. In *Latin American Philosophy: An Introduction with Readings*, edited by S. Nuccetelli and G. Seay, 105–19; also available online at <http://faculty.smu.edu/bakewell/BAKEWELL/texts/jamaica-letter.html>.
- . 1819. “Angostura Address.” In Bolívar 1951, pp. 103–22. [Reprint pp. 1–21 in Burke and Humphrey 2007].
- . *Selected Writings of Bolívar*, edited by Harold A. Bierck, Jr. New York: The Colonial Press, 1951.
- Draper, Hal. “Karl Marx and Simón Bolívar: A Note on Authoritarian Leadership in a National-Liberation Movement.” *New Politics* 7, no. 1 (1968): 64–77. <https://www.marxists.org/archive/draper/1968/winter/bolivar.htm>.
- Marx, Karl. “Bolívar y Ponte.” *The New American Cyclopaedia*, 1858. [References to copy in *The Marxists Internet Archive*, <https://www.marxists.org/espanol/m-e/1850s/58-boliv.htm>].
- Nuccetelli, S., and G. Seay eds. *Latin American Philosophy: An Introduction with Readings*. Upper Saddle Brook, NJ: Prentice Hall, 2004.

BOOK REVIEW

Rape and Resistance

Linda Martín Alcoff (New York: Polity, 2018). Paperback, \$24.95. ISBN: 978-0-745-69191-6.

Reviewed by Juan J. Colomina-Alminana
UNIVERSITY OF TEXAS, AUSTIN

Professor Linda Martín Alcoff’s book *Rape and Resistance* is both timeless and timely. It is timeless because it centers on the historical institutional factors that allow and condone sexual violence and assault. It is timely because it demands a recalibration towards the victims’ personal experiences and voices based on prominent, recent activist events that discredit and, even too often, criminalize victim reports by those very same institutions that are supposed to protect but simply ignore and silence them. Taking a stand with other epistemologists of resistance, Alcoff’s main purpose is, then, to develop “a program that focuses not simply on getting the word out, but on reforming and transforming the conditions of reception in the public domains in which our words emerge” (24). Unlike others, though, she does not push a relativistic agenda where every voice is given their space, but a local and “pluritopic” assessment of context where we can base “the transformative and subversive potential of survivor speech [as] absolutely dependent on its truth-value” (54). My comments revolve around this point.

One of Alcoff’s goals is to resist the widely held idea that sexual violence, and particularly rape, is a byproduct of war and militarism, and in general, of concrete deviant situations/events and individuals. And I cannot agree more. As Alcoff shows, we must “denaturalize” rape and sexual violence as an anomaly and, instead, move towards an argument that demonstrates sexual violation is “a systematic sanctioned practice . . . an orchestrated practice” (25). Only then can we make institutions and leaders accountable for such practices. As an example, one can take the case of rape in prisons. The inmate is incarcerated as punishment and, very often, left by themselves once inside. When raped, the mainstream assumes that is simply a byproduct of carceral life, and few ever question the double punishment of sexual violence. Alcoff does not provide further details on how to address this problematic in her book, but it is obvious how her approach to sexual violation would look into this case.

Second, Alcoff accounts for the vilification of the victims by the same institutions that are supposed to protect them—being silenced by certain sensationalist press/media (or all media in general) by the blaming of the victims (“What was she wearing?” “Why did you not defend yourself?” “What were you doing there so late?”), and what, lastly, she defines as a culture of “harsh public criticism and threats” against the victim of sexual violation in general (28). Alcoff labels this “the second rape” (29): the public humiliation and harassment that victims often suffer after reporting the events. (In fact, Alcoff links this with racism as well, since victim vilification may vary depending on the race and class of the victim and perpetrators, which highlights

the racialization of criminal activity. This is a separate issue, which is, of course, very important as Alcoff shows us, but I will not analyze it here.) Thus the “second rape” gives us the impression that “the public visibility of rape and sexual violence is today a complex phenomenon,” and of course it is (28). However, it seems to me that there is some tension between addressing this problematic, even by institutions and media that publicly and on paper say so, and the way that these same institutions actually handle these situations. I am thinking specifically about Title IX policies in universities that are, it seems to me, in place only to protect the “image and reputation” of the institution and not to protect students or victims, as has been obvious in recent polemic (and some still not public) cases, including at my own institution.

Alcoff’s book, though, puts emphasis not on the hidden agenda that many institutions have but rather upon what she sees as:

a contestation over the epistemic and discursive terms in which sexual violations can enter the larger public domain [and] the conventions about who can speak, what they can speak about, who will be accorded the title of expert, or credible witness, how the circulation of speech occurs, and what the subsequent effects of the speech will be in both discursive and extra-discursive, as well as legal and extra-legal, arenas. (29)

Of course, this raises some kind of tension, or even contradiction perhaps, between allowing and encouraging victims to speak up and the ways that their speech is handled. At this point, Alcoff provides some suggestions to alleviate the tension: We should all understand how victims’ speech is framed and circulated; we should build better practices about the selection and handling of stories, and about institutional policies (and even ethical norms) to encourage a more responsible way of covering stories (30); and we should not forget that these norms and practices could only apply to and affect institutions, and never individuals, if we actually want to preserve primary liberties and rights. In brief, Alcoff suggests that we, as a community that cares about victims, should “aim for a world in which the patterns in which our speech circulates can become more perspicuous as well as subject to critical analysis” (31).

Accomplishing this, Alcoff says, requires individuals to be educated in a way that encourages the development of what José Medina labels “meta-lucidity”: The second order capacity of reflection upon the ways in which one arrives to know something, the skill of being aware of one’s own biases, which includes, as Alcoff insists, awareness regarding those things upon which we are “meta-insensitive,” as well (that is, blind to one’s own insensitivities). The strategy that Alcoff proposes, then, is to educate with permanent scrutiny on how we learn, know, and reason at a meta-level, how to organize our attention and perception, and, even better, how we arrive to perceive things and produce habits that cross into our conscious experience and affect our comprehension of these very same sources of information that often are indicative of certain structures of inequality and discrimination.

For some epistemologists of resistance, such as Medina, the solution is simply to become acquainted with such conditions of perception in order to be aware of the “limitations of dominant ways of seeing,” to make sure that our perceptual habits are not unfounded or unjustified because the whole point is to be considered as epistemically responsible agents under conditions of oppression. For Alcoff, however, the point must be to change the contextual conditions in which institutions form and enforce those very same ways of seeing, and change existing conditions for many other ways of seeing.

The problem with Medina’s strategy is that perhaps one cannot confront by herself what is called “epistemic injustice,” since one can be pushed towards what Luvell Anderson has called “hermeneutical impasses.”¹ Or, in other words, how is one to modify her own epistemic habits simply because she witnesses some conflict between her beliefs and frameworks of interpretation? It is perfectly possible to think of a situation where one may be aware of some kind of injustice, but what about someone who is not? How should we address this, especially given that those who do not see things as unfair dominate institutions and institutional narratives? Medina’s strategy seems a little bit too simplistic and relativistic. Alcoff provides myriad examples of the increased visibility of cases of sexual harassment and aggression and the increase of public consciousness about it. However, there is a conundrum here, since we have seen very little justice for victims and only a symbolic punishment of very few notorious cases (with a lot of pushback from very prominent people and institutions). Alcoff says that this dissatisfaction comes from the fact that Medina got things right at a descriptive level: Epistemic friction will always exist because there will always be different interests and frameworks of information. However, Alcoff insists, the prescription does not work, since not everything can or must be allowed. So the only appeal is to some kind of coherence or consistency between the set of views at play.

According to Alcoff:

The potential benefits of epistemic friction in encouraging more meta-lucidity occur because we strive for coherence in our lives, some basic consistency. . . . The epistemic pay-off of friction occurs when there is communication across the silos and a motivation to rethink practices in each domain *in light of others*. . . . The pay-off of epistemic friction in enhanced meta-lucidity comes about because we are trying to reduce the friction. (33)

To make the claim for coherence, Alcoff states that we must put our attention onto the notion of “echoing/echoability” (34). The difficulty and special demands of echoing sexual violation seem to isolate the cases of rape. However, they do not. They are institutionally embedded and permitted, so the echoability of victims’ speech (the way that it expands its scope) must work in the same ways that other organized protests work. We should work together in developing nets of safety and knowledge by sharing methods of communication for victims that could

generate an understanding of “the limitations of dominant ways of seeing” (39).² That is, we require some kind of appropriation for the voice movement, where, first, one can report or share experiences in a judgment-free safe space and, second, the negative image of the survivor—which arises because their credibility is called into question and/or because of hidden institutional agendas—is eliminated. In other words, to eliminate the repressive culture against the victim and the reporters, we must first create a safe space for all where the raped person and her allies are neither criminalized nor retaliated against.

Since what is coherent or consistent may come from somewhere or someone else, and that can exclude our beliefs, sources, and views, we can’t simply let institutions and the people who represent them off the hook. In addition, Alcoff explicitly mentions that this is not a solitary enterprise, but appeals to individual cases (for example, the case of a Muslim activist, fearing repercussions for speaking up, who uses anonymity for fighting against certain traditions in her community, supposedly in the name of everybody who is against such practices) (42). In other words, the concerns are these: How can concrete victims’ reports have a truth-value if there are no available truth-conditions that sanction these concrete cases in particular, individual terms? How can we believe the reports when we still have an immoral institutional culture where the victim and their allies are going to be criminalized and punished? Because of this, we have much work to do in changing the truth-conditions (and the institutional culture) of our common ground. Because Alcoff has written an important and necessary book, we have some hope for how to shift the context of truth-value for survivors.

ENDNOTES

1. See “Hermeneutical Impasses,” *Philosophical Topics* 45, no. 2 (2017): 1–19.
2. See José Medina, *The Epistemology of Resistance* (Oxford: Oxford University Press, 2013), 47.

CONTRIBUTORS

Arturo Aguirre Moreno is research professor in the Department of Philosophy at Benemerita Universidad Autonoma de Puebla (Mexico); he was a postdoctoral visiting researcher at the Higher Council for Scientific Research in Spain (CSIC) and a Fellow at UNESCO/Keizo Obuchi. His line of research is contemporary violence within the framework of a forensic philosophy. He can be reached at aguirre.arturo@yahoo.com.

Eric Bayruns García is a PhD candidate at The Graduate Center at the City University of New York. He specializes in philosophy of race and epistemology. His current philosophical work focuses on the nature of race concepts and epistemological questions raised by race and racial injustice.

Damián Bravo Zamora studied a BA in philosophy and an MA in philosophy of science at Universidad Nacional Autónoma de México (UNAM). He obtained his PhD from Trinity College Dublin, writing a thesis on Kant’s critique of metaphysics and its relation to contemporary debates on paradoxes regarding absolute totalities. For two years, he was a research fellow at Universidad Autónoma Metropolitana, Mexico City, and currently teaches philosophy at UNAM and Universidad Anáhuac México.

Alan Yohualli Chavoya is a PhD candidate in philosophy at Northwestern University. His research focuses on Latin American and Latinx philosophy, critical philosophy of race, liberation philosophy, and contemporary political philosophy, with a primary focus on immigration.

Juan J. Colomina-Alminana is an assistant professor at the University of Texas at Austin. His work primarily deals with metaphysical and linguistic issues. His last book, *Formal Approaches to the Metaphysics of Perspectives* (Springer, 2018), defends a new account to knowledge: modal perspectivism.

Sergio Rodrigo Lomelí Gamboa works at the Centro Interdisciplinario de Investigación en Humanidades at the Universidad Autónoma del Estado de Morelos, Mexico (UAEM). He has a PhD in philosophy from the Universidad Nacional Autónoma de México (UNAM) and specializes in political philosophy with particular interest in social theory and heterodox Marxism.

Susana Nuccetelli is professor of philosophy at St. Cloud State University in Minnesota. Her articles in Latin American philosophy, ethics, and philosophy of language have appeared in many journals, edited volumes, and reference books. She has edited and co-edited several volumes, and is the author of *Latin American Thought: Philosophical Problems and Arguments* (Westview Press, 2002), and the co-author of *Engaging Bioethics* (Routledge 2017) and *How to Think Logically* (2e, Pearson, 2012).

Omar Rivera is associate professor of philosophy and affiliate professor of Latin American border studies and feminist studies at Southwestern University. His research centers on Latin American and Latinx philosophy, aesthetics, and phenomenology. He is the author of *Delimitations of Latin American Philosophy: Beyond Redemption*, forthcoming with Indiana University Press.

Stephanie Rivera Berruz is assistant professor at Marquette University. She received her PhD in philosophy from SUNY Buffalo in 2014. She is the recipient of the Woodrow Wilson Career Enhancement Fellowship (2017-2018) for her work on Latinx feminisms and Latin American philosophy. Her main research interests lie in Latin American philosophy and Latina feminism as well as philosophy of race, gender, and sexuality. She approaches these topics at their intersections as she is committed to the importance of diverse approaches to philosophical praxis. She recently published a co-edited anthology: *Comparative Studies in Asian and Latin American Philosophies* (2018), and her work has been published in *Hypatia* (2016), *Inter-American Journal of Philosophy*, and *Essays on Philosophy*.

Jorge M. Valadez, Professor Emeritus, Our Lady of the Lake University, has a philosophy PhD from Yale University. He has published in the ethics of immigration, deliberative democracy, multicultural philosophy, and indigenous rights. He is a Mexican immigrant, was a migrant farmworker in his youth, and is writing a book on planetary ethics and the future of humanity.

Alejandro Vallega works on Latin American thought, decolonial Latin American thought, and contemporary European philosophy. His focus is aesthetics and narratives of existence. He teaches Latin American thought at the University of Oregon graduate and undergraduate programs. Among his published books are *Latin American Philosophy from Identity to Radical exteriority* (Indiana University Press, 2014-2015) and *Tiempo y Liberación* (AKAL, 2019).

Allison B. Wolf, PhD, is a professor of philosophy at Simpson College. Her research is in feminist philosophy, philosophy of immigration in the US and Latin America, and feminist bioethics. Her work has been published in various journals and collections, and she is currently writing a book entitled *Just Immigration in the Americas: A Feminist Account*, under contract with Rowman and Littlefield International.



APA NEWSLETTER ON

LGBTQ Issues in Philosophy

GRAYSON HUNT, EDITOR

VOLUME 19 | NUMBER 1 | FALL 2019

FROM THE EDITOR

Grayson Hunt

UNIVERSITY OF TEXAS AT AUSTIN

This year's *Newsletter on LGBTQ Issues in Philosophy* features three original essays and one book review.

The first essay featured, "Assumptive Care and Futurebound Care in Trans Literature" is written by Dr. Amy Marvin. In it, Marvin adds her name to the list of feminists who have critiqued Carol Gilligan for "mainly centering the voices of white, middle-class women" a move that renders Gilligan's vision of care, an assumptive, "thus participating in the historical suppression of black feminist ideas even while challenging traditional philosophy" (Marvin, citing Patricia Hill Collins). Marvin notes that Gilligan's *In a Different Voice* also excludes trans women as ethical knowers. We should not be surprised at this since, as Marvin notes, "psychologists still tended to link trans people with pathology, as 1982 introduced the misgendering dichotomy between categorizing trans women as 'homosexual' or 'heterosexual males' picked up by later theories of autogynephilia." Gilligan, like many other feminist theorists, failed trans women, and Marvin's essay is no doubt a corrective to this failure. Marvin turns to Maria Lugones's discussion of arrogant perception and world-traveling and then to trans literature, notably Casey Plett's *A Safe Girl to Love* and Ryka Aoki's "To the New World," in order to argue that trans literature offers rich descriptions of the harm of assumptive and the value of futurebound practices of care that emphasize this dynamic of taking away and giving the other space, ultimately demonstrating the harm of assumptive care and the world-making power of future-bound care.

The second and third essays featured offer theorizings of trans and queer existence from the first-person perspective. Dr. Rachel McKinnon's essay, "Participation in Sport is a Human Right, Even for Trans Women," offers readers a defense of transgender athletes' right to sport and a critique of the faulty logic used to undermine the human right to sport for trans and non-binary people. McKinnon notes that the arbiters of sport disputes, the Court of Arbitration for Sport (CAS), ruled in 2015 that there are only two categories of competition (male and female) and that those two categories cover all athletes, meaning that in order to compete, athletes must compete in one of those categories, a move that in effect makes inclusion is sport difficult and often impossible for intersex, non-binary, and transgender people (especially trans women). McKinnon

concludes her article by challenging the "unfair competitive advantage" argument used to disqualify transgender women. Noting that there are already, amongst non-trans women athletes, huge differences in natural physical traits (an obvious example being height), and given the nature of elite sports where competitive advantages are what athletes train for, fairness does not require us to exclude those who have considerable performance advantages.

The final essay, written by Dr. Alice MacLachlan, is also about exclusion. In her essay, she reveals the many ways that queer families are absent from philosophical and bioethical discussions of assisted reproduction. For MacLachlan, the exclusion reveals the narrow approach to moral issues of alternative family-making and assisted reproduction, in part fueled by a heteronormative emphasis on technology, secrecy, and ancestry. MacLachlan argues that discussion of assisted reproduction within philosophy and bioethics tend to focus on four issues, only the first of which directly relates to her queer family's experience. She describes the four issues as "(i) the various rights and responsibilities of all involved insofar as these illuminate the necessary and sufficient conditions for moral parenthood; (ii) the growing role of technology in human reproduction; (iii) the value and significance of genetic ties, and knowledge of/access to one's genetic ancestors; and (iv) the ethics of secrecy and deception around someone's origins." MacLachlan notes that tendencies within bioethical and philosophical discussions of assisted reproduction (noting, as an example, David Velleman's work on the topic) has a normalizing effect that elides many queer families who are uniquely and differently situated compared to "bionormative parents (each of whom has a genetic relationship to the child) and adoptive parents (neither of whom does) and even step co-parents (they enter their child's life at different times while we became parents together)." MacLachlan demonstrates how paying attention to queer family narratives places collaboration at the center of creating a family.

In her timely review of Gayle Salamon's *Life and Death of Latisha King: A Critical Phenomenology of Transphobia*, Chris Jongchao Ma offers readers an overview of the 2018 monograph, identifies some of its key strengths, and asks questions provoked by Salamon's phenomenology of transphobia. The strength of Salamon's critical phenomenology is found in part in its capacity to reveal how a life may only emerge in a shared world, one which Ma notes is hostile to trans people of color. Referring to Merleau-Ponty's concept of horizon and Salamon's analysis of collaborative meaning, Ma notes how transphobic power relations extinguished Latisha's meaning as a transgirl. The concept of horizon, Ma says, "allows us to understand

the detrimental effects that power relations can have: my meanings may be canceled, may not be able to emerge, and other meanings might be read into me, against me, as 'power can delegitimize other horizons, can render them as untrue, invisible, or unthinkable'" (Ma quoting Salamon, 92). One question that I found particularly fruitful and necessary comes at the end of her review: "What role does the de-racialization of King play in the courtroom and in media, in relation to her non-existent transgender identity?" How has trans identity been racialized or white-washed in the US?

ARTICLES

Assumptive Care and Futurebound Care in Trans Literature

Amy Marvin
UNIVERSITY OF OREGON

INTRODUCTION: DIFFERENT VOICES, TRANS VOICES

Carol Gilligan's *In a Different Voice* critiques Lawrence Kohlberg's stages of moral development for universalizing the moral development of men and boys, arguing that Kohlberg casts women and girls as morally immature in comparison to a masculine telos toward justice.¹ In contrast, Gilligan centers the moral development of women and girls on their own terms, defending an ethics of care. An ethic of care, Gilligan describes, involves "seeing a world comprised of relationships rather than of people standing alone, [and] a world that coheres through human connection rather than through systems of rules."² Gilligan thus establishes a broad range for an ethics of care: "the vision that self and other will be treated as of equal worth, that despite differences in power, things will be fair; the vision that everyone will be responded to and included, that no one will be left alone or hurt."³ This emphasis on connection rather than an ethics of disconnected individualism is part of the broad tapestry from which feminist care ethics has been woven. Feminist care ethics does not necessarily entail giving up justice and autonomy, but it does take seriously the challenge that a relational view of the self poses to traditional philosophical theories that solely focus on an individual, disconnected moral agent while bracketing out human dependency, interdependency, and social ties.⁴

Though Gilligan's *In a Different Voice* strives to center ethical perspectives erased by traditional philosophical work, its project of care risks continuing to center some ethical voices over others. First, the project of care ethics risks circumscribing its understanding of care by ignoring differences among women, including race and class, that may result in different understandings and practices of care.⁵ For example, in *Black Feminist Thought* Patricia Hill Collins critiques Gilligan's project for mainly centering the voices of white, middle-class women, thus participating in the historical suppression of black feminist ideas even while challenging traditional philosophy.⁶ Second, this bias

within feminist care ethics risks obscuring the ways that rhetorics of care can be wielded as tools of oppression and imperialism. For example, Uma Narayan argues that rhetoric about care over dependents has functioned ideologically in combination with racism and imperialism to justify colonization, taking the form of "paternalistic caring."⁷ Despite Gilligan's emphasis on different voices, it is thus important to pay attention to which voices continue to be excluded from the ethical fold, as well as how these exclusions might obscure insidious deployments of care.

Among the women not included by Gilligan's work are trans women, and I begin this essay focusing on Gilligan rather than other care ethicists to consider the broader history of trans women's exclusion from ethical knowledge. The same year that Gilligan published *In a Different Voice* to center women and girls as ethical knowers, psychologists still tended to link trans people with pathology, as 1982 introduced the misgendering dichotomy between categorizing trans women as "homosexual" or "heterosexual males" picked up by later theories of autogynephilia.⁸

Feminist philosophy fared little better with taking trans women seriously as knowers during this period, indicated by Marilyn Frye's afterthought describing trans women as mindless robots in her otherwise excellent 1983 book *The Politics of Reality*. Frye argues that under a phallographic scheme subordinating women to men, women are not considered to have an authoritative, distinct point of view or manner of perception, instead "assumed to be robots hooked up to the senses of men," ultimately lacking a distinct soul or the ability to see beyond their appointed task of service and submission.⁹ Because women are in fact seers and authors of perception,¹⁰ technological advances in "male to female transsexual reconstruction"¹¹ aim to replace the threatening real women that men depend upon by "constructing actual robots" in the form of trans women.¹² While non-trans lesbians rise against this conceptual order as "women-seers,"¹³ trans women are framed by Frye as the true robotic replacement, internalizing male norms and continuing to serve men without wielding the potentially disruptive capabilities for perception, a point of view, and a soul beyond submission. Frye thus frames trans women as the realization of men's designs for women under patriarchy, mindless servile robots without ethical agency who do not question their subordination and are fundamentally questionable as knowers, let alone seers. During the time period Gilligan began to center the ethical lives of non-trans women and girls, trans women's ethical lives were largely obscured by discourses of pathology or placation, reduced to a figure of perversion, piteousness, or Stepfordism beyond ethical agency.

In this essay, I depart from this historical exclusion of trans women's ethical insights from care ethics by focusing on trans literature as a source of knowledge expressed by trans women about care. Specifically, I discuss short stories by Casey Plett and Ryka Aoki as sources of knowledge about assumptive and futurebound practices of care. I begin by arguing that the work of Maria Lugones considers the spatial dimensions of love and that this suggests that care involves giving space to another on their own terms. I then turn to Plett's story "Other Women" to unpack

assumptive care as a short-circuited form of caring in which a relationship to the other on their own terms is closed off by transphobic assumptions. I contrast this with Plett's story "Winning," which describes a futurebound care in which another's space is kept open, permitting action that allows them to flourish on their own terms. After this, I discuss Aoki's "To the New World" as an engagement with both racist assumptive care and futurebound care aimed at self-care. I conclude that these stories provide a rich space for considering trans ethical knowledge about care.

LOVE, CARE, AND SPACE

Before turning to stories in trans literature, I find it useful to philosophically justify the relationship between care and giving another space that I discuss throughout this essay.¹⁴ In "Playfulness, 'World'-Travelling, and Loving Perception" Lugones suggests a relationship between love and space through her emphasis on "worlds" when contrasting arrogant perception and loving perception. Lugones begins by discussing arrogant perception as a failure of identification with another and a failure of love.¹⁵ For Lugones, arrogant perception is not only directed towards women by men, but can also happen among women. Specifically, Lugones reflects on her arrogant perception towards her own mother after she was raised to treat her as a servant and "could not welcome her world," as well as the arrogant perception directed by many white women towards women of color.¹⁶ Lugones writes, "there is a complex failure of love in the failure to identify with another woman, the failure to see oneself in other women who are quite different from oneself."¹⁷ This failure of arrogant, walled-off perception towards another is not always immediately obvious to the perceiver, even though it can result in abuse towards others.¹⁸

Loving perception among women involves being aware of connection rather than the isolated, disconnected compartment of arrogant perception. Lugones emphasizes, "I am incomplete and unreal without other women. I am profoundly dependent on others without having to be their subordinate, their slave, their servant."¹⁹ Loving another requires being able to perceive their world as distinct from one's own but also as connected, including the ability to perceive one's own actions and their meaning from the positionality of another person. Lugones argues, "We are fully dependent on each other for the possibility of being understood and without this understanding we are not intelligible, we do not make sense, we are not solid, visible, integrated; we are lacking."²⁰ Loving perception thus requires an awareness of a shared space across different positionalities, refusing to level over the other like one would a servant, but also acknowledging a space of connection across different "worlds," defined as a multitude of situated positions one might inhabit, see from, or be seen from.²¹

The vantage point of arrogant perception, which Lugones describes as a way of being taught to love that is not love,²² involves an inability to allocate any space to another outside the perspective from which they are seen. The other is locked into an arrogant perceiver's perspective to the extent that Lugones associates this way of relating to another with abuse, with the example of "using, taking for

granted, and demanding her [mother's] services in a far reaching way that . . . left her little of herself to herself."²³ In this way arrogant perception may involve a restriction of space taken in both relational and material senses, denying the other a space to be seen or physically inhabit space on their own terms. In contrast, loving perception gives another more space to be understood from their distinct perspective. This way of viewing another does not constrict, but instead opens possibilities for difference, involving a practice of traveling to another's world that allows someone to not only "understand *what it is to be them*" [emphasis hers] but also opens up a space for the traveler to consider themselves from a different perspective.²⁴ Loving or arrogant perceptions are thus tied to the amount of space given to another on their own terms, acknowledging their distinct but connected living perspective and positionality.

Lugones's discussion of perception and world-traveling, while not explicitly about care, nonetheless suggests that care may also have its own spatial character. Lugones's work has been described as a reflection on the receptivity required for the "cared-for" to be heard by the "one-caring,"²⁵ and Maurice Hamington sees in Lugones's "World-Travelling" essay a relationship between the openness of world-traveling, the expansion of understanding, and the cultivation of "caring performances."²⁶ Lugones's work also shares an affinity with Gilligan's emphasis that an ethics of care focuses on connection, relationality, interdependence, and inclusion in the context of relationships. Lugones's work is thus helpful for considering feminist care ethics in relation to the space taken away from or given to another. I will argue in the following sections that trans literature offers rich descriptions of assumptive and futurebound practices of care that emphasize this dynamic of taking away and giving the other space.

ASSUMPTIVE CARE

In *A Safe Girl to Love*, Casey Plett provides several stories about trans women protagonists that include rich topographies of care, including the opening story, "Other Women," where Plett provides insight into traumatic forms of care imposed by cis (non-trans) friends and relatives. "Other Women" follows the main character, Sophie, as she ventures home to Winnipeg for the holidays after transitioning. Focused on Sophie's perspective, the story highlights how Sophie's cis friends' and family's skewed vision renders their practices of care assumptive and damaging.

At the beginning of the story, Sophie's relationship with her mother and extended family may seem promising. The first family member Sophie interacts with is her mother, Lenora, who seems friendly towards Sophie, even though we are informed that their relationship was strained after Sophie came out as trans. Sophie recounts, "It was hard after that. She stopped signing her e-mails 'Love, Mom.' Stopped going out in public with me. Those were rough months,"²⁷ and they only started talking again after Sophie moved away.²⁸ Additionally, Sophie's grandparents seem jovial, offering handshakes, hugs, food, and coffee.²⁹ After Sophie's initial visit with her grandparents, her mother is optimistic, saying, "It's nice they haven't called you by your old name, don't you think?"³⁰ It thus initially seems that

Sophie may be able to maintain some bonds of care with her family.

Sophie's interactions with family quickly reveal their tenuousness. Already Lenora had asked Sophie what she would wear to church on Christmas, and seemed uneasy about Sophie's plan to wear a dress. When asked if this was okay, Sophie's mother responded, "Well, I don't think your clothes will be a problem, no."³¹ Implicit in this response is the sense that it is Sophie herself as a trans woman wearing a dress to church who is the problem, and Sophie ends up sitting alone during the service at her mother's suggestion.³² Sophie's family may be "tolerant," but there is a rift between their worlds.

This rift is exposed during Christmas dinner. While Sophie's grandfather is dispensing blessings through prayer, he offers one to Sophie under her deadname,³³ saying, "We are especially thankful you saw fit to guide up from America our grandson, Leon."³⁴ Instead of accepting this misgendering blessing, Sophie interrupts by saying, "No," which is responded to by a sharp intake of breath by her grandmother and her mother stressing "hey now" in a quiet, angry voice. It is all downhill from here, with Sophie's apology to her grandfather met with a frown and a "well," and Sophie finds a note in her coat pocket from *Proverbs 3:5-6* written in her grandmother's script: "*Trust in the Lord with all your heart and lean not on your own understanding; in all your ways acknowledge him, and he will make your paths straight.*"³⁵

Here, I suggest, the novel does show Sophie being responded to by a practice of care, highlighting specific forms of care that may emerge when cis people try to care for trans relatives. One might be tempted here to say that these moments are in fact *not* care for Sophie, and instead constitute an ethical failure due to an absence of care. Drawing from Lugones's understanding of love, care, and space discussed in the previous section, Sophie's family directs a form of care towards her that is contrary to her own self-understanding, and the ways Sophie navigates her position in the world. In wanting to sit separately from Sophie during church, or praying for Sophie under her deadname, Sophie is *not* acknowledged as a complex person with a rich inner (and outer) life, except insofar as she is abjected or forced into silence about the ways she lives her life and understands herself. The letters and prayers offered to Sophie are thus only offered in ways that annihilate her life as a trans woman, denying her the space to live on her own terms due to a failure of maintaining fair, inclusive connections and, ultimately, care.

However, there is also reason to suggest that Plett's story provides knowledge not only about care's absence, but also about the transformation of care through the insidious nuances of transphobia as it conditions everyday relationships. For cis people living in a transphobic society, the horizon of caring practices may be restricted to the smallest gestures of accommodation. As evidenced by Lenore's optimism about Sophie not getting entirely reduced to her deadname, such a meager form of tolerance may be taken as indeed an attempt to care for Sophie, and for some cis people may even be experienced

with difficulty and hence as an effort for inclusion. Lenore interprets moments when Sophie is not misgendered, deadnamed, or barred out of her grandparents' home altogether as *landmark* moments of care and inclusion, without recognizing the more insidious topography of care practiced in these cis familial spaces. This is care as mediocre tolerance, exhibiting a facile pride in its receptivity to a trans other while unknowingly stripping their world away. Likewise, Sophie's grandmother, given her religious background, may see her note as a crucial intervention to bring her child back into the fold of family life and its relationship to the sacred. This care is unable to grant or even fathom that Sophie might have her own world, restricting her space during the entire encounter through its transphobic terms. The problem may not be that Sophie's family isn't caring, but rather that their caring *sucks*. In this case, transphobia may restrict our relationships to others through their assumptions about who we are and should be, including loved ones privy to our most vulnerable moments. In a transphobic society, even well-intentioned people can obliterate us with their care.

Sophie's friend Megan, who she frequently turns to for support, provides another example of assumptive care. Initially, Megan seems like someone who cares about Sophie. We are informed that Megan is the first person Sophie came out to as trans, to which she responded with encouraging words: "One day everyone will call you Sophie. And I'm going to be so fucking proud of you when they do."³⁶ Additionally, Sophie often spends time at Megan's place as a moment of respite from uncomfortable scenes with her family,³⁷ and Megan seems accepting and open about Sophie's transition when they are reunited at the beginning of the story.³⁸ Megan also portrays herself as supportive in relation to Sophie's friends who were angered upon hearing of Sophie's transition, reassuring Sophie with "I told [them] that if anybody gave you shit tonight that I would fucking murder them."³⁹ Megan thus seems like someone who might have a relationship of care with Sophie that does not involve projecting cis assumptions that restrict her vision of who Sophie might be and what caring for a trans woman friend might look like.

Megan's practice of assumptive care with Sophie is revealed in one of the worst violations of Sophie's autonomy and understanding of herself in the story. After attending a party where Sophie had to deal with a transphobic friend from her past, drank lots of alcohol, and fell over onto the sidewalk,⁴⁰ Megan drives Sophie back to her place. Megan and Sophie become intimate,⁴¹ but Megan is distraught when Sophie's parts no longer behave the way they used to. Sophie explains, "and then I didn't want her touching that part of me, I hated how she talked about it. Like if my dick didn't seem happy then I couldn't be . . . I wanted to lift her off of me and say *my penis isn't me.*"⁴² Despite Sophie's forceful nos, Megan proceeds to give her oral sex, violating consent and causing Sophie to experience intense dysphoria about her body.⁴³ After Sophie tells Megan to stop two more times, she finally ceases but also expresses anger at Sophie. Sophie recounts, "she looked up, furious. No, she said. I'm sorry, don't worry, I won't *bother* you again tonight. She looked like she was about to cry."⁴⁴

As Talia Bettcher emphasizes, denials of trans people's authenticity frequently share an affinity with sexual violence,⁴⁵ and Megan's refusal to stop also signals a refusal to grant Sophie the space for authenticity over her body and what it means for her. The temporality of their sex is restricted to expectations about cis bodies, culminating in Megan's tears of frustration when she interprets Sophie's "NO" as a slight. While Megan may think of herself as trying to care for Sophie, the wounded anger she expresses indicates that Megan does not understand her care was actually a violation, and her warped care is used in the service of violence through assumptions about sex, trans women's bodies, and caring intimacy. Through this assumptive care, what began with Megan tenderly reassuring Sophie, "You're so soft . . . you're so fuckin' pretty"⁴⁶ escalates into a violation of consent and an angry silence. Megan's earlier "tough love" of telling Sophie to stop being a dick is not unrelated to her inability to see Sophie as other than a dick in the context of sex, combined with the violence of her unwillingness to listen and stop.

"Other Women" thus indicates that practices of care by cis relatives and friends can often be a trauma and a terror rather than a straightforwardly benign force of inclusion and attention to particularity. The assumptions that cis people have about their trans family, friends, and lovers may find their way into their practices of care, and this assumptive care constitutes a real ethical failure. Assumptive care takes up the other's space by refusing to grant them a world, instead practicing care from the totalizing perspective of the perceiver in a way that can cause harm and violence. Through assumptive care, encounters with the other are restricted to predetermined assumptions without creating an open space through which they can respond or change the situation to better fit their terms. Plett's story thus provides knowledge about care when it is circumscribed by the assumptions of the person practicing care, ultimately short-circuiting because it is unable to escape expectations and truly connect to the other with their complex, lived positionality and needs.

FUTUREBOUND CARE

In addition to knowledge about ethical failures, trans literature also provides knowledge about rich ethical moments, such as those found in futurebound care. To draw out futurebound care, I will focus on Plett's story "Winning," which considers care between a trans mother and a trans daughter who returns to her hometown. The relationship between Zoe (daughter) and Sandy (mother) is frequently fraught, like many mother-daughter relationships. When Zoe came out to Sandy as trans on the phone, for example, Sandy responded with "oh no" and said she needed "a few weeks on this,"⁴⁷ after which Zoe didn't hear from Sandy again for months.⁴⁸ Sandy has also slapped Zoe before, and at one point digs her nails into Zoe's thighs to get her attention about trans women needing to dress cautiously.⁴⁹ It turns out that a trans daughter having a trans mother is not always easy, much like cis daughters having cis mothers, and Sandy often seems to be coming from a place of trauma when caring for Zoe. At one point Zoe wonders "if it was possible to love so fully, as Sandy had, unless you had been cut off at some point from that love yourself."⁵⁰

Nonetheless, Sandy also cares for Zoe, and Zoe cares for Sandy. Zoe was initially hesitant about staying with her mother again after living in New York City, but notes that they were able to generally "re-learn some good parent-kid relationship things."⁵¹ Sandy also makes gestures of care towards Zoe, for example, insisting that Zoe takes a nap to cure a hangover and take up time she could be using to help Sandy pack.⁵² The story may indicate tensions between mother and daughter, but it also emphasizes their attempts to maintain caring and attentive ties to each other.

This complex care allows Sandy's care to become futurebound. Earlier in the story, Zoe hears that while she was away her friend from high school, Frankie, lost her father, became pregnant, adopted her kid out, and got into cocaine.⁵³ Frankie holds a special place for Zoe because she provided her care and help with planning to transition. Plett writes, "Frankie once had looked at Zoe and cupped her chin and said God, you're beautiful. She was never exasperated with Zoe's tics. . . . She'd only taken care of Zoe, sisterly and lovingly."⁵⁴ Uncertain of Frankie's fate, Zoe texts her twice to ask if she needs anything but receives no response.⁵⁵ The subject of homelessness frequently appears in the text, including news of "student-bro types" shooting homeless people with BB guns⁵⁶ and the story takes place in the cold, dark part of the year.⁵⁷ At one point, Zoe spots a homeless woman who she thinks might be Frankie but she (unconvincingly) concludes "*it probably wasn't her.*"⁵⁸

Towards the end of the story, Sandy expresses a moment of here-I-care-and-can-do-no-other in response to Frankie's uncertain situation. When Sandy asks Zoe if she has seen Frankie yet, Zoe fills her in on what she had heard so far.⁵⁹ Sandy responds, "So you're telling me . . . that one of your old friends who no longer has parents and just had a baby who she had to give up is possibly on the streets right now and possibly has problems with serious fucking drugs and all you did was text her."⁶⁰ Zoe's two text messages, Sandy suggests, are an inadequate attempt at care. In anger, Sandy calls for ethical commitment and action, as she continues, "None of you get it around here! You don't know what friendship means! No, it could never mean life and death, could it? Humans aren't fucking games where you just try your best! . . . why am I about to waste time yelling. Let's get in the car."⁶¹ Sandy calls for the kind of care that results in action, and seeks to help Frankie when she is most vulnerable and isolated instead leaving her to the fate of abstract text message requests. As Eva Kittay stresses, one does not truly care unless they act.⁶²

Through Sandy's call to action, taking the form of hopping in the car with Zoe and trying to find Frankie, care becomes futurebound. Already, Frankie had been a source of futurebound care, providing Zoe with the help she needed to eventually actualize transition. Sandy's call to care, likewise, does not abandon Frankie to the rumors that she is alone and on the street, but instead cares *enough* for the situation to be otherwise. Futurebound care, as found in Plett's story, is that which makes living with and for another possible in the most radical sense, both envisioning and ethically committing towards a better possible future world. The other is not only given room for their perspective to

be centered on their own terms, but also the space they need is further opened such that the other's perspective can be granted a possible future of survival and growth. In contrast, assumptive care might project a future for the other, but its ameliorative potential remains limited by the transphobic horizons influencing the person practicing care while its potential for abuse and violence is enhanced. Futurebound care, in contrast to the closed character of assumptive care, invites a structure of open, ethical world-making with others.

Returning to Lugones, futurebound care not only avoids the trap of walling oneself off from the other, but also goes beyond just giving another person space by taking real action so the other can live and grow on their own terms. Care becomes connected amelioration, realizing the other's needs by taking the actions required to bring them to a more empowered space, while also avoiding the potential trap of projecting distorted self-imposed goals onto the other and getting trapped in assumptive care. The story ends before Zoe and Sandy are able to reach Frankie, and active listening beyond Sandy's initial call to action would be required for her futurebound care to not lapse into assumptive care, as would happen if Sandy merely assumed she already knows everything Frankie needs. Nevertheless, Sandy's call to get in the car is itself a moment of futurebound care that, if sustained, could potentially result in a reconnection with Frankie and a better realization of her needs and goals.

ASSUMPTIVE CARE IN ITS (RACIST, TRANSPHOBIC) FEMINIST DIMENSIONS

Thus far I have looked at Plett's stories "Other Women" and "Winning" as sources for knowledge about care. Assumptive care, as seen in "Other Women," is bound by the closed structure of transphobia, and hence easily results in a leveling over of trans experience and a violent denial of consent. Futurebound care, in contrast, involves an ethical action that opens up a space for the other to have a future more on their own terms.

We might leave my initial analysis of care in trans literature with the false idea that assumptive care and futurebound care are primarily related to one's trans status and not also about intersections between trans status and other crucial aspects of trans people's lives such as class and race. Tethering insights about assumptive and futurebound care to trans status alone would be a mistake, and erase the experiences of trans people across difference. I will thus expand my analysis by centering ethical knowledge from Ryka Aoki's story "To the New World," arguing that her story highlights assumptive care and futurebound care as shaped by trans status, race, and immigration, while distinctly linking futurebound care with self-care.

Aoki's "To the New World" describes Millie Wong's visit to a farmer's market in Los Angeles to pick up food for her grandmother's birthday dinner.⁶³ Though Millie's grandmother passed away, Millie sees the occasion as a way to maintain family ties after separating herself immediately after transitioning.⁶⁴ Here, Aoki focuses on race and gender in the context of Millie's everyday life as an Asian American trans woman. In the marketplace Millie

relates her study of other women's patterns of walking and speech with her immigrant family studying the TV and radio to lose their accents.⁶⁵ Studying the people around her also allows Millie to hear "snippets of conversation that convey gender, ethnicity, even social standing."⁶⁶ Millie is thus attentive to the social location of the people around her, and the way this informs even the smallest inflections in the language they use. The marketplace is a place where Millie both watches and is watched, and she soon runs into the other main living character of the story, who is the primary source of assumptive care.

The trap of assumptive care is set for Millie when she happens upon a loaf of sweet bread. Millie links the food back to her grandmother, with the narrator informing us, "Her grandmother had loved that sort of bread, especially towards the end when she couldn't chew very well . . . [Millie] remembered pulling off small bits of steamed bun and feeding them to her grandmother with fragrant, lukewarm tea."⁶⁷ Based on memory, the bread thus seems like a clear choice for celebrating her grandmother's birthday, and Millie prepares to purchase it.⁶⁸

Here we meet Millie's feminist friend, Sierra. Millie is blocked by Sierra's booming voice from buying her bread, as Sierra scolds her for "supporting the dairy industry."⁶⁹ While standing against the dairy industry may seem generally principled, Sierra's stances on food often link to her racist assumptive care. The reader is informed that Millie met Sierra while she was loudly decrying Fuji apples as excessively sweet. When Sierra expressed intrigue for Sierra's powerful voice, Sierra responded with a generalization about Asian women, comparing them to "those little *beep beep* horns on a Prius."⁷⁰ Sierra's actions throughout the story indicate that she makes generalizations about Asian women and Asian cultures frequently.⁷¹ When Sierra reads Millie as a cis Asian American woman, Sierra's assumptions about Asian women fuels her "advocacy" towards Millie as she insists Millie needs to "sound like you *mean* it!"⁷² Sierra is thus attempting to encourage Millie to find her voice even as she sweepingly casts the voices of Asian women as silent. Millie is skeptical of Sierra's generalizations, thinking that "she knew some pretty obnoxious Asian women," but stays silent nonetheless in the face of Sierra's call to loudness in an attempt to forge a friendship.⁷³

In this context, Sierra's words toward Millie take the form of care by one woman to another, as Sierra seems to genuinely believe that Millie as an Asian American woman could benefit from the (loud) stores of knowledge she has to offer. However, Sierra's care is also already assumptive, shaped by racist assumptions about Asian women as meek. Sierra is thus attempting to include Millie and foster a relationship with her, but only based on assumptions that harken back to controlling images of Asian women as passive.⁷⁴ Sierra's racist assumptions have restricted the horizon of care that she practices towards Millie, short-circuiting any relationship they might have via a racist white feminist perspective about what care towards Asian and Asian American women should look like.

Once Millie comes out to Sierra as a trans woman, Sierra's silencing call to speech transforms into an injunction that

Millie find her subordinate place in relation to Sierra and her view of feminism or else risk an exhibition of male privilege.⁷⁵ This shifts Sierra's conceptualization of Millie from "possible dating material" to "younger brother-sister".⁷⁶ Though Sierra does not see Millie as having "male energy," the omnipresent specter of "male privilege" heightens Sierra's already racist patronizing care towards Millie, informing Sierra's attempts to instruct Millie on "what it meant to be a socially and politically responsible woman."⁷⁷

In this context, though Sierra's care has shifted in response to Millie's trans status, it is also simultaneously based on (1) assuming that Millie has the same experience growing up as white cis men, (2) homogenizing Asian cultures, and (3) generally lacking a sense of awareness about race in conversations with Millie. For example, at one point Sierra "helpfully" informs Millie, "You can't go out alone at night, or walk into any old sports bar without fear anymore, you know!" Sierra's patronizing tone strikes Millie as odd in the context of growing up as an Asian American person, as she remembers being taught to avoid spaces frequented by "large groups of white men,"⁷⁸ indicating Sierra is oblivious of distinctions between Millie's experience and experiences of white men. Additionally, Sierra frequently homogenizes Asian cultures, and once responded to Millie by asserting, "men are men: Chinese men and Japanese men both abused women, like those women in World War II. Oh that was Korea? Whatever. It's all the same oppression."⁷⁹ This homogenizing move is also used to simultaneously exoticize Millie and dismiss her for being trans, showing Sierra's cluelessness about race and trans issues. When Sierra starts telling Millie of a Japanese Zen garden/spa that opened about two hours away, she clarifies that it is a space set up for cis women, saying, "You know, *women* women—but I thought you might appreciate the Zen part, being Asian and all. Very feng shui."⁸⁰ Here Sierra, who decries the possibility of Millie ever taking up space, is more than willing to take up all the space and time she wants through her racist assumptions.

It is in this context that Sierra's criticism of Millie's bread choice stands out as particularly imposing. Sierra's insistence that Millie not exhibit any male privilege and behave as a proper feminist is linked with her commitments to ways of eating food. Criticizing Millie's choice of bread, Sierra brings together racism and non-sequiturs, explaining,

Patriarchy is patriarchy. Do you know that dairy cows end up in McDonald's hamburgers? They're genetically engineered to produce milk. . . . It's worse than what happens to beef cattle. . . . I mean, it's like the Tibetan women—they have nothing but their stories and weaving, but the men in the Chinese government want to take that away from them too.⁸¹

Millie is aware that Sierra's logical connections are baffling, but she holds back to avoid "accusations of male privilege, and, in this case, of possible connections with the Chinese government."⁸² Generally, Sierra has impressed upon Millie that proper feminists are proper dieticians, and that eating properly could help Millie distance herself

from her insurmountable sin of a vaguely defined "male privilege." The narrator provides a window into Millie's reflection on feminist dieting: "She really was sad that she had been born with male privilege, and maybe by being vegan, in some way she could be closer to the woman she wanted to be. A caring woman. A strong woman. A vegan woman."⁸³ Millie is aware of both Sierra's wild connections and her racist assumptions, but Sierra has nonetheless influenced how Millie wants to be seen by others. However, this also entails that Millie must not purchase the bread for her grandmother to prove to Sierra that she is not tainted by male privilege, and Millie leaves the bread behind.⁸⁴ It is thus an imposing white feminism that links Millie picking up bread for her grandmother with a risk of demonization and misgendering through Sierra's racist, simplistic notion of male privilege. Sierra's assumptive care, taking up the mantle of feminism, is ultimately informed by racism and transphobia.

Emi Koyama has argued that contentions between anti-trans feminists and trans feminists have often involved a centering of whiteness that displaces women of color (including trans women of color). In response to controversies over trans inclusion, Koyama asserts that the prioritization of women's oppression above all other forms of oppression in radical feminism leads to the "assumption that the privilege transsexual women are perceived to have (i.e., male privilege) can be viewed as far more dangerous to others than any other privileges (i.e., being white, middle-class, etc.)."⁸⁵ Koyama's analysis thus parallels Sierra's confusing insistence on homogenizing oppression into patriarchy and linking Millie with male privilege, even when these connections lack cogency.

Aoki's story thus depicts a nuanced intersection between racism and transphobia in the context of assumptive care. Though Sierra may be attempting to care for Millie, her racist and transphobic brand of feminism short-circuits all of her attempts, and she is effectively unable to interact with Millie on her own terms. Fittingly, Sierra is totally unaware of why Millie is at the marketplace to begin with, or even that Millie's grandmother is not alive, brushing Millie off and exiting the conversation as abruptly as she entered (albeit by offering a hug).⁸⁶ At this moment Millie refocuses on her grandmother, and it is here that the story transitions from knowledge of assumptive care to a vision of futurebound care as it intersects with Asian American experience, experiences of immigration, and trans experience.

FUTUREBOUND CARE AND SELF-CARE

Though Sierra is the main living non-protagonist in the story, Millie's grandmother persists through death as the key figure to whom Millie turns. Lamenting that she did not pick up the bread upon her return home, Millie's thoughts turn from Sierra's booming voice to thinking about her grandmother.⁸⁷ Having abandoned the sweet bread, yet unconvinced that scrawny parsnips and tomatoes would do justice for her grandmother's birthday dinner, Millie digs around her freezer and finds two pork buns. This is coincidental because Millie had just thought about a time her grandmother reminisced about traveling to the US from Vietnam while she and Millie were in a grocery store searching for steamed pork buns, although the

narrator assures us, “The two pork buns she found there weren’t from providence; they were from being Asian and having a freezer full of ethnic food.”⁸⁸ Now Millie’s food dilemma escalates beyond sweet bread: Should she honor her grandmother by eating meat and go against not only veganism but also vegetarianism? Here, Millie chooses her grandmother over Sierra: “She paused, full of trepidation about backsliding to meat, about oppression and male privilege and Sierra. But, with Grandma’s picture looking right at her, she put the buns in the microwave.”⁸⁹ The shadow of Sierra’s assumptive white feminist care is temporarily dispersed by the more important meeting between magnetron and meat within the microwave.

The food having dispelled Sierra’s hold on space, Millie reflects on an affinity she feels with her grandmother based on a shared lived experience between being an Asian American immigrant woman and being an Asian American trans woman. Eating the steaming pork bun, Millie recalls telling her grandmother she was brave to leave Vietnam for the United States, to which her grandmother responded with laughter. Her grandmother responded, “Brave? No, not brave. You do because you have to. Oh, you give things up, but maybe find new things, too.”⁹⁰ Engaged by steamed meat and memory, Millie feels “a new connection” between herself, her grandmother, and her family, seeing “her own life and identity, for the first time, as an immigrant.”⁹¹ Connecting her own transition to her grandmother’s immigration, Millie reflects on the unexpected aspects of her life as an Asian American trans woman: “She thought about Sierra, and about her other friends, too: gay, queer, trans, Goths, poets—friends who sometimes she just didn’t understand, but who really meant well. *Oh, you give things up but maybe you find new things, too.*”⁹² The time Millie spends with her grandmother thus points her toward a shared trajectory of travel, immigration, and looking toward the future. It is here where Millie offers the toast that forms the title of the story: “*To the New World!*”⁹³

Through these dynamics, Aoki provides knowledge about three aspects of futurebound care. First, futurebound care is present in Millie caring for her grandmother. In suspending Sierra’s demand and preparing the kind of food that her grandmother would prefer, Millie conjures up vivid memories that create an open space for her grandmother to live on and engage with Millie’s world. Second, in opening up this space of care, Millie is also allowing her grandmother to care for her. The memories of her grandmother provide Millie with an open outlook towards her future and the possibilities it might hold for her new world. Third, through these two dynamics, futurebound care also becomes a means of self-care. By maintaining a relationship to her grandmother, Millie is also providing the care she needs to continue onwards in light of her transition. Aoki thus not only highlights intersections between race, trans experience, and immigration in the context of futurebound care, but also elucidates a complex framework in which the care of another is entwined with self-care. Futurebound care, like assumptive care, is inflected by our social position and our dynamic ties with others. This futurebound care even includes a place for Sierra. After offering the toast to the new world sustained by her grandmother, Millie thinks again of Sierra and the

care Millie might offer to her, noting that “Sierra seemed like she needed to talk” and could use some education about trans men.⁹⁴ Millie’s practice of futurebound care thus includes an open space for Sierra too, booming voice and all.

CONCLUSION: TRANS LITERATURE AND TRANS KNOWLEDGE

In this essay, I argued that looking at trans literature can provide knowledge about care ethics in its spatial dimensions. Specifically, I have drawn out insights about assumptive and futurebound care, which close or open another’s space, time, world, and future. To explain this, I looked at Casey Plett’s and Ryka Aoki’s short stories as a source of knowledge about care as it is short-circuited by transphobia and racism, or enabled by giving the other space beyond assumptions and taking action to help them flourish. I believe the rich discussions of care that I have been able to develop from just three stories in trans literature is evidence not only that trans women are complex ethical knowers, but also that trans literature is a helpful companion to trans philosophy and feminist philosophy more broadly. While I hope to revisit the field of trans literature more critically in future essays to discuss the political economy of its production, which itself can manifest in a short-circuited care, I appreciate that the reader cared enough to follow along as I explored some of the dimensions that has made trans literature vital for me over the past few years.

NOTES

1. C. Gilligan, *In a Different Voice: Psychological Theory and Women’s Development*, 18.
2. *Ibid.*, 29.
3. *Ibid.*, 63.
4. J. Keller, “Autonomy, Relationality, and Feminist Ethics,” 153-54.
5. J. Keller and E. Kittay, “Feminist Ethics of Care,” 544.
6. P. H. Collins, *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*, 8.
7. U. Narayan, “Colonialism and Its Others: Considerations on Rights and Care Discourses,” 135.
8. K. Freund et al., “Two Types of Cross-Gender Identity,” 8.
9. M. Frye, *The Politics of Reality: Essays in Feminist Theory*, 86-87, 88.
10. *Ibid.*, 86.
11. *Ibid.*, 85.
12. *Ibid.*, 89, 92, fn. 6.
13. *Ibid.*, 92.
14. Discussing assumptive care and futurebound care in the context of Lugones was suggested to me by Talia Bettcher.
15. M. Lugones, “Playfulness, ‘World’-Travelling, and Loving Perception,” 4.
16. *Ibid.*, 6-7.
17. *Ibid.*, 7.
18. *Ibid.*, 6.
19. *Ibid.*, 8.
20. *Ibid.*
21. *Ibid.*, 9-10.

22. Ibid., 7.
23. Ibid., 5.
24. Ibid., 17; emphasis hers.
25. V. Dalmiya, "Why Should a Knower Care?," 42.
26. M. Hamington, "Care Ethics and Engaging Intersectional Difference through the Body," 88.
27. C. Plett, *A Safe Girl to Love*, 11.
28. Ibid., 11-12.
29. Ibid., 13.
30. Ibid., 14.
31. Ibid., 20.
32. Ibid., 13.
33. A "deadname" refers to a trans person's previous name that is no longer used.
34. Plett, *A Safe Girl to Love*, 18.
35. Ibid.; emphasis hers.
36. Ibid., 29.
37. Ibid., 2, 19.
38. Ibid., 2-3.
39. Ibid., 19.
40. Ibid., 21-23.
41. Ibid., 25.
42. Ibid., 26.
43. Ibid., 27.
44. Ibid.
45. T. Bettcher, "Trans Identities and First-Person Authority," 114.
46. Plett, *A Safe Girl to Love*, 25.
47. Ibid., 207.
48. Ibid., 208.
49. Ibid., 211.
50. Ibid., 201.
51. Ibid., 188.
52. Ibid., 205.
53. Ibid., 191, 193-94.
54. Ibid., 195.
55. Ibid., 201.
56. Ibid., 191.
57. Ibid., 183-84, 189.
58. Ibid., 212.
59. Ibid., 222.
60. Ibid., 223.
61. Ibid.
62. E. Kittay, "The Body as the Place of Care."
63. R. Aoki, "To the New World," 55.
64. Ibid., 54.
65. Ibid., 55.
66. Ibid.
67. Ibid., 55-56.
68. Ibid., 55.
69. Ibid., 57.
70. Ibid., 56.
71. Ibid., 56-57.
72. Ibid., 56.
73. Ibid., 56.
74. Cf. Y. L. Espiritu, *Asian American Women and Men: Labor, Laws, and Love*, 107.
75. Aoki, "To the New World," 56-57.
76. Ibid.
77. Ibid., 57.
78. Ibid.
79. Ibid., 58.
80. Ibid.
81. Ibid., 57.
82. Ibid., 58.
83. Ibid., 57.
84. Ibid., 59.
85. E. Koyama, "Whose Feminism Is It Anyway?," 701.
86. Aoki, "To the New World," 59.
87. Ibid., 60.
88. Ibid.
89. Ibid.
90. Ibid., 61.
91. Ibid., 61-62.
92. Ibid., 62; emphasis hers.
93. Ibid.
94. Ibid., 62.

REFERENCES

Aoki, Ryka. "To the New World." In *The Collection: Short Fiction from the Transgender Vanguard*. Edited by Tom Léger and Riley MacLeod, 53-62. New York: Topside Press, 2012.

Bettcher, Talia. "Trans Identities and First-Person Authority." In *You've Changed: Sex Reassignment and Personal Identity*. Edited by Laurie Shrage, 98-120. New York: Oxford University Press, 2009.

Collins, Patricia Hill. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. 2nd ed. New York: Routledge, 2009.

Dalmiya, Vrinda. "Why Should a Knower Care?" *Hypatia* 17, no. 1 (Winter 2002): 34-52.

Espiritu, Yen Le. *Asian American Women and Men: Labor, Laws, and Love*. 2nd ed. New York: Rowman and Littlefield, 2008.

Freund, Kurt, Betty Steiner, and Samuel Chan. "Two Types of Cross-Gender Identity." *Archives of Sexual Behavior* 11, no. 1 (1982): 49-63.

Frye, Marilyn. *The Politics of Reality: Essays in Feminist Theory*. Trumansburg, NY: Crossing Press, 1983.

Gilligan, Carol. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press, 1982.

Hamington, Maurice. "Care Ethics and Engaging Intersectional Difference through the Body." *Critical Philosophy of Race* 3, no. 1 (2015): 79-100.

Keller, Jean. "Autonomy, Relationality, and Feminist Ethics." *Hypatia* 12, no. 2 (Spring 1997): 152-64.

Keller, Jean, and Eva Kittay. "Feminist Ethics of Care." In *The Routledge Companion to Feminist Philosophy*. Edited by Ann Garry, Serene J. Khader, and Alison Stone, 540-55. New York: Routledge, 2017.

Kittay, Eva. "The Body as the Place of Care." Presentation for the Association for Feminist Ethics and Social Theory (FEAST) Conference. Tucson, AZ. October 20, 2013.

Koyama, Emi. "Whose Feminism Is It Anyway?" In *The Transgender Studies Reader*. Edited by Susan Stryker and Stephen Whittle, 698–705. New York: Routledge, 2006.

Lugones, Maria. "Playfulness, 'World'-Travelling, and Loving Perception." *Hypatia* 2, no. 2 (Summer 1987): 3–19.

Narayan, Uma. "Colonialism and Its Others: Considerations on Rights and Care Discourses." *Hypatia* 10, no. 2 (Spring 1995): 133–40.

Plett, Casey. *A Safe Girl to Love*. New York: Topside Press, 2014.

Participation in Sport Is a Human Right, Even for Trans Women

Rachel McKinnon
COLLEGE OF CHARLESTON

I never wanted to be an "activist" for athlete rights. Truly, I just wanted to race my bike and see how far I could advance in the sport I love. So far, this has led to a UCI Masters Track Cycling World Championship in 2018. I still have Olympic aspirations. It's certainly hard enough to achieve such lofty goals, but it's a lot harder when people are actively doing everything they can to ban you from competing, even when you're following every rule under the strictest of scrutiny. They still call me a "cheater" just for existing.

There are a lot of arguments being bandied about on whether it's fair for trans women to compete in women's sport. What is often lost in the shuffle is that there are trans men athletes and non-binary athletes. "Where are the successful trans men athletes?" I've read that more times than I can count. But the most commercially successful trans athletes have been trans men: Chris Mosier had a NIKE television ad campaign run during the 2016 Rio Olympics; Patricio Manuel has had a feature in *Rolling Stone* magazine.

Mack Beggs, a trans male wrestler who rose to fame for being forced to wrestle in the high school girls' category, is all too often used as an example of what's wrong with trans women in sport. He's now happily wrestling on the men's team in college. Amazingly, people irrationally opposed to trans women in sport think that he's a trans woman. Trans men are often celebrated, but then forgotten. They are vanished from the social imagination, since they don't serve the dominant narrative that only focuses on trans women in sport. In some cases, like with Beggs, they're used as pawns against trans women, in contortionist twists of what can at best be loosely called "logic."

And while trans men at least receive some media coverage and accolades for their "bravery"—which is itself an entirely other problem, rooted in misogyny and sexism—non-binary athletes are left out completely. Sport itself is often not structured to include non-binary people. If one isn't easily categorized into the binary man/woman, where should one compete? Many rules require an athlete to choose one, and then can't change their mind for four years.

And this binaristic structure of sport comes from on high, from the International Olympic Committee (IOC) and the arms-length arbiter of sport disputes, the Court of

Arbitration for Sport (CAS). In a landmark 2015 CAS decision about intersex female runner Dutee Chand, CAS's decision reads in part, "there are only two categories of competition: male and female. These categories are together intended to cover *all* athletes who wish to participate in competitive athletics." Sport's institutional structure is such that all Olympic-eligible sports must at least abide by IOC policy. In this case, athletes wishing to compete in the Olympics may only do so in either the "male" or "female" categories.

Common "understanding" is that there is at least some rigid distinction between sex and gender. So one might question why I'm slipping between using man/woman or boy/girl, which are often understood as "gender" terms, and male/female, which are often understood as "sex" terms. The reason is twofold. First, sport does not make any distinction between "sex" and "gender": they are used interchangeably. In the aforementioned CAS case, the paragraph following the aforementioned quote reads in part, "A rule that prevents some women from competing at all . . . is antithetical to the fundamental principle of Olympism that 'Every individual must have the possibility of practising sport, without discrimination of any kind'."

Second, governments do not make any distinction between "sex" and "gender," at least not typically. While trans people are not universally able to secure legal recognition of their (transitioned) sex, they are able in many countries. The USA, Canada, UK, Australia, Germany, and a long list of other countries have practices of allowing trans people to change the sex designation on their legal documents, including their birth certificates. In Canada, for example, a trans woman can be *legally recognized* as "female." There is no legal distinction between being "female" and "a woman." Of course, different governments require different things of trans people to acquire this legal recognition, and there can be differences at the federal and state/provincial levels.

There are many examples. Consider the language of the UK 2004 Gender Recognition Act:

Where a full gender recognition certificate is issued to a person, the person's gender becomes for all purposes the acquired gender (so that, if the acquired gender is the male gender, the person's sex becomes that of a man and, if it is the female gender, the person's sex becomes that of a woman).

Note the use of "male gender" and "the person's sex becomes that of a man." The "gender" and "sex" terms are used in combination and interchangeably. Legal identification documents only list "Male" or "Female" or their abbreviations "M" or "F." One's passport doesn't say "Woman" or "Man." The idea that either sport or "female-only" social spaces such as bathrooms are segregated on the basis of sex and not gender, leaning on a rigid distinction between sex and gender, is simply not the case.

It's simply the case that *sport* and *governments* do not make a rigid distinction between sex and gender. Why does this matter, though? Some people who oppose trans women's

inclusion in sport are increasingly, begrudgingly willing to grant that trans women are *women* in terms of gender, but remain steadfast in their denial that trans women are not *female* in terms of sex. The idea is that we segregate sport on the basis of sex and so, begrudgingly, while trans women may be women, they are not *female* and thereby shouldn't be permitted into female-only sports categories.

My point here is to call attention to the lack of any such distinction in organized sport—including at the Olympic and Commonwealth levels—or in governments. So those who seek to argue for trans women's exclusion on the grounds that they (wrongly) think that sport is segregated on the basis of sex are, well, totally wrong.

But it gets worse for them. Prior to 1999, and in practice before the Atlanta 1996 Olympic Games, the IOC had various "sex verification" policies to determine whether an athlete was male or female, irrespective of their legal sex. However, this only ever applied to those seeking to compete in the women's category. We never scrutinize men and those competing in the men's category. Women's sport is the "protected" category. And in the early days, an athlete thought to be too masculine—which was differentially applied to women of color, since norms of femininity were set by white women's femininity—would be required to appear before a panel who would inspect her genitals. Anything other than a (white) normative vulva was deemed sufficient evidence that the athlete was not "really" a woman, and so would be excluded from competition.

Eventually, genital inspection was deemed to be insufficient, so the IOC introduced the practice of chromosomal testing. Specifically, they applied the Barr Body test, which (by "failing") merely tests for the existence of a Y sex chromosome. The presence of a Y chromosome constituted a "failure" of the test, and the female athlete would be deemed ineligible and barred from competition (with women). Over time, as scientific understanding of the prevalence of intersex conditions increased, test cases increased. Humans are not sexually dimorphic, meaning that people fall neatly into "XX" (female) and "XY" (male).

Humans are bimodally distributed around XX and XY configurations, but there are many other possibilities including XXY, XYY, XO (a null chromosome), and others. Moreover, some XX people have Congenital Adrenal Hyperplasia (CAH), which causes much higher production of testosterone, typically leading the person to develop phenotypically male. Conversely, some XY people have Complete Androgen Insensitivity Syndrome (CAIS) where while their body produces "typical" male levels of testosterone, their testosterone receptors are *insensitive* to the hormone, and this typically leads the person to develop phenotypically female.

Recognizing this, the IOC last engaged in chromosomal testing of athletes in the Atlanta 1996 Olympics, and formally banned the practice in 1999. Since then, and noted explicitly in the 2015 CAS decision, all sex verification policies are banned. Sport organizations are no longer in the business of determining whether an athlete is male or female. Instead, the CAS panel notes, "The distinction

between male and female is a matter of legal recognition" whereby "whether a person is female is a matter of law."

All Olympic-eligible sports are thereby required to respect an athlete's legally recognized sex. If a trans woman's identification documents say "female," *then she is really female*. Those who oppose trans women's inclusion in sport typically say that trans women are not "really" female, but they have no legal standing. They can scream "but biology" all they like: it doesn't change the facts. Sport takes an athlete's legally recognized sex. In many jurisdictions, trans women are legally female.

To reiterate, the first line of "argument" from those who oppose trans women's inclusion in women's sport is that sport isn't about gender, it's about sex. And while they may begrudgingly grant that a trans woman is a *woman*, they deny that she is *female*. But sport and governments make no such sex/gender distinction.

The second line of argument is that trans women are not "really" female, and thereby should be excluded from "female" sport. But sport—from the IOC down and all sports under the auspices of CAS—must respect an athlete's legally recognized sex: they may not have their own sex verification policies. And since trans women can be legally recognized as female, they are *really female for the purposes of sport*.

In my reading of things over the years, these are the primary arguments of those opposing trans women's inclusion in women's sport. And neither argument holds any water. More recently, some have become at least a little responsive to their losing this battle and have transitioned into what is now the most prevalent argument used against trans women's inclusion in sport: alleged performance advantages.

But I want to explain why this is largely irrelevant. Trans athletes' rights to compete are not contingent on showing that there isn't a competitive advantage. Additionally, because proving a negative is literally impossible, people who oppose trans women's inclusion can forever demand "more study" and the need for "more evidence" before they'll relent. But that day will never come, for they'll continue to manufacture potential sources of evidence even in the absence of any scientific evidence suggesting such a thing exists. My favorite so far is the claim that trans women have "muscle memory," by which these people mean that trans women's muscles "remember" pre-transition endogenous testosterone. This isn't a thing. They're making this up.

What does matter is the human rights framework. Some balk at the idea that sport is a human right. Others frame this "debate" about trans women in sport as pitting trans women's rights against (cis) women's rights in a kind of conflict of rights. But I'm here to tell you that there is no conflict of rights.

Sport is a human right. After a preamble, the IOC's Olympic Charter lists seven Fundamental Principles of Olympism. The fourth begins with, "The practice of sport is a human right."

That's the first full sentence. And they mean competitive sport: the IOC is only concerned with competitive sport. There is no right to *win* in sport; there is no right to *make a team selection* in sport. The right is to participate and try to win, to try to make the team, to strive for your best in the Olympic spirit, of mutual understanding, and of fair play.

(Cis) women do not have a right to exclude women they don't like or don't feel comfortable with. Sport and society has a long history of excluding women of color, often trading on these same claims of alleged competitive advantage or not feeling "safe." But these are not rights. Thus, extending the right for (trans) women to compete in sport with other women is *not* in conflict with other rights. We can't make up rights and claim a conflict. Rights are socially constructed but institutionally enshrined, often in law or policy. These alleged rights that people claim are in conflict with trans women's inclusion in sport don't exist in any institutional or legal sense.

Suppose that you're now convinced that participation in (competitive) sport is a human right. And suppose that you're now convinced that trans women are legally female, and thereby belong in women's sport. You might still object to trans women's inclusion on the basis of some notion of "fairness" and alleged unfair competitive advantages that you think trans women enjoy.

This is why the human rights framing is what controls the issue, not whatever scientific evidence we may want to argue over, allowing us to distract from the core issue. The IOC is an *international* organization, as is CAS. When they speak of "human right(s)," this puts us into the realm of international human rights law and principles. Two frameworks are often invoked: the United Nations [Universal Declaration of Human Rights](#) and the European Court of Human Rights [European Convention on Human Rights](#). CAS and the IOC are situated in Switzerland, which falls under both.

Both frameworks require the elimination of discrimination against women on the basis of sex (or gender; remember, these are interchangeable). This isn't to say that discrimination is never justifiable. "Discrimination" can be a neutral term, simply referring to the distinguishing between different groups or categories. In common use, "discrimination" refers only to the unjustified, unethical, or illegal forms of discrimination. But international human rights frameworks include provisions for when we can justify what is otherwise discriminatory.

There's a four-fold test. First, the policy must be in service of a worthy social goal. We have prisons and override the right to freedom of movement, partly on the grounds that doing so is in service of the worthy social goal of "promoting public safety." We can argue about whether this is *effective*, but the first test is merely to ensure that policies are in service of a worthy social goal. In sport, policies are in service of the worthy social goal of *fairness*.

Second, the policy must be *necessary* for the promotion of the worthy social goal. If we can achieve the worthy social goal without infringing upon human rights, then we must.

We can only potentially be justified in overriding human rights if doing so is necessary for promoting a worthy social goal. Arguably, the death penalty is not necessary for promoting public safety, and thus arguably fails this second test.

For our purposes, one issue is whether excluding trans women from women's sport is *necessary* for promoting fairness in competition. More on that below.

Third, the policy must be *effective* at promoting the worthy social goal. Even if a discriminatory policy might be judged necessary for promoting a worthy social goal, if it isn't effective at doing so, then it fails to be justified. A primary justification for the death penalty is to deter other crime. Arguably, evidence suggests that the death penalty is not effective at such deterrence. And since there are other methods capable of preventing someone from reoffending, the death penalty is neither necessary nor effective, and so is not justified.

For our purposes, another issue is whether excluding trans women from women's sport is *effective* at promoting fairness in competition. More on that below.

Finally, the benefit from promoting the worthy social goal must be *proportional* to the harm caused to the group or individuals discriminated against by the policy. Generally, policies that discriminate against already vulnerable or stigmatized social groups, even if they are necessary and effective in service of a worthy social goal, will fail to be sufficiently proportional. Appeal to the small size of a group will not suffice, either: the proportionality test is not a utilitarian calculus whereby a large group can benefit greatly at the expense of a few.

For our purposes, another issue is whether the harm to trans women caused by excluding them from women's sport is *proportional* to any proposed benefit to (cis) women. More on that below.

In an unprecedented move, the UN Human Rights Council released a statement calling on the UN High Commissioner for Human Rights to look into discriminatory policies in sport, including restrictions on endogenous testosterone in women, calling out the International Association of Athletics Federations (IAAF) by name. The statement reads in part:

Expresses concern that discriminatory regulations, rules and practices that may require women and girl athletes with differences of sex development, androgen sensitivity and levels of testosterone to medically reduce their blood testosterone levels contravene international human rights norms and standards, including the right to equality and nondiscrimination, the right to the highest attainable standard of physical and mental health, the right to sexual and reproductive health, the right to work and to the enjoyment of just and favorable conditions of work, the right to privacy, the right to freedom from torture and other cruel, inhuman or degrading treatment and harmful

practices, and full respect for the dignity, bodily integrity and bodily autonomy of the person.

The statement also explicitly refers to the aforementioned international human rights framework:

Noting with concern also that the eligibility regulations for the female classification published by the International Association of Athletics Federations that came into effect on 1 November 2018 are not compatible with international human rights norms and standards, including the rights of women with differences of sex development, and concerned at the absence of legitimate and justifiable evidence for the regulations to the extent that they may not be reasonable and objective, and that there is no clear relationship of proportionality between the aim of the regulations and the proposed measures and their impact.

What's most crucial about sport being a human right is that the *default* position is inclusion. The default is not "Exclude trans women until we have more evidence about there not being a competitive advantage." Rather, the default must be "Include trans women unless we have sufficient evidence to justify discrimination in an international human rights framework.

This is why the human rights framework controls this "debate." The practice of sport is a human right (IOC Charter), sex is a matter of legal recognition (CAS), and trans women can be legally recognized as female. Therefore, trans women have a human right to participate in competitive sport as women.

Trans women don't have to justify our inclusion. The burden of argument is entirely on those who seek to exclude us. And, as I'll briefly prove, that burden has not yet been met, and is unlikely ever to be met.¹

In order for excluding a group of people based on an alleged competitive advantage, as noted just now, such a policy would need to be in service of a worthy social goal, necessary and effective at promoting that goal, and the benefit to society is proportional to the harm caused by the policy. Trans-exclusionary policies fail on every measure except that the policies are at least plausibly in service of the worthy social goal of "fairness in competition."

How much advantage do trans women have in sport? I'm here to say *it doesn't matter*. It really doesn't. Let's assume for the sake of argument that there is an inherently biological cause of the gender performance gap where we see an approximately 10–12 percent difference between peak (cis) men's performances and peak (cis) women's performances. Let's just assume this, even though I think it's false. Let's also assume for the sake of argument that trans women are physiologically coextensive with cis men. Let's also just assume this, even though I think it's false.

My point here is *even if we assume this*, and we assume that trans women are physiologically the same as cis men, this is not enough to justify the exclusion of trans women from women's sport.

Some will recoil: but isn't this why we have men's and women's sport? Don't we sex-segregate sport *because* men are stronger and faster? No, I don't think so. I don't think the history of sex segregation and banning women from sport (back to ancient Greek Olympic Games) is because men are stronger: women were banned outright. Prior to the 1984 Olympic Games, there was no women's marathon event. Prior to the 1972 Olympic Games, there was no women's event in the 1500m or any event longer than 800m. Women weren't allowed into the Boston Marathon when, in 1967, Kathrine Switzer broke the rules to participate.

In the 1992 Olympics, the sports shooting competition was not sex-segregated (just like equestrian still is). A Chinese woman, Zhang Shan, won the gold medal. In the following 1996 Olympics, the IOC sex-segregated the event and *didn't offer a women's category*. The defending gold medalist was thus banned from competition. The current women's chess world champion was not permitted to compete for the men's world championship. Billiards, darts, bowling, and a long list of other sports are sex-segregated with no plausible physiological explanation.

None of these policies are or were because there's an alleged fundamental biological advantage that men have over women. This isn't *why*, historically, sport is sex-segregated. Sport both reflects and leads social attitudes. Societies the world over were and continue to be sex-segregated, relegating women to second-class status. Post hoc rationalizations abound, but simple sexism explains why we sex-segregate sport. We should have no illusion otherwise.

However, this is all irrelevant. My point, again, is that *even if we grant* that (cis) men are inherently stronger than (cis) women, *and we grant* that trans women are physiologically co-extensive with cis men, excluding trans women is not justified in an international human rights framework.

Why is this? Because claims like "men are stronger than women" are, strictly speaking, false. There are many women who are stronger than many men. These claims, instead, are either: the *average* man is stronger (taller/faster/etc.) than the *average* woman; or the *best* man is stronger (taller/faster/etc.) than the *best* woman. At present, these latter claims are true. But both elide the massive ranges within men and women. The shortest, weakest, slowest man is often the same as the shortest, weakest, and slowest woman.

Elite women athletes are considerably stronger than the *average* cis man, and certainly the *average* trans woman. The average height of the 2016 Rio Olympics women's high jump podium was 6'1.7". The tenth place woman in the final is 5'5"; the gold medalist is 6'3.6" and was the tallest in the competition. The global average height for men is around 5'9". Moreover, height is not uniformly distributed around the world. The average Dutch woman is 5'6.5", whereas the average Indonesian woman is 4'10".

We permit *huge* differences in natural physical traits within sport and call that "fair." So *even if* we say that the average trans woman is 4" taller than the average cis

woman, we already permit much larger height differences within women's sport and call it fair. This is true for any natural physical trait one selects, including endogenous testosterone. And while there's no evidence of a relationship between endogenous (internal, natural) testosterone and performance, my point is that even *if there were*, it's insufficient to justify excluding trans women.

We demonstrably permit competitive advantages within women's sport that are far greater than 10–12 percent. "Fairness" does not require that no athlete have a competitive advantage over her competitor. In fact, that's the entire purpose of training, coaching, equipment, nutrition, and so on. And fairness does not require us to exclude those who have considerable performance advantages.

So where does this leave us? The "fairness" claim is that trans women are like cis men, and cis men have a 10–12 percent performance advantage over women. If we simply assume that this is true, it's not enough to justify excluding trans women. Such a policy, while in service of the worthy social goal of fairness in competition, is not necessary for fairness: we permit even larger advantages within women's sport. It also thereby can't be effective. It's failed two of the *required* tests to justify a discriminatory policy under an international human rights framework.

And, finally, trans women are a heavily stigmatized and marginalized group, particularly trans women of color, and particularly trans women of color from the global south. And since the benefits to society of discriminatory policies that further harm an already marginalized, stigmatized group tend not to be proportional, excluding trans women also fails the proportionality test.

Excluding trans women is not justified under an international human rights framework, under which the participation in sport as a human right is subsumed. And given the comparably huge competitive advantages within (cis) women's sport we permit under fairness, I think that it's unlikely (to impossible) that any new evidence will ever change this. The default is inclusion, not exclusion. We don't need "more study" or "more evidence" in order to decide to grant what is everyone's human right: the right to participate in competitive sport.

So the scientific evidence is irrelevant. No competitive advantage that could ever be found to be attributed to trans women will be enough for an exclusionary policy to be both necessary and effective at promoting fairness in sport, given the massive competitive advantages we already permit in women's sport. The human rights argument is the one that controls this "debate."

And there *is no debate*.

NOTES

1. For an in-depth treatment of this argument, see Rachel McKinnon and Aryn Conrad, "Including Trans Women Athletes in Competitive Sport: Analyzing the Science, Law, and Principles and Policies of Fairness in Competition," forthcoming. Available at <https://bit.ly/2Hbeavr>.

Conceiving Differently within the Ethics of Assisted Reproduction

Alice MacLachlan

YORK UNIVERSITY

"Some parents who choose to conceive using gametes from others do so outside a clinic (e.g. a couple may use sperm provided by a friend). In such arrangements, intending parent/s may be exposing themselves to health risks and uncertainties regarding legal parenting rights and responsibilities. Little is known of these families and therefore this chapter will focus on those who conceive in a clinic."¹

– Lucy Blake, Martin Richards, and Susan Golombok, "The Families of Assisted Reproduction and Adoption"

Over the last few years, I've had various professional reasons to dive deeper into academic debates on the ethics of alternative family-making and assisted reproduction. I always do so with intellectual curiosity, significant personal investment, and no small degree of trepidation: it's hard not to search anxiously for myself across the pages, wondering what kind of verdict will be given about *my* family and my choices.

Almost always, the short answer is—pretty much none. As the quotation above illustrates, my family is both conspicuously absent and also fairly consistently referenced *precisely in order to flag our irrelevance* in discussions of assisted reproduction and sperm donation—we are alluded to as a possibility, cautioned as a risk, then dismissed as unruly data. Sometimes, given what ethicists have to say about the families they *do* discuss, this is a relief; at other times it leaves me with a pang of invisibility. But, over time, I've started to notice that academic silence around known gamete donation² doesn't just have an emotional effect (on me) but, more generally, a rhetorical one—I believe conversations about the ethics of family-making are shaped by the forms of family-making that are taken into account. And this has led me to wonder: what gets left out along with my family? And how might these conversations change if we were included?

My wife, Amy, and I are mothers of two daughters, Emmylou and Martha, aged six years and eighteen months, respectively. I carried both children; they were conceived with my eggs and with sperm from a friend of ours, outside of a clinic and without the use of other fertility treatments. My wife legally adopted Emmylou through a step-parent adoption in 2013. Then, in 2016, we were part of a successful legal battle to change family law in Ontario so that she could be recognized on Martha's birth certificate in 2017 without subsequent adoption.³ While the donor has legally waived all parental rights and responsibilities (first through a legal contract pre-conception, then the step-parent adoption, and finally under the new family law), he and his family are dear friends of ours; they care about our

kids just as we care about theirs. In many ways, I consider them family. They are part of the extended network of kin—whether through blood, choice, or happy accident—that my kids and their mothers are astoundingly lucky to have surrounding us.

Much as I might like to think of my family as unique, our story isn't. Lots of queer parents turn to friends and kin when building their family—we belong to more than one fairly large Facebook group organized around families with known gamete donors, and stories like ours are mundane enough to have featured as a plot point in various TV shows and movies. So why don't we figure in these academic debates? One easy answer is that the relevant ethical issues, for the most part, don't apply—or don't apply as directly—to our families.⁴ Philosophical and bioethical discussions of assisted reproduction tend to focus on four issues: (i) the various rights and responsibilities of all involved insofar as these illuminate the necessary and sufficient conditions for moral parenthood; (ii) the growing role of technology in human reproduction; (iii) the value and significance of genetic ties, and of knowledge of/access to one's genetic ancestors; and (iv) the ethics of secrecy and deception around someone's origins.

These are undoubtedly significant issues—but only the first touches directly on my family's experience. My wife and I came to parenthood together but through slightly different routes, meaning that I am sometimes described as the genetic-gestational mother and she, the social mother. This asymmetry distinguishes us from bionormative parents (each of whom has a genetic relationship to their child) and adoptive parents (neither of whom does) and even step co-parents; they enter their child's life at different times while we became parents together. Our asymmetrical, wholly shared parenthood provides an interesting case for debates about what fixes *moral* parenthood: that is, the particular rights and responsibilities towards a child we attribute to and only to their parents. At least some, though not all, variations of a leading (causal) theory of parenthood would shift my wife's standing to our sperm donor, for instance.⁵ From my perspective, this a pretty damning knock against such theories—but that is a question for another day. Moreover, since neither of our parental statuses is contested (by the donor, the state, or our children—so long as we don't count things said in the heat of piano practice), the ethical question feels relatively inert, in our case.

The other three issues are even less salient to us. As far as (ii), technological intervention, is concerned, we set out to conceive with the assistance of a \$.99 drugstore syringe (an artifact no more advanced than a broken condom). I've heard of folks using a known donor who employ nothing more than a spoon or the (now infamous) turkey baster. And as for (iii), the significance of genetic ties, the most impassioned arguments *against* gamete donation in the realm of philosophy are made by J. David Velleman, who focuses on the importance of *knowing* and *relating* to one's genetic relatives for human development and flourishing; the assumption is that donor-conceived people simply can't have this. Without the chance to observe (genetic) family resemblances and hear (genetic) family stories,

Velleman's argument goes, the donor-conceived individual grows up not knowing—or not entirely knowing—who *they* really are and, further, lacks the resources to make proper sense of what they do learn about themselves.⁶ While I am not persuaded by Velleman's arguments, it is also true that they don't apply as forcefully to my children, who do know their donor, his parents, and his children (their genetic half-siblings) and who have had and will have opportunities—should they desire them—to hunt for resemblances and affinities, points of difference and of overlap, within the context of ongoing caring relationships.

Velleman himself would likely say his arguments *do* apply to my family, as he notes that “the reasons for concluding that children should have access to information about their biological parents support the stronger conclusion that, all things being equal, children should be raised by their biological parents.”⁷ But he presents his case as an exhaustive disjunction—access to information or full parenthood—that ignores my family's existence. There is a vast difference between “having access to information” about a donor (e.g. health history, a photo, list of interests and aptitudes) and growing up with them as a well-known, beloved adult and family friend, within a personal relationship that has developed over time.⁸ It's not clear why having a genetic parent be one's legal guardian and default caregiver should be required to cultivate the kind of relationship Velleman describes as necessary for personal development—and Velleman never provides an argument for why it does. Again, this is not to endorse his overall argument *against* gamete donation; merely to note that its rhetorical power depends on the erasure of my family's reality.

And finally, there is much to be said about (iv): the ethics of secrecy around gamete donation, and the shock of discovering that one's social parent is not one's genetic parent after assuming, or being explicitly told, that they were. Certainly, secrets can be painful and deception even more so—even when they are an understandable and perhaps necessary choice for particular families in particular circumstances.⁹ I have spent some time on the We Are Donor-Conceived website (www.wearedonorconceived.com) and have been profoundly moved by the very heartbreaking testimony provided there.¹⁰ But, again, these moral quandaries are less relevant to queer families who use donor conception—whether with a known, open, or anonymous donor—where secrecy is less common. After all, even if same-sex parents use anonymous donor-conception, it's generally known that *some* donor, surrogate, or bio-parent was involved, and the default approach between queer parents and their children is *more* rather than less openness about conception.¹¹ In many cases, like ours, ‘queerspawn’ kids grow up knowing the mechanics and the meaning of how they came to be—as soon as or even before they can fully grasp both—told as a story of love and collaboration.¹²

But here's where I want to challenge the rhetorical framework at play. Instead of justifying our invisibility by pointing out our irrelevance, we could ask: How might academic understandings of what is or isn't ethically significant when it comes to non-traditional family-making *change*, if we took into account the wider variety of ways

that non-traditional families are made? And so, I find myself asking: if I were to develop the ethics of collaborative reproduction by starting with my own family's story, for example, what might emerge or transform about this field?

Put simply: I think we complicate things.

Academic research tends to characterize non-normative family-making in terms of *absence* or *lack*: lack of the father or the mother or appropriate reproductive organs or authentically genetic ancestors or significant genetic information. Bionormative families are taken to have something that other families lack—a particular parent, or set of histories, or genetic web—and implicitly become the model for completion, even wholeness. And, then, it's still easier for "lacking" to slide subtly into "less than."

Yet our experience has been one of abundance, not absence. Our children have their own family in us and our families of origin, and yet another family that is not theirs but with whom they have a special genetic connection, and the chance to explore what that means to them and to their genetic half-siblings, who also have a stake in its significance. In fact, my children also have something that children of bionormative families *lack*; the opportunity to experience and learn about themselves through genetic family resemblances and resonances *and* those that are not the dictates of coded genes, but which emerge solely out of and reflect the work of an ongoing loving relationship. When our eldest daughter (born and raised in Toronto) was young, about half the words that came out of her mouth were shaped by my wife and my in-laws' Newfoundland accent. She carries traces of it to this day, along with her love of superheroes, reading and drawing comics, DIY carpentry, and jokes repeated ad nauseum—all of which she gets from my wife. She loves to categorize and annotate our family's samenesses and differences; she and I have long hair and wear dresses, but the baby and my wife do not. Her sister and I like kicking a soccer ball around while she and her mama would rather draw. All three of them are right-handed and have blue eyes; I'm a brown-eyed lefty. The girls like Baby Shark on repeat, while Amy and I have our limits. Someday, as she grows up, Emmylou will start to map her own categories onto their wider social meanings: masculine vs. feminine, athlete vs. artist, adult vs. child. My hope is, our abundant family configuration will give her the puzzle pieces to attach her own meanings to family vs. genes, and to resist messages like Velleman's, that tell her whatever she sees of herself in me and my family of origin—or indeed, in our donor and his—must *necessarily* be more authentic and meaningful for her identity than the panoply she shares with Amy, with her Nana, or with her Grandpa and Nanny Sue.

In fact, children from exactly the kinds of families that the researchers in the epigraph avoid may be in a uniquely well-placed position to better understand the significance of genealogy vs. family than most of us. My children will experience what it is like to be loved by a social-genetic parent, a purely social parent, and—in another capacity—to be cared about by a genetic parent. Children in other overlooked *abundant* configurations (e.g., three-parent families,¹³ poly families, sets of partners co-parenting

together, and beyond) will have still richer comparative understandings of what various caregivers and histories bring to their own story and identity as they grow up. Moreover, this understanding is less likely to be shaped by the cultural message of *absence*—with the implication that the goods of social parenting are there to compensate for the loss of biological, genetic goods.¹⁴ In fact, I suspect the unique and distinct goods of purely social parents will only become visible when they are not framed as compensation for a prior or deeper loss.

I don't raise these cases to add a top layer to the hierarchy of bionormativity, as it were. My intention is not to argue that more parent-permutations necessarily means better parenting, or that some families "win" at family-making through sheer numbers alone. But we do put the lie to the idea that bionormative families have something other families *lack*—and in doing so, draw attention to the rhetoric of bionormativity that frames itself as wholeness vs. incompleteness. What I am calling "abundant" non-normative family configurations are not necessarily richer or more conducive to human flourishing than bio-mom-and-dad families, any more than bio-mom-and-dad families are necessarily richer or more conducive to human flourishing than single-parent families, blended families, non-parent caregiver families, or those families created through anonymous gamete donation, surrogacy, or adoption. Sadly, both social and biological parents can fail our children. Instead, shifting the picture to include a variety of configurations—smaller and larger, original and blended, social and/or biological—destabilizes the idea of that there is a single "whole" family, which alone has the potential to provide exactly those goods that growing humans need. Indeed, some of the relevant goods may not even be identifiable within current framings. Once we refuse the rhetorical move of starting from a disjunction (whole vs. lacking, default vs. compensatory, bionormative vs. less-than) we open ourselves up to a possible plurality of family goods and good families—and thus, put ourselves in a better position to do the finer-grained work of understanding what goes wrong and why, when families fail or harm.

In other words, I think including families like ours offers up different paradigms for understanding and answering the central ethical questions of family-making. It may also draw attention to new, undertheorized ethical themes. When I reflect back on our experiences, the salient ethical issues were not secrecy or technology, who was or wasn't a true parent, but the murkier and yet more familiar ethics of collaboration and dependency, and the relationship between risk and trust within collaborative endeavors: whether it was safer to house risk in the institutional protections of state and clinic or privately through interpersonal goodwill; how the asymmetry of our collaborative enterprise reacted to hurdles and disappointment, even miscarriage; how we navigated our differing attitudes to privacy and disclosure; how to think and talk about our family in relation to others who are more or less closer to the bio-norm, in ways that celebrate our configuration without disparaging others (something that has been a struggle, even in the course of this essay).

No one can reproduce entirely on their own, unless they're an amoeba, but our collaboration was larger than many. Beyond the donor were his wife and their children, who regularly hosted us as we were trying to conceive—and then there were lawyer friends who reviewed the contract and adoption for us, queer friends who offered detailed insemination advice, family members who gave us their frequent flyer miles, visiting grandmothers who offered sweet and surprisingly non-awkward encouragement, friends whom we called on to cat-sit, colleagues who covered classes, the family doctor who taught me to track my fertility cycles, and so on, and so on. In our case, it didn't just take a village to raise a child; it took a village to make one. Emmylou and Martha were conceived with a thousand tiny acts of kindness, and several large and significant ones. Paying attention to stories like ours centres the reality of *collaboration*—not technology or secrecy or ancestry—at the heart of what it is to make a family.

ACKNOWLEDGMENTS

For a short piece, this has received a lot of assistance. I'm grateful to Carolyn McLeod and Françoise Baylis, whose work in this area first introduced me to the debate; to Olivia Schuman, Daniel Groll, and Michael Montess, whose conversation and writing have shaped my thinking on the topic; to Deborah Finding, Amy Noseworthy, and Susanne Sreedhar for their comments on earlier drafts, and to Grayson Hunt for giving this essay a home. Finally, I'm grateful to Amy, Emmylou, and Martha for letting me do philosophy about them.

NOTES

1. Found in *Family-Making: New Ethical Challenges*, ed. Françoise Baylis and Carolyn McLeod (Oxford, 2014), 69.
2. Some quick terminology: known donors are personal family or friends; open-identity donors are strangers, connected through a clinic or sperm bank, who agree to share personal information.
3. Alice MacLachlan and Amy Noseworthy, "By Ignoring Parental Rights, Ontario Puts Our Daughter's Welfare at Risk," *The Globe and Mail*, June 2 2016. <https://www.theglobeandmail.com/opinion/by-ignoring-parental-rights-ontario-puts-our-daughters-welfare-at-risk/article30244298/> (accessed May 27, 2019).
4. And of course, when it comes to descriptive empirical research, we're less easy to track—since our family-making doesn't necessarily require a fertility clinic, we can't always be found on client lists or through donor registries. But my focus here is normative debates about parenthood and assisted reproduction in applied ethics, which don't rely as heavily on empirical data.
5. For a compelling argument for why causal theories of parenthood do not need to draw this conclusion, see Lindsey Porter, "Why and How to Prefer a Causal Theory of Parenthood," *Journal of Social Philosophy* 45, no. 2 (2014): 182–202.
6. J. David Velleman, "Family History," *Philosophical Papers* 34, no. 3 (2005): 357–78.
7. *Ibid.*, 362 fn.
8. Indeed, there are multiple degrees of knowledge between access to a donor-profile and a lifelong personal connection to one's donor that may also be more or less conducive to personal development. For a fascinating discussion of the epistemology of various donor arrangements and the sense of "know" in question when philosophers discuss the need to know, see Olivia Schuman, "What Kind of Interests Does the 'Right to Know' Protect?" (unpublished).
9. The activist in me can't help but note that a culture which prizes bionormativity and implicitly downgrades other families to second-best may bear some responsibility for why parents choose to keep donation secret.
10. While the voices of donor-conceived individuals online largely express dissatisfaction at anonymous donor-conception, their views are not monolithic. This testimony is taken from another

website, anonymous.org: "I have never met my donor "father," and I have no desire to do so. I do not see this lack of contact with my biological father as something missing in my life, and I have no hurt at the fact of my creation. What does cause me hurt, however, is the idea, constantly repeated by small numbers of donor-conceived children, and in popular media representations of the issue, that there must be something wrong with your life if you do not know your biological father." <https://anonymous.org/i-have-never-met-my-donor-father-and-i-have-no-desire-to-do-so/>. I first read this quote in a chapter draft of Daniel Groll's book, *The Need to Know* (forthcoming).

11. I want to be careful in my language here. I mean to indicate that this is the case for cisgender same-sex parents, and not all same-sex parents are cisgender. Donor conception and/or surrogacy is also typically assumed for same-sex parents where one parent is trans, since the possibility of normative conception may not be available (or, if it is, be widely known to be available). Different-sex couples where one parent is trans and one is cis may not face the default assumption that they used a donor, and in this case they are more like cisgender different-sex parents, who face the moral choice whether to be open about donor conception from the start, or not.
12. For those who are interested, I highly recommend the children's picture book *What Makes a Baby*, by Cory Silverberg and Fiona Smyth (Triangle Square Press, 2013) as a resource for such storytelling.
13. See, for example, this public radio story about a married couple and their asexual best friend who decided to parent together: <https://www.wbur.org/herenow/2017/11/06/three-parent-family> (accessed May 27, 2019).
14. Which is not to say they won't face other cultural hurdles—no one likes to be different in grade school.

BOOK REVIEW

Life and Death of Latisha King: A Critical Phenomenology of Transphobia

Gayle Salamon (New York University Press, 2018). 192 pp. Hard cover \$89.00, ISBN 9781479849215; paperback \$24.00, ISBN 9781479892525.

Reviewed by Chris Jingchao Ma
VILLANOVA UNIVERSITY

On February 13, 2008, fifteen-year-old Latisha King at E.O. Green Junior High School in Oxnard, California, was shot by her fellow student, fourteen-year-old Brandon McInerney. Latisha King, previously known as Larry King, had identified and come out as a gay boy and was routinely bullied for wearing high heels and makeup. Her acquaintances and other people in the courtroom and in most media coverage (including the current Wikipedia page on "The Murder of Larry King") refer to her as Larry and use the pronoun "he." On the day of her murder, King had just announced her new name as Latisha. After she made the announcement and typed it on a computer screen, King was shot twice by McInerney and died two days later. Gayle Salamon's new book *Life and Death of Latisha King: A Critical Phenomenology of Transphobia* (2018) is a unique philosophical project that presents a detailed study of the courtroom hearings of this case. Instead of examining the actions of the crime or before the crime, Salamon examines the things said in court about Latisha King and especially about her gender identity, in order to reveal the transphobic understanding of gender held by everyone in court, not only McInerney. This trial is particularly relevant to philosophical

inquiries of gender identity and embodiment because much of the debate hinges on what gender is and whether King's gender expression is an aggression in itself. If King's gender expression can be established as an aggression, then McInerney's action of manslaughter is a reactionary attack, a provoked murder, or even a self-defense taken too far. With her analysis of the court hearings, Salamon shows how transgender identity is weaponized, that is, how transgender expression is seen as an aggression and an attack towards cisgender heterosexual people. Transphobic understanding of gender is fundamentally contradictory and incoherent: it unevenly and retrospectively attributes volition and intention to queer and trans subjects; they are found guilty of committing a gender identity and hence deserve retribution. We find this recurring thesis throughout this book: "queer or trans gender becomes a target of homophobic and transphobic aggression through first being read as itself constituting an act of aggression" (10). Latisha King's gender is understood as her voluntary choice and action and an *extra* demand on others to look at her and to acknowledge her; yet at the same time, her non-conforming gender is denied any legitimacy or recognition. "Violence against Latisha, daily teasing and name calling, shoving her into the lockers as she passed by, was part of the normative fabric of the school," and these everyday instances that deny her existence happened long before McInerney's final action (156). The daily transphobic violence of bullying is made invisible, while Latisha King is made constantly visible as the cause of this violence.

For readers interested in phenomenology, this book offers a discussion and an application of the method of critical phenomenology. Salamon refers to Lisa Guenther and Werner Marx for critical phenomenology and phenomenology as ethics. Following Guenther's emphasis on both first-person experience and transcendental intersubjectivity, Salamon hopes to elucidate with phenomenology that the subject is enmeshed in the world which "is constituted by and appears as a result of certain conditions" (17). The phenomenological method requires us to withhold our daily rapid judgments and to actively withhold what we already know about this world in which we are enmeshed. For this specific project and for these kinds of projects, that is, philosophical inquiries that stem from social justice and ethical concerns, Salamon suggests that phenomenology can be helpful as a method of unknowing: "to revise or undo knowledge that I already have, perhaps to question the epistemological regime that brought that knowing about in the first place" (149). As Marx suggests, phenomenology does not simply describe our experience but aims at revealing a structural whole of our experience so that it can remind us of human mortality and sociality. In this way, phenomenology is not only a descriptive method, but can also be an ethical one.

For readers interested in trans philosophy, queer theory, feminist philosophy, and philosophical discussion of gender and self, this book weaves theoretical discussions with a fine-grained analysis of a court hearing where the definition of gender is of utmost legal and political importance. The opening statement of the trial, notes Salamon, defines gender as sex and describes King's sex as male, yet King's anatomy is never brought up in court,

exposing that the language used in court is inadequate to understand King's identity (115). For the majority of the court hearing, King's "gender" is described, presented, and imitated through gestures, mimes, and an exhibition of objects. These reenactments and scrutinization of an absent person's gender allows Salamon to examine the meaning of gender as gestures in this case and in the broader societal imaginary. Salamon first points out that King's gendered expression ("feminine") is often read together with her sexual orientation ("gay"). Although in our lived experience these two aspects are often intertwined and inseparable, here, the conceptual conflation is dangerous as it "encourages people to look at gender expression as an act, and often as an aggressive act, akin to a sexual advance or even a sexual assault" (30). Following Gail Weiss and Maurice Merleau-Ponty, Salamon examines gender as a gestural phenomenon and gestures as "a bodily means of communication" (31). She argues that embodied gestures have a style that is not reduced to biology nor completely within our voluntary control. However, King's gender, here epitomized by her gendered style of walking, is read as her willful action of defiance and hence aggression, while cisgender gestures are not read as voluntary and merely taken for granted as natural. "In contrast, then, with the cisgendered boys and girls against whom she was compared, Latisha's gender was seen as a choice, as a behavior over which she had control, and also construed as something over which she refused to exert control" (32). Consequently, Salamon argues, if one believes that transgender subjects make the conscious choice of performing their gender, it follows that this conscious performance makes a social demand because a voluntary act is seen as an act *upon* others. Hence, King's gender expression is seen as a demand for response, or in the words of her teachers: she demands attention. Salamon notes, "whereas normative gender identity, in this logic, asks nothing and demands nothing of others—it is, in effect, non-social—trans gender is understood as a provocation to the extent that it is a shared social project" (30).

To understand transgender as voluntary and demanding allows cisgender people to mask their desire to look and to accuse the trans queer subjects as the cause of their discomfort—"it's not that I wanted to look at her; she made me." Subsequently, queer and trans subjects are understood to be responsible for the everyday verbal and gestural harassment. In fact, the daily bullying happening to King does not register as harassment to the teachers, while she, by putting on lip gloss, is seen as the threat to her own safety in this logic (37). Adding the dismissed factor of race back into the picture (Latisha King was mixed race and identified as Black, while McInerney held white supremacist beliefs, but racist hate crime was ruled out as a possibility in the case), Salamon compares this case to the beating of Rodney King in that in both cases, an aggression is retrospectively read into the queer bodies, trans bodies, and black bodies as "fantasies of aggression projected onto a victim" (90).

In the second half of the book, Salamon takes a turn of perspective and discusses gender as an object of perception. As embodied gestures, gender entails a

bodily way of communicating some meanings between subjects. Another recurring theme throughout the book is a discussion of this collaborative meaning in our shared world. Drawing on Edmund Husserl and Merleau-Ponty, Salamon argues that we live “a lifeworld that is social, in which the engendering and interpretation of meaning is a relational activity” (92). The “I” only emerges in a context of a horizon shared with other embodied subjects. In another place, she argues that the “I” does not constitute the world from its subjective position—which would be the Cartesian cogito that Merleau-Ponty refutes—but perceives the world in which it lives (125). In this shared world, therefore, my meaning is better understood as a collaborative project. This, however, does not guarantee the optimal equality in collaboration. On the contrary, the concept of horizon allows us to understand the detrimental effects that power relations can have: my meanings may be canceled, may not be able to emerge, and other meanings might be read into me, against me, as “power can delegitimize other horizons, can render them as untrue, invisible, or unthinkable” (92).

Latisha King, like many queer and trans individuals, was unable to emerge as intelligible or possible. As a “boy” who wears “girls’ clothes” and makeup (Salamon points to how we assign a gender to certain clothing items and then fear their power of making that gender come true), King cannot be reconciled with the rest of students who wear their “correct” gender-related attire and enjoy their “correct” gender-related activities. She thus stands out from the everyday, mundane anonymity. The concept of anonymity, for Alfred Schütz, is “the fabric from which daily life is woven” (112). The shared social fabric allows us to live in the everyday life flow, but it is a protection from which King is denied: “she was unable to retreat into anonymity and thus unable to inhabit the shared and reciprocal life of those around her” (117). Standing out from the daily fabric, King therefore stands outside of common sense—the epistemic equivalent of anonymity—and fails to be intelligible. One of her teachers describes seeing her first as a well-dressed girl talking to her friends, “then I saw it’s Larry.” Salamon identifies here a Husserlian moment of retroactive crossing-out: the object perceived at first—Latisha being a well-dressed girl—is crossed out as it is unintelligible and therefore non-existent when the proper “Larry” is seen. This, Salamon argues, is the everyday violence “to void or cancel her existence” and to deny her life, and this happened long before the final shooting (156).

This book is particularly relevant and urgent today. As Salamon notes in the coda of the book, we currently see in our news that transgender people are portrayed by transphobic politicians as deceivers, liars, and purposeful sexual harassers in women’s bathrooms and transphobic politicians claim that cisgender women’s privacy and safety must be defended by exposing and endangering trans women, as if trans women pose the true danger to cisgender women. At the same time, we also see trans-exclusionary feminist authors and lesbian activists denouncing trans people because they are seen as imitations of the “real” gendered experience, no matter how hard they “try.” It is therefore crucial to parse out the relation between gender identity, sexual activity, sexual attraction, and subjective will. To contextualize her argument, Salamon acknowledges

that there are multiple ways to conceive the relation between these concepts, and that other philosophers, such as Talia Mae Bettcher, have thought of gender identity and sexual attraction as inevitably connected. Bettcher poses the theory of “erotic structuralism” where she argues that “a gendered eroticization of self” is an indispensable component as we think of sexual attraction. Salamon argues that the crisp distinction drawn in this book is a response to the way in which these concepts are mobilized in transphobic discourses, and especially the case study here.

Reading this book may invite us to think with Salamon and further ask these questions: We have seen the legal and political mobilization of the concepts of gender identity and sexual orientation, how can we think about the conceptual genealogy of these terms and their political valence? In other words, why are these terms mobilized and transphobic discourses produced now? Another set of questions concerns King’s racial identity, dismissed as irrelevant early on in the legal procedures. How can we use critical phenomenology to think about black femininity, black masculinity, and black transness, which are often deemed as protruding, excessive, and aggressive? As C. Riley Snorton points out, “‘gender,’ under captivity, refers not to a binary system of classification but rather to what Spillers describes as a ‘territory of cultural and political maneuver’” (12). What role does the de-racialization of King play in the courtroom and in media, in relation to her non-existent transgender identity?

Readers interested in critical phenomenology would find in this book an excellent exemplar and a resource to think with. Written with a clarity fit for readers who are not experts in phenomenology, this book would offer immense insight and inspiration for readers interested in trans, queer, feminist theory and politics.

REFERENCES

- Bettcher, Talia Mae. “When Selves Have Sex: What the Phenomenology of Trans Sexuality Can Teach Us About Sexual Orientation.” *Journal of Homosexuality* 61, no. 5 (2014): 605–20.
- Salamon, Gayle. *The Life and Death of Latisha King: A Critical Phenomenology of Transphobia*. NYU Press, 2018.
- Snorton, C. Riley. *Black on Both Sides: A Racial History of Trans Identity*. Minneapolis London: University of Minnesota Press, 2017.

CALL FOR PAPERS

The APA Newsletter on LGBTQ Issues in Philosophy invites members to submit papers, book reviews, and professional notes for publication in the **fall 2020** edition. Submissions can address issues in the areas of lesbian, gay, bisexual, trans, gender, and sexuality studies, as well as issues of concern for LGBTQ people in the profession. The newsletter seeks quality paper submissions for review. Reviews and notes should address recent books, current events, or emerging trends. Members who give papers at APA divisional meetings, in particular, are encouraged to submit their work.

DEADLINE

The deadline for submission of manuscripts for the fall edition is May 1, 2020.

FORMAT

Papers should be in the range of 5,000–6,000 words. Reviews and Notes should be in the range of 1,000–2,000 words. All submissions must use endnotes and should be prepared for anonymous review.

CONTACT

Submit all manuscripts electronically (MS Word), and direct inquiries to Grayson Hunt, Editor, APA Newsletter on LGBTQ Issues in Philosophy, graysonhunt@austin.utexas.edu.

BIOS

Dr. Amy Marvin received her PhD in philosophy from the University of Oregon in March 2019, where she focused on feminist philosophy and did organizing work for trans philosophy. You can see what she is currently up to on Twitter at @amyrmav.

Dr. Rachel McKinnon is an associate professor of philosophy at the College of Charleston in South Carolina. In addition to her work in epistemology, philosophy of language, and feminist philosophy, she is an expert on trans athlete rights and barriers that trans athletes face while competing. In 2018 she became the first ever openly trans woman to win a world championship in an Olympic sport, winning the UCI Masters Track Cycling Championships sprint event for women 35–44. She continues to be an academic, activist, and athlete promoting fairness and inclusion in sport.

Dr. Alice MacLachlan is associate professor of philosophy at York University (Toronto, Canada), where she directs the graduate program in philosophy. She writes and teaches in moral, political, and feminist philosophy, focusing on philosophical issues arising in the aftermath of conflict. She has also worked in LGBTQ philosophy on topics ranging from the ethics of coming out to the changing nature of family. Her recent publications include “Fiduciary Duties and the Ethics of Public Apology” (*Journal of Applied Philosophy* 2018), “‘Trust Me, I’m Sorry:’ The Paradox of Public Apology” (*Monist* 2015), “Gender and the Public Apology” (*Transitional Justice Review* 2013), and “Closet Doors and Stage Lights: On the Goods of Out” (*Social Theory and Practice* 2012). She is co-editor of *Feminist Philosophy Quarterly*.

Chris Jingchao Ma is a doctoral candidate in the Philosophy Department at Villanova University. Her dissertation research brings psychoanalysis and phenomenology together to offer a social and dynamic understanding of gender identity, sexuality, and the gendered body.

Dr. Grayson Hunt is the associate director of LGBTQ studies at the University of Texas at Austin. He completed his PhD in philosophy at the New School for Social Research in 2013. He specializes in transgender studies, queer theory, and continental philosophy. His work has been published in *New Nietzsche Studies*, *The Journal of Speculative Philosophy*, and *Hypatia Reviews Online*. He teaches Introduction to LGBTQ Studies and Transgender Feminisms at UT.



APA NEWSLETTER ON

Philosophy and the Black Experience

STEPHEN C. FERGUSON II AND DWAYNE TUNSTALL, CO-EDITORS

VOLUME 19 | NUMBER 1 | FALL 2019

FROM THE EDITORS

Stephen C. Ferguson II
NORTH CAROLINA STATE UNIVERSITY

Dwayne Tunstall
GRAND VALLEY STATE UNIVERSITY

In this issue of the *APA Newsletter on Philosophy and the Black Experience*, we are excited to publish the third "Footnotes to History" on Charles Leander Hill (1906–1956). Hill was the first person to translate Anthony Wilhelm Amo's work into English, and the first English-speaking scholar to write a philosophical commentary on Amo's dissertation. He was also a scholar of Philip Melancthon's thought.

We are also excited to publish Anthony Neal's article, "Freedom Gaze: Explicating the African Freedom Aesthetic." In this article, Neal begins by discussing how some events shape our self-perceptions to such an extent that they create cultures. He then contends that the African Freedom Aesthetic among African Americans is one the cultures formed from a people's perception of their circumstances and place in the world. This aesthetic arises from an American context, particularly slavery and the aftermath of legalized segregation in the late nineteenth century and the first half of the twentieth century. Then, he explains why the African Freedom Aesthetic is foundational to any attempt to explicate what African American philosophy is. He ends by distinguishing between the African Freedom Aesthetic and the black aesthetic of Addison Gayle, Amiri Baraka, and Larry Neal.

SUBMISSION GUIDELINES AND INFORMATION

The *APA Newsletter on Philosophy and the Black Experience* is published by the committee on the status of Black philosophers. Authors are encouraged to submit original articles and book reviews on any topic in philosophy that makes a contribution to philosophy and the black experience broadly construed. The editors welcome submissions written from any philosophical tradition, as long as they make a contribution to philosophy and the black experience broadly construed. The editors especially welcome submissions dealing with philosophical issues and problems in African American and Africana philosophy.

All article submissions should be between 10 and 20 pages (double spaced) in length, and book reviews should be between 5 and 7 pages (double spaced) in length. All submissions must follow the APA guidelines for gender-neutral language and *The Chicago Manual of Style* formatting. All submissions should be accompanied by a short biography of the author. Please send submissions electronically to apa.pbe.newsletter@gmail.com.

DEADLINES

Fall issues: May 1

Spring issues: December 1

CO-EDITORS

Stephen C. Ferguson II, drscferg@gmail.com

Dwayne Tunstall, tunstald@gvsu.edu

FORMATTING GUIDELINES

- The *APA Newsletters* adhere to *The Chicago Manual of Style*.
- Use as little formatting as possible. Details like page numbers, headers, footers, and columns will be added later. Use tabs instead of multiple spaces for indenting. Use *italics* instead of underlining. Use an "em dash" (—) instead of a double hyphen (--).
- Use endnotes instead of footnotes. Examples of proper endnote style:

John Rawls, *A Theory of Justice* (Cambridge: Harvard University Press, 1971), 90.

See Sally Haslanger, "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" *Noûs* 34 (2000): 31–55.

FOOTNOTES TO HISTORY

Charles Leander Hill (1906–1956)

Stephen C. Ferguson

NORTH CAROLINA STATE UNIVERSITY

Charles Leander Hill was born on July 26, 1906, in Urbana, Ohio. Hill was one of seven children born to David Leander and Karen (Andrews) Hill. He attended Urbana public schools and graduated magna cum laude from Wittenberg College in 1928. In the same year he was ordained a Deacon in the AME Church. He was conferred the Bachelor of Divinity degree by the Hamma Divinity School in 1931.

As a student at Hamma, he acquired mastery of several languages including Hebrew, both ancient and modern Greek, Latin, German, Spanish, French, and even Sanskrit.

He then studied in Germany (1931-1932) as an American-German exchange student at the University of Berlin. As John H. McClendon has observed, "The legacy of research and study in Germany, on the part of African American philosophers, is an under examined aspect in the history of African American philosophers. In addition to [W. E. B.] DuBois (University of Berlin 1892-1894), those African American philosophers traveling to Germany to further their education in philosophy before Hill included: Gilbert Haven Jones at Universities of Göttingen, Leipzig, Hallé and Jena (1907-1909), Alain Locke at the University of Berlin (1910-11), and William Stuart Nelson at the University of Marburg and the University of Berlin (1926-27)."¹

During this year, he began his research on Philip Melanchthon (German co-reformer of the Protestant Reformation with Martin Luther). Also, during his time in Germany, Hill discovered the Ghanaian philosopher Anthony William Amo's inaugural dissertation on apathy. Hill was the first person to translate Amo's work from Latin into English; he additionally provided the first commentary on the philosophical substance of Amo's thesis. Later, Hill decided to publish his translation and commentary as an article entitled "William Ladd, The Black Philosopher from Guinea: A Critical Analysis of His Dissertation on Apathy" to *The AME Review*, and it later appeared in the October/December issue of 1955.

After returning from Germany, he received an S. T. M. from Wittenberg in 1933. In 1934, he became an Elder of the AME Church. During this time, he had already assumed the position of Dean of the Turner Theological Seminary of Morris Brown College, in Atlanta, Georgia, where he remained until 1944. During his time at Morris Brown, he taught philosophy and theology. Morris Brown College, founded in 1881, is an AME-affiliated institution and the only college in Georgia established by African Americans. The College is named after Bishop Morris Brown, the second elected and consecrated Bishop of the AME Church.

In 1938 he earned a PhD in philosophy from Ohio State University. His dissertation topic was titled "An Exposition and Critical Estimate of the Philosophy of Philip Melanchthon." Fluent in several languages, including Latin and German, Hill had begun research on this co-reformer of Martin Luther during his studies at Hamma Divinity School using primary documents in Latin and German. During his tenure at Turner Theological Seminary, he was an exchange professor at Atlanta University.

From 1944 to 1947, he taught sociology at Allen University in Columbia, South Carolina, and simultaneously was pastor of Bethel AME Church in Columbia. At the end of 1947, Hill was inaugurated as the thirteenth president of Wilberforce University (Wilberforce, Ohio).

Along with functioning at the helm at Wilberforce, Hill took on the responsibility of teaching classes in philosophy. In fact, unable to find a suitable text for classroom instruction,

Hill wrote a history of modern Western philosophy that was published in 1951. With *A Short History of Modern Philosophy from the Renaissance to Hegel*, Hill became the first African American philosopher to publish a book on the history of modern philosophy.

Unfortunately, on December 8, 1956, Hill's brilliant career as an educator and philosopher came to a sudden end. Hill died after a stroke at the relatively early age of fifty years old.

As one of the leading scholars on Phillip Melanchthon, Hill belonged to several learned societies both in the United States and abroad, including the Royal Institute of Philosophy of Great Britain, American Philosophical Society, American Association for the Advancement of Science, and the American Academy of Social and Political Society.

The Charles Leander Hill Manuscript Collection is housed at Wilberforce University.

PUBLICATIONS

Dissertation

Hill, Charles Leander. "An Exposition and Critical Estimate of the Philosophy of Philip Melanchthon." PhD diss., Ohio State University, 1938.

Books

A Short History of Philosophy from the Renaissance to Hegel. Boston, MA: Meador Publishing Company, 1951.

The Evangel in Ebony. Boston, MA: Meador Publishing Company, 1950.

Melanchthon, Philipp, and Charles L. Hill. *The Loci Communes of Philip Melanchthon, with a Critical Introduction by the Translator, Charles Leander Hill*. Boston, MA: Meador Pub. Co., 1944.

Melanchthon: Selected Writings, translated by Hill and posthumously edited by Elmer Ellsworth Flack and Lowell J. Satre.

Articles

Hill, Charles Leander. "William Ladd (sic) [Anthony Amo]—The Black Philosopher from Guinea: A Critical Analysis of His Dissertation on Apathy." *The AME Church Review* 72, no. 186 (1955): 20-36.

———. "Melanchthon's Propositions on the Mass." *Lutheran Quarterly* 6, no. 1 (February 1954): 53-57.

———. "Some Theses of Phillip Melanchthon." *Lutheran Quarterly* 6 (August 1954): 254-48.

———. "American Democracy." *The Negro Journal of Religion* 6, no. 2 (March 1940): 9.

Unpublished Works

A Philosophy of the Episcopacy: A Scientific Brochure. 1934.

Secondary Sources

McClendon, John H., III. "Charles Leander Hill: Philosopher and Theologian." *The AME Church Review* 119, no. 390 (April-June 2003): 81–105.

Stokes, Arthur P. "Charles Leander Hill: Profile of the Scholar." *AME Church Review* 117, no. 379/380 (Fall 2000): 26–34.

NOTES

1. John H. McClendon III, "Charles Leander Hill: Philosopher and Theologian," *The AME Church Review* 119, no. 390 (April-June 2003): 89.

ARTICLE

Freedom Gaze: Explicating the African Freedom Aesthetic

Anthony Neal

MISSISSIPPI STATE UNIVERSITY

First we must define the terms 'noun' and 'verb,' then the terms 'denial' and 'affirmation,' then 'proposition' and 'sentence.'

Spoken words are the symbols of mental experience and written words are the symbols of spoken words. Just as all men have not the same writing, so all men have not the same speech sounds, but the mental experiences, which these directly symbolize, are the same for all, as also are those things of which our experiences are the images. This matter has, however, been discussed in my treatise about the soul, for it belongs to an investigation distinct from that which lies before us.¹

Signs, symbols, experiences, and frameworks, this is how humans perceive the world. Attempts to describe or express the meaning of a perceived moment of experience can be affected by many factors. The attempts to perceive something or the products of what is perceived are known as the aesthetic, or at least what gets perceived is perceived through the aesthetic. But what does it mean to perceive something, and why don't we all perceive experiences the same? Beyond the nature of perception and what is perceived, what causes a change or shift in an aesthetic? In this paper, I will firstly define "perception" and secondly determine how some events shape human perception such that cultures are created. Finally, I will argue that there is an African Freedom Aesthetic among Africans in the diaspora that is demonstrable in creative productions. I regard African American culture as the foremost exemplar culture of this phenomenon and examine the African Freedom

Aesthetic through an examination of the specific linguistic signs and symbols of this particular aesthetic. I will also show why I think this aesthetic to be foundational to any attempt to explicate what African American philosophy is.

To truly philosophize in earnest, by my understanding, is to create separation, which simply means to identify a thing or phenomenon. Next, the entity must be interpreted, and then it is represented to or communicated in the phenomenal world. We do this first through conception, next through the releasing of air in order to make noises we call verbal symbols or words, and lastly through markings or written symbols. All of this is contingent upon spatial-temporal dynamics such as where and when one perceives a thing or phenomenon and is always contextual. It is also contingent on the appetites and aversions of the ones who perceive the world. This filtering of phenomena through time and space, as well as through our appetites and aversions, is the origin of perception. So, as the Greek word *αισθησις* means sensation, consciousness, knowledge, or perception, then aesthetics is the science or study of perception.² Since all philosophers base their philosophizing on some foundational assumptions, I believe that it is good to make that assumption known and not keep it hidden from your audience. This paper, as well as my previous philosophical research, is grounded in the basic assumption that to separate, interpret, and represent an experience or phenomena to others with the greatest amount of clarity, it is necessary to begin with the study of the aesthetic. This is because, as suggested earlier, the world is seen very differently in different spaces, under different conditions, and among different people.

Blackness in the American context is one such condition. To truly communicate its meaning to others, we must tease out its differences from other cultures, creating the ability to discuss it as a particular, rather than attempting an understanding of Blackness in general. Some use a Marxist epistemological framework to form their analysis of what is Blackness; however, this type of analysis does not capture the full scope of Blackness.³ While narrowly focused on the metaphysics of the conditions African Americans find themselves, this type of analysis seems to miss the phenomenon of Blackness and therefore fails to give an accurate account of the subjective actions and reactions of African Americans to the American experiential moment.⁴ In order to get at the variegated nature of Blackness, some attempt to separate it from other cultures must occur. This separation is performed by categorizing or defining exactly what is meant by Blackness in the American context. Of course, Blackness in this respect is understood to mean African American, but has in the past been referred to as Negro and by the pejorative nigger. Its defining characteristics have sometimes been arbitrarily ascribed, to say the least, but they can range from skin color and hair type to having one Black relative, regardless of how distant. The meaning of Blackness has been permanently and ironically predicated as Black in the negative connotation of Black, which can be understood as not white and therefore not deserving of whatever is considered normative treatment to humans. This description may seem illogical, but it is meant to demonstrate a clear vision of Black existence in the American moment. Representations of Blackness by

others, whether in the arts or letters, has mainly to do with the accentuation of differences to ensure the separation between ethnic groups, which has historically meant placing Blacks in an illogical and precarious status that ensures their devaluation.⁵

Blackness, as constructed in this way of being, in this experiential/aesthetic moment has formed the perceptions for those beings, which can be considered to inhabit such an existence. This is to say, Blackness is at once a sign of death, a symbol of depravity, an experience of oppression, and a framework for struggle, while also acting as a veil behind which exists a multiplicity of accurate expressions of life. As humans, our epistemological foundation for all knowledge is always subjected to our perception and moment of existence. "Moment," although usually connected to time only, is in this paper referring to the relationship of space and time in the creation of culture and culture's relationship to the general thought process of a people, which is significant in the shaping of those people's perceptions. A person's ability to be conscious, the degree to which they are conscious, and whatever the aim of their consciousness are direct reflections of, and are dependent upon, when and where they occur.⁶ I will call this phenomenon "the experiential moment." A working definition of experiential moment arises from the concept of the lived experience, which I take to mean a phenomenological account of an individual's or group's reality. The experiential moment deepens this description of lived experience by attempting to account for those significant experiences which have charted or changed an individual's or group's experience in a substantial way such that their consciousness is also affected, i.e., the experiences of being racialized, slavery, and war.

Consciousness always means conscious of a particular phenomenon.⁷ In this sense, consciousness always has an aim. In my earlier work, I define "consciousness" as awareness intentionally focused on a physical object or mental creation that is subjectively understood.⁸ The phrase "subjectively understood" is pivotal in that it gives credence to the more thorough meaning of perception. This definition of consciousness highlights a level of contingency inherent in the concept of consciousness. It is a contingency based upon the perception of the subject as the subject participates in a subject-object relationship. In this way, the perception, or gaze, one has of the world now becomes a type of dialectic relationship, in that the cognition of what is perceived depends on the perceiver as well as that which is being perceived in this experiential moment. To put it plainly, as it pertains to Black people, Black people see white people, and they see how white people respond to seeing Black people. Black people see the structures or institutions in this society and the responsiveness of the institutions towards Black people. The African Freedom Aesthetic is a direct response to this dialectical relationship of this experiential moment. It is when this dialectical relationship is trivialized or ignored that mischaracterizations of it occur. These mischaracterizations are grounded in wrongful historical idealizations and thwart any enterprise concerned with a true understanding of a particular people within a particular experiential moment, in this case African Americans.

Clarity can be gained on this matter by connecting the concept of the gaze found in the works of Laura Mulvey (803–816) and bell hooks (115–131) to this present work, with an aim towards distinguishing between their use of the concept and how I use it here. Mulvey and hooks contend that the gaze involves three entities, each of which enters into the moment of the gaze with a different level of awareness and has a status that is determined by the manner in which they enter the gaze moment. The gaze moment is crucial, because it is the point of interaction between the entities. It is also the moment that the power differential between the entities becomes apparent, and the entity with least power becomes objectified. The gaze moment sets the stage for all other subsequent activity and conceptualizations of the ones who gaze or who are the ones being gazed at. There is either a rejection of the moment or a reluctant acquiescence, but there is a reaction to the moment by all entities who participate in the gaze moment.

Although there is no necessary order to the listing, the first entity is the creator of the gaze moment. Here, I am describing the gaze moment in its film application. The creator of the gaze moment is the entity responsible for the creation of the movie or film. Their status is active owing to their having created the space upon which to have the gaze moment. It is their bias that creates a positive or negative gaze. Mulvey and hooks only discussed the negative gaze in their writings on it. The other two entities are the female and the male. Mulvey and hooks focus upon the power dynamic that exists between male and female, with female being viewed as less powerful and hence subordinate to male. Of course, they each utilize three entities—the creator of the gaze moment and the two entities being gazed at, the male and the female. However, in this work there exists only two entities (Black and white). The shift in focus from gender to race is important, because I essentially flatten the conceptualization of the moment as non-sexed, but race-determined. I do this intentionally to demonstrate that I believe the nexus of racial oppression (the conceptual level) is at the genus level while the actual experience may be more specific. It is at the genus level that Blackness becomes the DuBoisian veil.

In this experiential moment, even when Blackness is not viewed as a type of depraved existence, it is mostly viewed as other even by Blacks themselves. This is to say Blackness is viewed as the antithesis of the human. Black people exist in an illogical state in which whenever they try to gain an understanding of themselves as selves in a world where they are viewed as abnormal or other. Over the centuries, this continuous illogical state has been a factor in the creation of a differing aesthetic. However, I think it necessary to state that it is not the Black body that is the origin of this aesthetic, for that body has historically been conceived as an object that cannot be a subject of its own experiences. I also think that it is necessary to state that Black existence is neither logical nor illogical; Black existence, at least as a manifestation of being itself, simply is. We can only classify our perceptions, conceptions, and descriptions of phenomena as logical or illogical.

As far as being is concerned, if it can be accepted that being is not static, but is always in transition, and this transition is affected by experience, then experiential moments become crucial to the understanding of being. This means that in order to give a proximally accurate depiction of being in a given experiential moment, then the particular being along with all of its available experiences must be considered. This is necessarily true of the individual as a particular member of a group, and it is also necessarily true of the group, as the group becomes an individual when under observation.

Plessy v. Ferguson (1896) is an example of a significant experiential moment that changed the lived experience for African Americans in a substantial way and affected their consciousness, for this is the court decision that legalized the concept of separate but equal.⁹ This decision went beyond affecting class in an economic sense; it had such a broad ranging effect on Black people as to determine everything from your birthplace to where you could be buried. It determined nutritional intake, genetic makeup (based upon procreation partners), religious practices, intellectual pursuits, the scope of ethical behavior. It also determined Black people's access to life, liberty, and the very pursuit of happiness. Whites could define the limits of life, liberty, and happiness for themselves and for everyone else, usually reserving the most restrictive definitions of these terms for those with African ancestry. It was in this experiential moment that the Black perception of Blackness was formed.¹⁰ It was a perception formed of a distant African past, the rejection of a not-so-distant past in captivity, and an immediate present defined by a struggle to teach the meaning of humanity through the development of a human community, in spite of being denied its benefits.¹¹

It can be said that this particular use of experiential moment here can be confused with a concept of epoch or milieu. For this reason, one needs to consider the horizontal temporal-spatial existence of Black people during this historical period, while offering a phenomenological and even processual interpretation of this historical period as part of the analysis of the experiences had by Black people living during this historical period. To this end, any analysis of Blackness must, of course, be emic in nature, bearing in mind that the experience of Blackness in the American moment is particular. However, the interpretation and representation of this experience must also consider the African Freedom Aesthetic. From a phenomenological perspective, African Americans must not be taken as object, but instead as subject with a particular intentionality just as other groups are. As such, and giving consideration to the African Freedom Aesthetic, the realization of the difference between Blackness and Black people should be viewed in a similar manner as the distinction between a thing as seen or as it appears and what it is. Sometimes the two are simultaneously aimed at when studying the experiences of Black people, but in most cases, they are not the same phenomenon and they should never be taken as being identical. An example of this can be found in the well-known poem, "We Wear the Mask," by Paul Lawrence Dunbar:

We wear the mask that grins and lies,
It hides our cheeks and shades our eyes,—
This debt we pay to human guile;
With torn and bleeding hearts we smile,
And mouth with myriad subtleties.

Why should the world be over-wise,
In counting all our tears and sighs?
Nay, let them only see us, while
We wear the mask.

We smile, but, O great Christ, our cries
To thee from tortured souls arise.
We sing, but oh the clay is vile
Beneath our feet, and long the mile;
But let the world dream otherwise,
We wear the mask!¹²

The African American Freedom Aesthetic is also affected by the fusion of horizons of many peoples' experiences. Black people in the American moment cannot exist wholly apart from others; they are affected by being culturally immersed in their times. In this American moment, Black people have developed music, literature, and systems of spiritual formation. However, all of these developments were influenced to some degree by surrounding cultures. To be sure, even under tremendously adverse conditions, African Americans were able to prehend positive experiences useful to their aim such that they participated in their own development or auto-poieses.¹³ The many rebellions, attempts at revolution, and engagements in wars go a long ways towards demonstrating that regardless of the plight, African Americans did not see themselves as a determined or predestined people for the oppression they encountered.¹⁴

So what makes it necessary to put forth this description of the aesthetics of Black people in this moment? Most of my work is enveloped in the philosophical quest to understand Blackness in the American context for the purpose of determining just what can be called an African American philosophy. As such, I am concerned with the methods of just how such a study should be performed. Since I am not the first to make this attempt, I have given large amounts of time considering the work of others who have also made this journey. For the sake of time and space, I will certainly not outline all of the attempts to create an African American philosophy, but I will discuss the attempt I find to be most problematic. Before doing so, I would like to identify three schools of thought pertaining to African American philosophy and their problems:¹⁵

1. African Americans doing philosophy is African American philosophy.
2. Any philosopher of race focusing on the nature of the Black experience, whether analytic or continental in method, with the aim of offering correctives to bring about eudaimonia. It is mostly political in nature.
3. Constructive or corrective philosophy, which rejects Eurocentric schools of thought as normative and focuses on what would be African American philosophy if not for slavery.

It is the third school, which I find to be most problematic, especially as far as my research is concerned. This group is most problematic because it is necessary for this group's philosophy to put forth a theory of what is the meaning of Blackness in the context of America. Within this group are those Afrocentric theorists whose foundational assumption is based on what is called Location Theory, and they also maintain that there is an essential nature to being Black.¹⁶ Without explicating the details of this theoretical framework, it suffices to say that if it is maintained that there is an essential nature to being Black, then any understanding of African Americans' Blackness as being historically and spatially conditioned is not a viable option.

Why does this matter? Its importance is rooted conceptually in the foundational questions of African American philosophy, beginning with the question, "What does it mean to be Black?" This question loses significance if at any time a group of Black people appears in history we can simply refer to them as African. To do so would allow us to make certain assumptions about what it means to be African and that the term "African" provides a basic understanding of the culture of the group to which it is applied, regardless of where they are located and when they were or are alive. This is simply not so. Since this is not so, then the basic question must be revised so that one can ask what does it mean to be Black in America? This question acknowledges that being Black in America and being African or being Black in Africa are not the same, or better even there is no transcendent essential nature to being Black. Blackness in the American context is different in significant ways from Blackness in the African context. Although I would admit that there are significant similarities, to speak of them as if there are no significant differences between the two ways of being Black reduces any claims of the effects of slavery and oppression on both groups.

This is why I must refer to the American moment. It was the American moment that created a particular consciousness, the African Freedom Aesthetic, which is the main lens of African American philosophy. African American philosophy is distinguished from other genres of philosophy by society's objectification of its main subject of discourse, which is the Black body. In other words, African American philosophy is necessarily humanistic in its discourse and as a field of inquiry because of the requisite goal of reclaiming the human status for Black bodies. In many ways African American philosophy was set in motion, just as Greek philosophy, in its rejection of the previous worldview and/or perceptual frameworks. The relevance of African American philosophy is derived from the context of enslavement. While other groups certainly were enslaved in the Americas in addition to Africans, it was the freedom struggle of those people of African descent that in many ways shaped the freedoms that all Americans enjoy through a prolonged social movement. Not all Africans, but particularly those in the American diaspora.

How then does one do African American philosophy? In other words, who can and is doing African American philosophy? This is simple! Anyone can do African American philosophy, but there should be some distinction made between studying African American philosophy (i.e., the

writings of any African American philosopher) and doing African American philosophy. However, this is a major bone of contention among philosophers who are drawn to this discourse. I am contending that African American philosophy must have as a central concern the conflict between the lived experience of oppression and the desired experience of freedom. African American philosophy must toil with the variance between normative experience being equated with humanity, which would be the same as equating it with whiteness in the American context, and the Black experience, which was thought to be insignificant. It must be suspicious of blind patriotism just as it should be suspicious of any type of blind faith. The questions and issues of African American philosophy are not limited to these. Nor are these questions limited to African American philosophy, but these issues and questions in conjunction with an attempt to understand the experience of Blackness must be central and not peripheral to African American philosophy. In short, African American philosophy begins by rejecting the definition of the Black body as a tool defined by the institution of slavery by defining Black bodies as human and also free, and by expressing the experience of Blackness in a phenomenological manner (which simply means expressing this experience based on the perception of the expresser). Particularly when studying American political thought, since much of this thought was shaped either in contention or in conjunction with the African American freedom movement and African American philosophy, it would seem that the full picture of American political thought cannot be had without a firm understanding of African American philosophy which is bolstered by understanding the African Freedom Aesthetic.

I would like to end this paper by briefly distinguishing the African Freedom Aesthetic from "the Black aesthetic." The Black aesthetic is time locked as a product of the Black arts movement beginning in 1965 and is mostly the creation of Addison Gayle, Amiri Baraka, and Larry Neal.¹⁷ It was not meant to explain as much as it was meant to be a creative impetus for a particular kind of protest literature and artistic creations. The name "African Freedom Aesthetic" serves more of a descriptive function for a particular meta-philosophical framework and cultural memory, which has shaped a group of people, namely, African Americans. This shaping began instantaneously when Africans arrived in America, but it reached its zenith in 1896 with the *Plessy vs. Ferguson* decision, bringing the question of what is Blackness front and center.¹⁸ While developing the African Freedom Aesthetic, African Americans became a discreet people, immersed in many cultures, adverting some while also intentionally subsuming others, all with the aim of becoming free.

NOTES

1. Aristotle et al., *The Categories on Interpretation*, 115.
2. Heidegger, *Introduction to Phenomenological Research*, 5.
3. Ferguson, *Philosophy of African American Studies*, 7.
4. Neal, *Common Ground*, 35.
5. Davis, *Darkest Hollywood*, 12.
6. Heidegger, *Introduction to Phenomenological Research*, 35.
7. *Ibid.*

8. Neal, *Common Ground*, 2.
9. Hartman, *Scenes of Subjection*, 9.
10. *Ibid.*, 10.
11. Scott, *Crimes Against Humanity in the Land of the Free*, 75.
12. Quoted in Black, "Literary Subterfuge: Early African American Writing and the Trope of the Mask," 387.
13. Whitehead, *Process and Reality*, 32.
14. Thurman, *Deep River and the Negro Spiritual Speaks of Life and Death*, 39.
15. Gordon, *An Introduction to Africana Philosophy*, 91.
16. Welsh-Asante, *The African Aesthetic: Keeper of Traditions*, 53.
17. Johnson, *Being and Race*, 22.
18. Neal, *Common Ground*, 1.

BIBLIOGRAPHY

- Aristotle, et al. *The Categories on Interpretation*. Trans. H. P. Cooke. Cambridge, MA: Harvard University Press, 1938.
- Black, Daniel P. "Literary Subterfuge: Early African American Writing and the Trope of the Mask." *CLA Journal* 48, no. 4 (2005): 387–403.
- Davis, Peter. *Darkest Hollywood*. Ohio University, 1996.
- Ferguson, Stephen C. *Philosophy of African American Studies*. Palgrave, 2015.
- Gordon, Lewis. *An Introduction to Africana Philosophy*. Cambridge, 2008.
- Hartman, Saidiya V. *Scenes of Subjection*. Oxford University Press, 1997.
- Heidegger, Martin. *Introduction to Phenomenological Research*. Indiana University Press, 2005.
- hooks, bell. *Race and Representation*. South End Press, 1992.

- Johnson, Charles. *Being and Race*. Indiana University Press, 1990.
- Mulvey, Laura. "Visual Pleasure and Narrative Cinema." *Film Theory and Criticism: Introductory Readings*, edited by Leo Braudy and Marshall Cohen. Oxford University Press, 1999.
- Neal, Anthony. *Common Ground*. Africa World Press, 2015.
- Scott, Imani, ed. *Crimes Against Humanity in the Land of the Free*. Praeger, 2014.
- Thurman, Howard. *Deep River and the Negro Spiritual Speaks of Life and Death*. Richmond, IN: Friends United Press, 1975.
- Welsh-Asante, Kariamu. *The African Aesthetic: Keeper of Traditions*. Praeger, 1994.
- Whitehead, Alfred North. *Process and Reality*. The Free Press, 1978.

CONTRIBUTOR BIO

Anthony Sean Neal is an assistant professor of philosophy in the Department of Philosophy and Religion and a Faculty Fellow in Shackouls Honors College at Mississippi State University. He is author of *Common Ground* (Africa World Press, 2015) and *Howard Thurman's Mystical Philosophy: Love Against Fragmentation* (Lexington, 2019). Recently, he was inducted into the Morehouse College Collegium of Scholars. His areas of specialization are African American philosophy and African American philosophy of religion, with additional interests in critical theory, neo-Platonism, philosophy of religion, and political philosophy.



FROM THE CHAIR

Peter Boltuc

UNIVERSITY OF ILLINOIS SPRINGFIELD

The gist of my term as the chair is dictated by the announcement “Changes to APA Committees” posted on February 23, 2018, especially by the following passage: “the board has made the decision to wind down the Committee on Philosophy and Computers, Committee on Philosophy and Medicine, and Committee on Philosophy and Law.”¹ The above webpage allows APA members to post comments, but there were no comments from (or on behalf of) the other affected committees and only one comment at all, by Fritz J. McDonald, at the time a member of this committee, dated February 26, 2018. The gist of Fritz’s posting can be summarized by its last sentence: “While information technology grows more and more central to our lives, the APA decides to eliminate the committee on this area.” The fact that there was no follow-up does not necessarily mean that nobody agreed with the important aspects of McDonald’s comments, but it indicates something more practically important—that there was no overwhelming interest in reopening this issue at the time. I understand the issue to be closed.

Two things would be important to take up. First, the causes and reasons for elimination of this committee. Second, practical issues that pertain to its activities, which include organization of sessions at the APA meetings on philosophical questions in information technology, legacy of the *APA Newsletter on Philosophy and Computers*, and the future of the Barwise Prize. I am a big fan of academic freedom, and democracy in general, and so I think those issues should be put under broad consideration that goes far beyond membership of the committee. The role of the committee, and even more so of its chair, is to formulate productive topics for consideration, create the forum for stakeholder discussion, and listen to any conversations that may ensue.

1. SOME OF THE REASONS AND CAUSES.

1a. *First, my take on the historical background.*

Technological changes created groups of enthusiastic early adopters around the turn of the twentieth century in the areas such as computer programmers, web-development, online teachers, and so on. Early adopters were also the core of philosophers interested in computers, some of them—such as Jon Dorbolo, Robert Cavalier, Anthony Beavers,

Ron Barnette, Bill Uzgalis, and Marvin Croy—have shaped, directly or through the CAP movement, this committee—at least for the first half a dozen years. There was much enthusiasm and people were devoted to *the cause*, though detailed understandings of the cause varied. The committee met regularly and the links between CAP and the committee were strong. The thriving of the committee can be seen in the spring 2004 issue of the newsletter, especially in the report from then chair Marvin Croy.²

The strong link between the committee and CAP resulted in the fact that the distinction between the two remained vague. There were worries at the APA that the committee was not truly open to the philosophers not associated with CAP. This resulted, around 2003-2004, in the influx of committee members unconnected with CAP—I think I actually benefited from this change by being able to join the committee on this wave. Those important and potentially beneficial changes gradually decreased cohesion of the group. Some of the members participated in no activities or were not available for physical meetings (much later this was the case even with some of the committee chairs). As can be seen on the committee’s website, one of the chairs failed to even post obligatory reports throughout the whole term in office,³ and one of these years we even failed to award the Barwise Prize (despite a valid vote taken by committee members). Also, the link with APA executive directors and persons responsible for the APA’s electronic presence gradually weakened, although the committee always had some of the experts that could have been helpful in making decisions on the electronic technology and web presence. The original mission of the committee has been drafted as a rather basic document, so as to build a coalition even with those members of the APA board who, at the time, did not see computers as relevant for philosophy, except for typing their articles. There was always a feeling that we needed to wait a bit longer to show the committee’s true colors.

I do not think we should belabor on the historical background, but some level of clarity is essential and further clarification, especially from the colleagues involved in the first years of the committee, would be very welcome.

1b. *The mission of the committee.*

Already during the chairmanship of Michael Byron, around 2008, we started getting encouragement from the APA leadership to update the committee’s mission. This was not viewed as an urgent task, though it started some reflection. Much of it was lost during the following years. The second nudge from the APA leadership is visible in the committee

report for 2015-2016.⁴ At the very end of his term as chair, Tom Powers received a clear message from the APA for the committee to update its mission, preferably including its name. It would have been an easy way to sneak out of the now endangered with extinction class of “philosophy of” committees and to align the written mission with what our real activities were. Unfortunately, proper attention to the mission statement was put second to day-to-day operation. Committee leadership (Marcello Guarini, then the chair, and I, then the vice-chair) revised the committee charges as late as 2018, which was after the APA Executive Committee resolved to discontinue it; as the saying goes, it was *too little too late*. As the 2017-2018 report indicates, those revised charges were “well received by the committee,” which makes it a bit murky whether they have been formally adopted. Assuming the affirmative, committee charges now are as follows:

The committee works to provide forums for discourse devoted to the critical and creative examination of the role of information, computation, computers, and other computationally enabled technologies (such as robots). The committee endeavors to use that discourse not only to enrich philosophical research and pedagogy, but to reach beyond philosophy to enrich other discourses, both academic and non-academic.

As one of the first steps as the chair, I submitted those charges to an up-or-down vote by the committee and they have been adopted unanimously. Those current charges are now on the agenda of the APA Board in its November meeting with our hope for approval. The reason for the up-or-down vote has been the lack of time for philosophical discussions on this. I do not view the above as perfect, but we have been trying to formulate the perfect mission statement for almost a dozen years and time has run out on us. The reason for working on the charges at all, the charges for a committee being closed down in a matter of months, is for the sake of clarifying what we represent, the issues we have developed or needed to be working on. For the most part, this definition is meant to be descriptive of what we have been doing at the committee sessions, in the newsletter, and in other ways. Some of those tasks need to be articulated in order to be explicitly taken over by the APA when the committee is gone, while others may need to be pursued by a follow-up group or groups after the committee’s retirement. The above is just an introduction to the more practical discussion.

1c. The reasons for the Executive Committee’s decision.

Back to the announcement from February 23, 2018. The first reason for closing the three “philosophy and” committees is that those committees were created to address pressing needs, and those needs no longer exist. The claim is addressed to the three rather different committees and so it is overly broad to allow for fruitful discussion. However, the second reason, which I list below, is addressed specifically to this committee.

The second reason is that the board understands the committee’s mission to deal with “the use of computers

by philosophers for instruction, writing, and publishing,” which at the time the committee was created “was relatively unexplored territory.” Is it a misunderstanding? Well, it may be a *de facto* misunderstanding of our activity, but—guess what?—this is what the committee’s official charge was through 2017, and it was not corrected despite the committee being asked, at least twice, for a major update. For the sake of clarity, here is what the charge was in 2017:

The committee (created by the board in 1985) collects and disseminates information on the use of computers in the profession, including their use in instruction, research, writing, and publication, and it makes recommendations for appropriate actions of the board or programs of the association.

The Executive Committee was right to take the charge of the committee at face value—the fact that it had not been revised properly lies on the committee, primarily on the committee’s leadership for the last dozen years. Of course, every committee member, and especially myself as a long-standing member and newsletter editor, could have moved the mission changes forward—but we’ve failed. That’s the answer to Fritz J. McDonald’s well-meaning comments. The Executive Committee does not deal with Platonic images of the committee; it does not even evaluate it primarily based on the content of its sessions, newsletter, or oral testimonies. The Board of Officers is supposed to focus primarily on its reports and even more so on the mission statement. Organizations unable to pass a basic test of revising their antiquated mission statement are likely to be dysfunctional also in other ways, or so they seem.

2. SOME OF THE PRACTICAL ISSUES FOR THE REMAINING YEAR.

2a. Organizing sessions on philosophical questions in information technology at APA meetings.

Currently we have five proposals for the 2020 APA Sessions. The session “Philosophical Approaches to Data Justice,” organized by Daniel Susser, has been accepted by the Eastern Division. The sessions “The Unreasonable Effectiveness of Logic in the Computational Sciences,” organized by Gary Mar, and “Women in Tech: Things You Need to Know,” organized by Susan Sterrett, have been submitted to the Central Division. A session titled “Machine Consciousness and Artificial General Intelligence” and a Barwise Prize award session are being finalized to be submitted for the Pacific Division.⁵ With this level of interest in organizing solid sessions related to the committee’s actual mission, it would be a waste to lose this capacity.

If there is a silver lining, it comes in here. In correspondence with my predecessor, Marcello Guarini, the APA offered to give the status of an affiliated group, if a group was built out of the current and former members or activists of this committee. As I understand, it is not trivial to gain such status. More importantly for the issue at hand, such a group has the right to propose a session for each of the APA divisional meetings. I understand that it should not apply for more than one session for each divisional meeting, and that such applications are prioritized just below those by

the committees. I think that this is a relatively good deal—perhaps too good to pass up.

However, we need a thorough, democratic discussion about whether to create an affiliated group. Even more importantly, we would need to define what such a group would need to focus on and who would want to give the time to develop it.

2b. Active legacy of the APA Newsletter on Philosophy and Computers.

It is clear that the *APA Newsletter on Philosophy and Computers* may not be published outside of the APA, not under this name. While *circa* 2012 we had a request from the APA to turn our newsletter into an APA journal, it was long before the APA established its official journal.

The issue at hand—and I am talking here with my hat as the newsletter’s editor on—is to preserve and enhance the influence of our newsletter’s legacy.

The *APA Newsletter on Philosophy and Computers* has played multiple roles.

A. Newsletter as a documentary of the committee’s past.

It documented accomplishments of the committee, recorded its history, and recognized people active at the committee. This function was predominant during Jon Dorbolo’s editorship (the first five years or so), but it has been a vital function of every issue of the newsletter—including the current issue. For this, it is important to gain the APA’s commitment of keeping available the newsletter “forever,” which in practical terms means, at least, while the organization exists. It would also be good to allow somebody, at least our “affiliated group” (should we create one), to mirror those newsletters on their website, as part of our shared legacy.

B. Newsletter as a repository of major philosophical masterpieces.

Several major philosophers have decided to publish their original articles with us, largely trusting that the name of the APA and the open access status of the newsletter would guarantee their work’s survival and high visibility.

Those masterpieces include two original articles by Hintikka, organized by M. Kolak; an important paper by John Pollock, published posthumously, in our newsletter, by Terry Horgan, charged by Pollock’s family to find the most appropriate place for this 53-page-long article (Terry also wrote a substantial introduction); an important article by Lynne Rudder Baker, with commentaries by Amie Thomasson and other top philosophers of the younger generation; original works by Gilbert Harman, Bernard Baars, Stan Franklin, Susan Stuart, Greg Chaitin, and many up-and-coming scholars; and, also, Barwise Prize winners such as Luciano Floridi (we published several of his articles since 2002, and a number of important commentaries on his work), J. Moor, T. Bynum, W. Rapaport, J. Copeland and G. Piccinini.

Those and many other outstanding articles need not only secure preservation but also promulgation. Many journals today, including the open access ones, help organize anthologies based on their content. It would be a great project to undertake, maybe by working directly with a publisher or to be undertaken by the new affiliated group together with the APA. Those topics should remain on the table.

Due to the changes in the manner in which APA Newsletters have been presented at the APA website, which took place *circa* 2013 (that eliminated webpages and kept the newsletters only as PDFs), currently the articles published in this and other newsletters are practically non-web-searchable. I have been working with my former office assistant on producing a list of all the articles published in the newsletter, which—if completed—may serve as the beginning of an easier-to-search catalogue.

C. Newsletter as a living journal.

Finally, there is a question of producing content of the sort this newsletter has been. A follow-up group may want to do so, without the APA affiliation. Within the APA, as Amy Ferrer recommended in a recent email, we “might consider working with the APA Blog to develop a periodic blog series—they do that for some committees and I expect would be willing to work with affiliated groups as well.” Many other options exist as well, but starting a new journal by a different group may not rank high on the committee’s busy agenda.

Again, we need a broad discussion among the many stakeholders—at the APA, in the committee, and out in the community—to work out the best ways to clarify and satisfy at least the top two of the above objectives. However, first, we want to work on building some approximation of a consensus on what “we” are going to be starting July 2020—and whether “we” want to be anything, as a group.

2c. Future of the Barwise Prize.

The Barwise Prize has been approved by the APA, at the request of CAP, as a unique committee-based APA prize. The APA does not mean to stop awarding it. It is meant to “officially be put under the oversight of the larger APA prize committee—the Committee on Lectures, Publications, and Research.” (as stated by Amy Ferrer in a recent communication). Amy continues in the same message, “we will continue to ensure that appropriate specialist expertise is part of the selection process, and we can certainly discuss a role for the new affiliated group in that process. Perhaps the affiliated group could be given a set portion of the seats on the Barwise Prize selection committee, for example.” This is a step in the right direction, and also an invitation for further discussion. This approach should be appreciated and acted upon by the committee.

Again, we should gather the relevant stakeholders. I think that primary group of stakeholders in a position to shape up the future of the Barwise Prize are the past winners of this prize. But, of course, the option of the “affiliated group” being given “a set portion of the seats on the Barwise Prize

selection committee" is very much worth keeping on the table.

2d. New proposals.

We should be very much open to new topics and initiatives. This should pertain to the current committee members, the ones whose term expires, and the broad community of stakeholders.

3. IMPORTANCE OF COMMITTEE ACTIVITIES AND MEMBERSHIP.

3a. Positive message from the APA National Office.

The quotes from Amy Ferrer that appear earlier in this note all come from her May 14, 2019, email in response to Marcello Guarini's question about the future of the newsletter and the Barwise Prize. It was addressed to Marcello in his capacity as the chair with me CC'd as the vice-chair. Thus, in my capacity as the current chair, I think it is my call to share its important parts with the committee and other stakeholders, even more so since the letter sends a positive message and opens up the space for further productive collaboration.

Hence, whatever the feelings of some of the stakeholders associated with the committee, I am willing to argue that the APA should still be viewed as a reliable partner, and to some degree potential home for several of the initiatives related to *philosophy and computers* that are currently carried on by this committee.

3b. Membership and the stakeholders.

Starting today, only Jack, Robin, Susan, Daniel, Gary, and I are official members of this committee. This is a small group of people. Hence, while we do not expect extraordinary feats from anybody—including the chair of this committee—we must expect of all the members to fulfill their duties. As I have learned last week, from an excellent online training for committee chairs, all members have an active duty to work for the committee, to a reasonable extent, and the committees have the option to ask the APA to replace those failing to do so.

Second, it is customary to ask active members of the cohort whose terms just expired to stay on the email list and continue with their ongoing tasks—in our predicament it is very important since this cohort is not being replaced. We should all volunteer what we are going to do for the committee and follow up on those promises.

4. SUMMARY—PREPARING FOR THE NOT-SO-DISTANT FUTURE.

First, I think, we want to build a common vision on what "we" are going to be after June 30, 2020—and whether "we" want to be anything, as a group. The committee needs to organize a dialogue with all the stakeholders this committee serves, such as past and present committee members, the Barwise Prize winners, participants in our sessions, authors publishing in the newsletter, readers,

audiences, and many others. We need to see if they care to continue the workings of this committee, in a new venue, and what the follow-up activities would be. And if there is no interest, then, well, we would have done our due diligence and move on with our lives.

One final clarification, this is the note of the incoming chair, with my personal opinions and proposed projects. It has to be submitted on the first day of my term as the chair to the publication schedule at the APA. It will be consulted with the committee throughout the fall, and I am sure many improvements shall be made to those plans and ideas. But we need an action plan swiftly, and here is a draft. It is informed by my various roles on this committee for the last fifteen years, which may be an asset, but also a hindrance in designing truly new things. Hence the need for all committee members, as well as all the stakeholders for whom this committee operates, to address their visions, initiatives, and productive concerns.

Sincerely,

Piotr (Peter) Boltuc
 Professor of Philosophy and Associated Faculty of Computer Science, University of Illinois Springfield; and University Professor of Online Learning, Warsaw School of Economics

NOTES

1. <https://www.apaonline.org/news/388037/Changes-to-APA-Committees.htm>
2. M. Croy, "From the Chair," *APA Newsletter on Philosophy and Computers* 3, no. 2 (2004): 2. Available at <https://cdn.ymaws.com/www.apaonline.org/resource/collection/EADE8D52-8D02-4136-9A2A-729368501E43/ComputersV03n2.pdf>
3. <https://www.apaonline.org/members/group.aspx?id=110436>
4. <https://www.apaonline.org/members/group.aspx?id=110436>
5. Those are the titles and expectations as of July 2019. All of our proposals have been accepted. The list of the actual session titles, participants, and in most cases days and times of the sessions have been listed in the note from the editor and the announcements that appear at the end of this issue.

FEATURED ARTICLES

How Many Thoughts Can Fit in the Form of A Proposition?

S. G. Sterrett
 WICHITA STATE UNIVERSITY

1. INTRODUCTION AND OVERVIEW

Let us agree on this much: people use sentences to communicate. On the view that sometimes it is thoughts that are communicated, then, sentences can be used to communicate thoughts. This was Frege's view. However, sentences are used not only to communicate thoughts, but to do other things as well. And, sometimes, in conversation and in writing in natural language, people rely on more than the sentence itself to communicate a thought. This, too, was Frege's view. The study of language is not the study of logic.

Still, developing a logic can start with the study of language, and progress by clarifying how logic is different from language. Unlike the study of language, logic studies sentences only inasmuch as they are used to communicate thoughts, and logic is about using nothing more than sentences to do so. This approach is sometimes part of an anti-psychologistic program in logic, for these two differences between language and logic involve separating the communication of thoughts from psychological aspects of communication.

It can be a bit misleading to call this approach anti-psychologistic without qualification, though, for it is not against the use of psychology in places other than logic: in fact, it draws attention to the fact that psychology is involved in many cases of human communication. An anti-psychologistic view of logic based upon a conviction that psychology has no place in logic was not the only motivation a mathematician might have for distinguishing logic from natural language. For, by Frege's time, it was becoming clear to many mathematicians that natural language, no matter how well-suited it might be for conversation, prose, and poetry, was not always up to the task of providing a language in which to prove theorems and show how the truth of one thought depends upon the truth of another. Relying on natural language, different mathematicians produced proofs whose conclusions were in conflict. What was needed was a means of referring to thoughts that allowed one to determine how thoughts were related to each other.

The anti-psychologistic aspect of Frege's approach is the conviction that relations between thoughts are not a matter of human psychology of any sort whatsoever, general or individual, and hence that any formal calculus of sentences meant to reflect the relationship between thoughts should not involve psychology. In Frege's writings, there was never any wavering, never the slightest hint of compromise, on this point.

Another, less crucial goal, was that the means of expressing a thought using a sentence in this formal calculus be at the same time a means of communicating thoughts that did not depend upon the contingencies of an individual human's psychology. This latter goal was something aimed at, but even Frege came to realize that there was no guarantee that it could be achieved in every case using the system of logical notation he had developed.

Sentences are used to enable people to grasp thoughts, but there can be cases where a sentence enables some people, but not others, to grasp the thought it expresses. Ultimately, communication of thoughts relies on some common understanding, including "a store of concepts" held in common; Frege wrote that a common store of concepts is handed down from one generation to the next,¹ and that learning a language necessarily involves nonverbal communication. Once the common understanding of relevant concepts is achieved, though, it becomes possible for communication of thoughts to proceed such that the sentence alone communicates the thought, without any reference to particulars such as the context in which it was uttered or particulars of the psychological state of

the speaker or hearer. When, however, this common understanding is lacking, as Frege fretted might be the case with certain concepts such as the mathematical/logical notion of a "course of values" of a function used in his formulation of Basic Law V, Frege did not count on everyone grasping the thought expressed by the sentence. Hence different people might disagree on the truth of a sentence, yet their disagreement not be a matter of one person holding a grasped thought true and the other person holding the same thought false. The disagreement arises because they are not grasping the same thought.

That different people hearing or reading a sentence might not grasp the same thought by it was an unhappy situation that, as time went on, Frege learned to accept as inescapable in practice—in the general case. For the more specific and very urgent issue of giving a foundation for arithmetic, however, he at least began the project with the hope that the necessary common understanding might be achieved among mathematicians, if only the right formalism for expressing the thoughts used in arithmetic might be developed. That was the goal of the *Begriffsschrift* (Concept Script): to develop a formalism for statements that expressed the thoughts needed to prove theorems of arithmetic. It was not until he got down to working out the project in detail in the *Grundgesetze* (*Basic Laws of Arithmetic*) that he hit the little snags he regarded as temporary imperfections that might eventually be perfected.²

The formalism Frege developed in the *Begriffsschrift* is generally regarded as containing the fundamental features of modern symbolic logic—and so as an historically significant breakthrough from previous logics, including subject-predicate logic and Boolean logic. Thus reverential appellations such as "the father of modern logic" are heaped upon Frege. In light of such an accomplishment, many have wondered at Frege's seeming antipathy to formalism in arithmetic, and to his criticisms of Hilbert's formalization of geometry, especially his criticisms of implicit definitions of concepts such as the concept of point. Some sympathetic examinations of Frege's views have already been offered.³ One point on which Frege did strenuously diverge from Hilbert was on the use of (what we would now call) uninterpreted statements as premises in proofs.

The thesis I will put forth in this paper—that, ultimately, Frege came to the view that "ideally, one sentence, one thought; one thought, one sentence"—bears on questions about how to explain and understand Frege's criticisms of formalism in arithmetic and Hilbert's use of uninterpreted statements and implicit definition. These criticisms Frege made of his contemporaries are not ad hoc reactions or piecemeal bits of philosophy, but reflect a more unified view about the relation between sentences and thoughts that was slowly being clarified as he tried to attain his original goals of putting mathematics on a solid foundation. The view about the relation between sentences and thoughts that developed over time was as much a result of frustrations encountered as of initial convictions vindicated, and one that he came to accept as practically necessary over time, rather than one chosen as a starting point.

2. LANGUAGE AND LOGIC

According to Frege, it is the thought that matters—at least to the logician. Sentences express thoughts, he said.⁴ But he also found it frustrating to have to use sentences to communicate thoughts. He lamented that he could not put a thought in the hands of his readers as a mineralogist might put a rock-crystal in the hands of audience members. “One fights against language,” he wrote in a footnote to his essay “Thought,” and I am compelled to occupy myself with language although it is not my proper concern here.”⁵ According to him, the logician is concerned only with the thought expressed by the sentence. The thought, however, cannot be handled on its own; it can only be dealt with as wrapped in a linguistic form.

If sentences express thoughts, then what is the problem? Twofold: sometimes—often, in fact—(i) the content of a sentence goes beyond the thought it expresses. Sometimes “the opposite” happens, instead: (ii) the thought expressed goes beyond the content of a sentence; the “mere wording . . . does not suffice for the expression of the thought.”⁶

The examples Frege gave of sentences in natural language in which the content of the sentence goes beyond the thought expressed by it (i.e., of case (i) above) consisted of pairs of sentences that expressed the same thought, although one sentence in the pair was a transformed version of the other. Although the sentences expressed the same thought, one of the sentences had a content that the other did not. An example is the following pair: “Alfred has not come” and “Alfred has not yet come”—the latter sentence differs from the former by the addition of the word “yet” and it creates expectations in the hearer that the first does not. Some other examples of transforming one sentence into another that expresses the same thought but has a different content are the following: (a) replacing “but” for “and” in a sentence, (b) adding “still” or “already” to emphasize part of a sentence, and (c) changing the verb from active to passive and the accusative into the subject.

In all these examples of case (i), the difference between the original and transformed sentences may not be trivial from the standpoint of what the hearer comes to understand or expect upon hearing it, so the sentences may well be said to differ. But, according to Frege’s own remarks on such examples, the original and transformed sentence do express the same thought. This is because these transformations “do not touch the thought, they do not touch what is true or false.”⁷ Logical relationships are relationships between thoughts; the relationships between thoughts in which Frege was interested were relationships of one thought’s dependence on another for justification. So, in the context of an endeavor in which we are concerned only with relationships between thoughts, rather than with expectations created in the hearer, we are not concerned with variations on a given sentence that do not affect its truth value. This point is made by saying that the logician is not concerned with the difference between two such sentences.

“Alfred has still not come,” says Frege, is not false even if Alfred’s arrival is not expected. “Alfred has still not

come” is a different sentence than “Alfred has not come,” but the two sentences express the same thought.⁸ That is because a thought, for Frege, “is something for which the question of truth can arise at all.”⁹ That different sentences of our natural language can express the same thought is no problem for the logician: the logician just doesn’t distinguish between them. In the *Begriffsschrift*, Frege had explained the purpose of his logical symbolism by comparing it to a specialized mechanical aid for seeing: a microscope. The comparison was meant to emphasize that logical symbolism is designed for a special purpose, as is a microscope. Logic is as poor a tool for capturing all the distinctions important to understanding conversation and poetry as is a microscope for viewing a landscape. The point here is that one must be modest about the aims of a particular logical symbolism. Or, rather, one must be clear about its purpose. For purposes of capturing what is relevant to the kind of content of a sentence of interest in doing logic, sentences that express the same thought should not be distinguished.¹⁰ Let’s call this activity—the activity of, for the purposes of logic, identifying sentences that express the same thought—pruning sentences.

What about the opposite situation, i.e., case (ii), when the thought expressed goes beyond the content of a sentence? To Frege, the task is clear: since logic is concerned only with thoughts, we need to augment such a sentence so that the thought determined by the sentence is unique. The kind of sentence that logic is concerned with is the kind that expresses a thought. Case (i) was the case in which different sentences may express the same thought—this the logician tolerates by not distinguishing between the various sentences, and perhaps selecting one as canonical and not using the others—but case (ii) is not so easily accommodated, for it is intolerable that the same sentence should express different thoughts.

Hence, when a sentence contains indexicals (e.g., “I,” “this,” “that,” “yesterday”) or proper names (e.g., Dr. Lauben, Venus) the logician is in trouble if stuck with only the sentence to go on. Thus, Frege says that “The words ‘This tree is covered with green leaves’ are not sufficient by themselves to constitute the expression of a thought, for the time of utterance is involved as well.” He continues: “Only a sentence with the time specification filled out, a sentence complete in every respect, expresses a thought.”¹¹ Let’s call this activity extending a sentence, on analogy with a gardener’s activity in doing the opposite of pruning a small offshoot—here, each of a set of multiple offshoots is encouraged to grow and extend itself to become a distinguishable branch in its own right.

What, then, he asks, should a logician make of a sentence such as “Dr Lauben was wounded”? This sentence expresses different thoughts to different people, depending upon the meaning each associates with “Dr Lauben,” which in turn may depend upon whether they are acquainted with him and what they know about him. Frege gives a long example in which various people associate different definite descriptions with the proper name “Dr. Lauben,” and are in different states of ignorance or knowledge about whether people with whom they are acquainted fit those descriptions. A fellow named Leo and a fellow named

Rudolph both hear Dr. Lauben say aloud, "I was wounded." Later, Rudolph hears Leo report aloud, "Dr. Lauben was wounded." Whether or not the statement made by Dr. Lauben and the statement made by Leo express the same thought to Rudolph is going to depend upon whether or not Rudolph knew that the man he heard saying, "I was wounded" was Dr. Lauben. The point of these examples is that "Dr. Lauben" is a proper name, but there may be different modes of determining the man to whom it refers ("the way that the object so designated is presented"). The logician does need to take these differences into account, Frege says, for different modes of determination for "Dr. Lauben" will result in different thoughts being expressed by the sentence "Dr. Lauben was wounded."¹²

Here, the problem is not a matter of difference in truth value of the different thoughts, for, says Frege, either the thoughts expressed by the sentence are all true or the thoughts expressed by it are all false. The problem is that knowledge of the truth of these thoughts can differ due to different hearers' mode of determination of the person to whom a proper name refers, and this indicates that the thoughts are different. In "Thought," Frege addresses this kind of case—i.e., the kind of case wherein the same sentence can be used to express different thoughts—by adding a restriction on sentences that will be permitted in a logical treatment of any topic involving proper names. The restriction is this: restrict the meaning (sense) of proper names so that no sentence expresses more than one thought. Let's call this kind of activity extending a sentence, too, for, as in the other examples of case (ii), it is analogous to dealing with multiple offshoots by encouraging each offshoot to take its own shape, and so distinguishing each offshoot from each other. However, we do not ever use more than one proper name for an individual—we may have multiple modes of determination that happen to determine the same object, but no proper name has as its meaning more than one mode of determination. In Frege's words: "So we must really stipulate that for every proper name there shall be just one associated manner of presentation of the object so designated. It is often unimportant that this stipulation should be fulfilled, but not always."¹³

3. NATURAL LANGUAGE AND THE FORMAL GARDEN OF PROPOSITIONS

Thus, Frege requires that the sentences of one's natural language that are the concern of logic be in some cases extended (distinguished from each other) and in some cases pruned (identified with each other) so that the relationships that hold between the resulting sentences—sentences the logician can, so to speak, hold in his hand and show to his audience—express the relationships that hold between the thoughts they express. As described in the previous section, sentences that express the same thought are not distinguished from each other (metaphorically, the several branches are pruned down to a single branch). Sentences that do not determine exactly one thought are extended (so that they determine only one thought) or disambiguated such that several sentences, each of which determines exactly one thought, are obtained.

The result is that, for the pile of sentences with which the logician deigns to work, each such sentence expresses exactly one thought, each thought is expressed by exactly one such sentence, and the relation of consequence between such sentences expresses the relation of consequence between the thoughts they express. Of course, this is not true for all the sentences of one's natural language—the point is that it is true of all the sentences the logician is working with after extending and pruning them per the prescriptions just described. Frege eventually came to see such prescriptions as necessary.

We can call the items that result from this process propositions, once they meet such prescriptions; it is irrelevant whether or not the resulting items also happen to be sentences of a natural language. In the *Begriffsschrift*, in explaining the value of the notation he introduced as a replacement for subject-predicate form, Frege said the symbolism he was presenting was a useful tool, if the task of philosophy was to "break the power of words over the human mind" and to free thought "from the taint of ordinary linguistic means of expression."¹⁴

Some readers may take issue with the point just made above, that for the pile of sentences with which the logician deigns to work, each thought is expressed by exactly one sentence, each sentence expresses exactly one thought, and the relation of consequence between sentences expresses the relation of consequence between the thoughts they express. I am well aware that not everyone who has encountered Frege's writings has the impression that Frege avoids the situation wherein a thought is expressed by more than one sentence. Nevertheless this is what Frege says in "Thought." He wrote "Thought" over twenty-five years after writing the much-emphasized and more widely studied "On Sense and Reference" and almost forty years¹⁵ after the publication of *Begriffsschrift*, the work in which he introduced the formalism suitable for doing arithmetic in a "calculus of pure thought." In the *Begriffsschrift* (which predated a distinction he later drew between sense and reference), he did begin to lay out a view that was later revised. As I see it, the vision and ideal he had are not rejected, but rather are better realized, in the view he later laid out in "Thought." In "Thought" he explains more fully, and with examples, the process that I have referred to as the extending and pruning of sentences in the natural language required to obtain the kind of propositions that are fitting for the study of logic.

There's a similar progression in Frege's work concerning his attitude towards the relation between sentences and the thoughts they express. Frege's break with the traditional subject-predicate form of his predecessors, which he discusses in the *Begriffsschrift*, is accompanied by the statement in that early work that this break with tradition is warranted, "that logic hitherto has always followed ordinary language and grammar too closely."¹⁶ In the much later "Thought," Frege writes that although he is not in the "happy position" of the mineralogist who can exhibit the gem he is talking about, he is resolved to a kind of resentful contentment: "Something in itself not perceptible by sense, the thought, is presented to the reader—and I must be content with that—wrapped up in a perceptible linguistic

form." It is not, however, a totally peaceful contentment: "The pictorial aspect of language presents difficulties. The sensible always breaks in and makes expressions pictorial and so improper."¹⁷ The contentment he has achieved is the serenity of accepting what he cannot change.

Frege's explanation of this point—that there are differences between the linguistic forms one needs in natural language (where sentences have additional functions not relevant to logic, such as the function of generating expectations in a hearer that enable conversations to be carried on effectively, and the function of generating ideational associations), and the logical forms one needs to establish the truths of arithmetic—also illuminates his critique of Hilbert. For once one sees the view he expresses in "Thought" about the relationship between sentences and the thoughts they express as the view he was in the process of working towards when he responded to Hilbert's *Foundations of Geometry*, Frege's response to Hilbert's formalization of geometrical axioms seems quite natural.

Hilbert's axioms of geometry were (what we would call) uninterpreted: they were neither true nor false, until they received an interpretation. Frege's complaint was that the notion of an interpretation of a proposition was fundamentally incompatible with the notion of proposition required to do logic. It's easy to see why he thought so: logical relations hold between thoughts. A proposition—the kind of extended and pruned sentence logicians deal with—expresses a thought, and only one thought. On this view, the notion of interpretation has no place in logic.

In his correspondence with Hilbert, Frege wrote that "one feels the broad, imperspicuous and imprecise character of word language to be an obstacle, and to remedy this, one creates a sign language in which the investigation can be conducted in a more perspicuous way and with more precision." He used a slightly different horticultural metaphor, the process of lignification, to illustrate a point about symbolism: Instituting a new symbolism is like the tree's new growth hardening—after it has had a chance to take on the shape appropriate to performing its function.

Then, additional new growth depends upon those hardened sections to support the delivery of nutrients to the newly forming branch tips. The point is that trees do not grow into a predetermined suit of armor made of bark. The rigidity provided by the bark comes only after new branch tips have had a chance to grow in a natural formation. The sign language of a science is not set independently of inquiring as to what signs are best suited to it; if developed appropriately, these signs can be used to hook imperceptible thoughts and wrap them in a perceptible form so there is something that can be held in one's hand, so to speak, and worked with. Signs always involve a compromise compared to what one wishes to communicate, for, after all, signs are perceptible and the thoughts they express are not. It is fundamental to Frege's view that having the right formalism available is important to being able to capture the kinds of imperceptible thoughts in which one is interested. Frege's remark to Hilbert that the need for symbolism comes first, and only later the satisfaction of that need, reflects this conviction.

In correspondence, Hilbert expressed agreement with this last statement.¹⁸

Hilbert used axioms as implicit definitions of the concepts contained in them, though, and Frege didn't like that any more than he liked the fact that Hilbert's axioms required interpretation in order to express a thought. However, as critical as Frege might have seemed of Hilbert, he did evaluate Hilbert's formalization of geometry with the idea of showing how one might achieve what Hilbert was after in a proper manner.

In fact, he outlined a way to make sense of Hilbert's method of showing axioms independent of each other. Frege's reconstruction of Hilbert's independence proofs, however, only work for (what Frege called) real propositions, which Hilbert's axioms were not.¹⁹

Frege's method works as follows: one maps ("set(s) up a correspondence between") words of a language (in which, of course, the reference of every word is fully determinate) onto other words of the same language, subject to some restrictions. These restrictions include mapping proper names to proper names, concept-words to concept-words of the same level, and so on. The signs whose references belong to logic (e.g., negation, identity, subsumption, and subordination of concepts) are not mapped to different signs. Then, one can show that a thought G is independent of a group of thoughts β , if one can obtain from β and G , respectively, a map to a group of true thoughts β' and a false thought G' . In Sterrett 1994 I argued that this was in fact somewhat like the approach Hilbert actually took, and so it was striking that Frege distinguishes his method from methods that employ interpretations of statements. The significance of the difference between Hilbert and Frege, I concluded there, had to do with differences in their accounts of how words come to mean what they do. I will not repeat that discussion here, as it is readily available elsewhere.²⁰

It should be clear by now that Frege is not drawing a distinction between referring to a thought and referring to the perceptible linguistic form in which the thought is wrapped. Frege's point was that the only way he's got to show anyone what thought he is referring to is by wrapping it in a perceptible linguistic form. Hence in talking of the thought G' to which G is mapped (via the mapping of words of a language as outlined above), Frege can hardly be talking about making substitutions of words in, and obtaining transformations of, anything other than sentences. Not just any old sentences of a natural language, however. These sentences or propositions are the result of extending and pruning sentences of the natural language so that each proposition expresses one and only one thought, and so that propositions that express the same thought are not distinguished from each other. I use the term "proposition" here because Frege isn't including all sentences of natural language. He doesn't talk about the forms he has to wrap thoughts in other than as the forms in which the thoughts are wrapped; these forms are not self-subsistent. He was certainly against the idea of developing symbolic forms first and then looking for thoughts that might fit into them. And I don't think he ever meant to talk about these symbolic forms other than as used to express thoughts.

Thus, for Frege, the notion of logical consequence arises for relationships between the imperceptible thoughts that are wrapped up in perceptible forms, not to the forms of the wrappers themselves. One cannot communicate thoughts except by capturing them in such a perceptible wrapping, so proofs and derivations proceed by way of rules that apply to propositions or statements. However, these propositions always express a thought: they are never empty wrappers. They are not in need of interpretation.

4. DEPARTED THOUGHTS

Hence, Frege says that if by sentence is meant the “external, audible, or visible that is supposed to express a thought,” then it does not make any sense to say that one sentence is independent of another. The context in which Frege wrote this was in arguing that Hilbert had erred in the specific way he had gone about trying to establish the independence of the parallel postulate from the other axioms of geometry. It was in this context that Frege said that Hilbert makes a mistake in calling anything “the axiom of parallels,” for, as Frege put it in the passage quoted above, it is not the same in every geometry: “Only the wording is the same; the thought-content is different in each particular geometry.”²¹

Frege means here to warn against mistaking the “external, audible, or visible that is supposed to express a thought” for the thought. Logic is concerned with thoughts and how they are related to each other. So there is a realm of thought: it cannot be perceived by the senses, but it is like perceptible things in that it does not need an owner, as ideas do.²² Frege’s favorite example of a thought in his essay entitled “Thought” is the Pythagorean theorem. Different people can grasp the thought, and it can be communicated by wrapping it in a perceptible linguistic form.

But I don’t think Frege intends to alert the reader to the existence of a logical calculus of “the external, audible, or visible that is supposed to express a thought.” This would be a study of the relationships of linguistic forms, something Frege thought of interest for many purposes—understanding conversations and writing poetry, for instance—but decidedly not the subject matter of logic. Logic is about thoughts, it is about the laws of thought, the laws of the laws of science. It is about deriving proofs so that we can see how one thought depends upon another. It involves the linguistic forms in which these thoughts must be wrapped in order to be communicated, but only in the context of investigating which thoughts depend upon which other thoughts. It is not about relationships of dependence between perceptible linguistic forms. If there are such things as forms that exist as shed snakeskins left behind from departed thoughts, they are not the concern of logic; they are not the items of a calculus of pure thought.

For Frege, there is no such thing as a realm of linguistic forms within which no thoughts are wrapped but which are related to each other in virtue of their form by logical laws. The logical relations are not logical relations between linguistic forms.

5. THE UNITY OF THOUGHT AND EXPRESSION

Frege did discuss examples of different sentences that expressed the same thought, even in “Thought,” arguing

that “the content of a sentence often goes beyond the thought expressed in it.” But his response to this observation was not to posit a new kind of logical law or a new kind of logical relation to account for how such sentences were related. In “Thought,” he did not regard such situations as puzzles; he did not then consider them relevant to logic. Rather, his response to this observation about natural language was that the logician does not distinguish between such sentences.

Was Frege this blasé about different sentences that express the same thought because, on a view sometimes attributed to Frege, he thought that there are really two distinct things, sentences and thoughts, and thus that the distinction between sentences is a distinction that can be made only in the realm of what is derivable, and not in the realm of what is provable? I don’t think that this is how Frege’s views on sentences that express the same thought ought to be viewed.

Recall that what Frege said about pairs of distinct sentences that express the same thought was only that some such transformations between sentences must be recognized as admissible. But this wasn’t a matter of recognizing relationships that obtain in a realm of equipollent propositions. In his 1906 letter to Husserl, in fact, Frege suggested that equipollent propositions could all be communicated by a single standard proposition.²³ In closing the letter, he remarks that the question of whether equipollent propositions are congruent “could well be debated for a hundred years or more.” But he isn’t concerned about the answer; he writes, “I do not see what criterion would allow us to decide this question objectively. . . . But I do find that if there is no objective criterion for answering a question, then the question has no place at all in science.”²⁴ Placing significance on the difference in the relations that hold between sentences and the relations that hold between thoughts is attributing significance to exactly what, I think, he actually said ought to be de-emphasized.

We have seen that what Frege said about sentence transformations that do not affect the thought expressed was that sentences with differences that don’t affect the thought expressed don’t need to be distinguished when doing logic. All that the existence of transformations that yield two or more sentences expressing the same thought means to the logician is that, if propositions or statements admit of such transformations, one must recognize as admissible those transformations that do not affect the thought expressed. Once we see this point of Frege’s, the apparition of the notion of derivability according to which things are not always as they seem disappears: i.e., the notion of derivability on which a thought when wrapped in a different wrapper might have different derivability relations disappears. There is a realm of thoughts (thoughts are not the property of individuals as ideas are, but they are not perceptible either), and it is distinct from the realm of perceptible things.²⁵ Logical laws are used in showing the relationships that exist between thoughts, via a proof. Thus, the realm in which logical rules apply involves both of these realms, since it includes both thoughts and signs; and this in turn is due to the unavoidable situation that

communication of thoughts requires that thoughts be wrapped in perceptible forms.

What about Frege's statement that a thought can be "carved up" in different ways? Doesn't the fact that the same thought could be carved up in different ways mean that the same thought could be expressed by different sentences? Yes and no—the difference being a matter of which language you are talking about. In natural language: Yes, the same thought can be expressed by different sentences that analyze the thought into subject and predicate differently; typically this will happen whenever the same sentence is transformed from the active to the passive voice. But in the formal language of the *Begriffsschrift*, the answer is no: the carving really captures the structure of the thought relevant to the kind of inferences one wants to be able to draw. That is, the whole point of the *Begriffsschrift* was that subject-predicate logic did not get at the structure of thought relevant to making inferences! In contrast, the formalism of the *Begriffsschrift* was created to ensure that all of the structure relevant to making inferences that were a matter of pure logic could be expressed.

The point that the situation of having only subject-predicate logic available is restrictive in spite of allowing many options might be explained using the metaphor of a plant, as follows: that situation is like having only a certain kind of analysis of the plant available to you, for instance, having only the option of describing a plant in terms of dividing it up into the edible food it bears and the part of the plant that produces the edible food. What's limiting about this is not a matter of how many ways there are to carve up the plant, for in fact the edible-food and plant-that-produces-food way of carving up a plant permits many different ways of carving up the plant. Depending upon what part of the plant a creature is interested in consuming, one could analyze the plant into an edible product and the remainder of the plant that produces it in different ways, just as the subject-predicate form allows one to express a thought in different ways depending upon what one chooses as the subject of the sentence. Rather (using the plant metaphor) the limitation is this: the available ways of analyzing the plant does not necessarily allow us to analyze the plant structure in the way required for investigations in natural science.

Analysis of a plant based on edible parts of the plant does exhibit something about the structure of the plant, of course, but it also obscures some of the structure of the plant. What we want is a general method of carving up the plant in a way that allows the flexibility and precision to exhibit various kinds of structure in the plant, a way that permits the many different kinds of carving ups of the plant needed for making inferences we want to draw to conduct research about a variety of questions that interest us.

On this analogy, what's wrong with subject-predicate logic is that the kinds of "carving up" of a thought it permits—and it may permit a number of alternatives—might not include the structure of the thought that is relevant to making the kinds of inferences in which one is interested. In contrast, the formalism of the *Begriffsschrift*, in which concepts are modeled on functions, is meant to introduce

a formal language in which one can carve a thought in any way needed for making scientific inferences. The formalism provided in that work is supposed to be enough to permit making any inferences that are a matter of pure logic. This is not to say that the kind of structure sought for even when using the formalism of the *Begriffsschrift* may not be relative to the kinds of inferences one is interested in making (hence the formalism needed for chemistry and physics is left open in the *Begriffsschrift*; in my biological metaphor, the added formalism needed to carve the plant into its relevant parts might be the gene concept). It is to say that the formal language does not, as subject-predicate logic does, limit one to carving the plant into two parts according to a criterion that may never permit one to delineate the structure of the plant relevant to the inference in which one is interested.

The advance Frege offered was not a way of dissecting a thought into formalism and unformed thought, but, rather, consisted in a formalism that permitted carving thoughts in more useful ways than previous formalisms allowed. The separation of thought from sentence underlying the distinction between provability and derivability is not something we find in Frege. To describe such a disconnect as part of Frege's view misdescribes Frege's notion of a proposition in the same way that Aristotle's notion of form would be misdescribed by using Plato's notion of form. That is, in Plato's philosophy, forms exist in a realm separate from the things of which they are forms. Aristotle, too, used a metaphor from biology to break from Plato: that there are male and female animals, he said, does not imply that male and female exists as something separable from male and female animals.

To use another metaphor: in a certain science fiction television series, there is a creature that can transform itself into various shapes, called a shape-shifter. These shape-shifters can separate from their shapes and meld together somehow in a realm in which they are shapeless. But this is, after all, fiction. To make the metaphor of shape-shifters who take on various shapes fit Frege's account of sentences as the forms within which thoughts are wrapped, let us leave the details of this particular science fiction story behind and stipulate that shape-shifters take on human forms, that a given shape-shifter cannot take on every form, and, in fact, that the forms a particular shape-shifter takes on are not taken on by any other creature. (This corresponds to Frege's requirements that, in his formalism, thoughts are expressed by sentences, that more than one thought is expressible, and that no sentence expresses more than one thought.) Clearly, once we've figured out the shapes between which a particular shape-shifter can transform itself, we no longer need distinguish between those shapes.

The analogy to thoughts and the linguistic forms they take on is this: just as, in the science fiction story, a creature is apprehended via the senses by its sensible form, so a thought is expressed via a sentence. That, in effect, is Frege's unperturbed response in his essay "Thought" to the examples in which there are several sentences that express the same thought, such as two sentences that differ only in the manner used to designate an object. That is, in contrast

to the view that Frege is saying that there are two different calculi, one for thoughts and one for the linguistic forms in which they can be wrapped, Frege shows that he intends to avoid such commitments by stipulating that, when proper names are used, only one manner of presentation (e.g., for Venus, either “the morning star” or “the evening star,” but not both) be permitted. Thus I do not think that, as is often supposed, Frege developed a calculus of sentences associated with something called derivability in addition to the calculus of thoughts associated with provability. His remark to Husserl (quoted earlier) that he does not think there is room in science for the question of whether equipollent propositions are congruent bears this out.

Looking back from the present, some people attribute to Frege’s *Begriffsschrift* the achievement of having developed a calculus of sentences related by derivability, which are accurately described in modern parlance as syntactic relations. This is not so, and Frege is explicit enough about what he was doing to make that clear. Frege’s *Begriffsschrift* was to be a calculus of thoughts. There were reasons that the calculus had to involve symbolic formalism—to clarify thoughts, and to express them—but the calculus was not a “topic-neutral” calculus of symbolic or syntactic forms. That may be what a modern logician sees in looking at the *Begriffsschrift*, but it doesn’t sound much like Frege’s description of the *Begriffsschrift*. What it does sound like, however, is Frege’s description of Leibniz’s vision, which, he said, “was too grandiose for the attempt to realize it to go further than the bare preliminaries.” (in Beany, p. 50)²⁶ Frege thought Leibniz’s vision of a universal calculus an excellent guiding vision, but what he said about his own achievement in the *Begriffsschrift* with respect to Leibniz’s visionary aim was that “even if this great aim cannot be achieved at the first attempt, one need not despair of a slow, step by step approach.” The project, Frege said, “has to be limited provisionally” at first. And he identified the *Begriffsschrift* as one of the “realizations of the Leibnizian conception in particular fields.”²⁷ He spoke of additions that would have to be made to extend it to geometry and then to the pure theory of motion, then mechanics, and then physics. These latter fields involve natural necessity as well as conceptual necessity.

In his correspondence with Hilbert, Frege writes that he thinks Hilbert is (mistakenly) treating geometry as if it were like arithmetic. Frege thought it an error to regard geometrical knowledge as having the same kind of basis as arithmetical knowledge. This is important, for it meant that Frege didn’t think sentences or propositions of geometry were related to each other in the same way that statements of arithmetic were. The *Begriffsschrift* was to help in showing that arithmetical truths were truths of logic, but even this does not mean that the rules in the *Begriffsschrift* applied to topic-neutral sentences, for Frege did not take a formalist approach to arithmetic either. What I mean by this is that he did not allow (what we would now call) uninterpreted statements of arithmetic any more than he did statements of geometry. In the *Basic Laws of Arithmetic*, he reiterates his requirement on axioms, i.e., that all the terms in them must be defined. That he is not always able to meet the requirement should not be cited as evidence that some of the concepts are implicitly defined

or are uninterpreted and to be interpreted at a later date. Rather, Frege’s explanation of such undefined concepts is found in a statement he makes in preliminary remarks in the *Basic Laws*: “It will not always be possible to give a regular definition of everything, precisely because our endeavor must be to trace our way back to what is logically simple, which as such is not properly definable. I must then be satisfied with indicating what I intend by means of hints.”²⁸ The principles of the *Begriffsschrift* may apply to every science, but according to Frege they do not include all the principles nor, even, all the formalism needed to do geometry, kinematics, physics, or chemistry. These await future development, he said.

Thus, we must avoid the anachronism of splitting asunder a propositional form from a thought. For Frege, a proposition is a thought wrapped in a perceptible linguistic form, i.e., a propositional form. The perceptible linguistic form it is possible to wrap a thought in may not be uniquely determined for a given thought, but the thought must be wrapped in some perceptible linguistic form or other. Hence the proposition cannot survive such a dissection. Even in developing a calculus in which the ideal is “one proposition, one thought; one thought, one proposition,” a thought and its expression are not split apart. Throughout his correspondence with Hilbert, Frege seems concerned to speak of the proposition as a whole, i.e., a “real” proposition expressing a thought. The kind of axioms Hilbert proposed, which were neither true nor false, and so which, on Frege’s view, did not express thoughts, were not, on his view, proper subjects of logic.

On Frege’s view, a thought is necessarily wrapped in linguistic form if it is to be communicated, studied, or used in reasoning. Thoughts are individuals for Frege, somewhat as trees and humans were individuals for Aristotle. Aristotle was concerned (at least in some of his works) to hold out for the identity of an individual in spite of the different things that could be predicated of it, but in a way that didn’t call for dissecting that individual into a self-subsistent form and something else. Similarly, what Frege thought was called for with respect to thoughts was a method of expressing an individual thought that exhibited the structure of the thought in such a way that we could see its relation to other thoughts, but in a way that didn’t call for dissecting it into a self-subsistent linguistic part and something else. Frege also seemed to recognize different kinds of relations between thoughts, that the relations that were crucial might be different for different investigations and different disciplines. The *Begriffsschrift* was meant to provide a calculus in which to express thoughts that met the needs of the discipline of logic, i.e., a calculus in which the logical relations between thoughts would be exhibited. Frege continually warned against the tendencies of some of his contemporaries to take the approach of attempting to separate the propositional form of a proposition from the thought it expresses and treat it as self-subsistent. The admonition to refrain from attempting such fatal dissections, though, is quite general. It is as old as Aristotle and as new as post-analytic philosophy.

NOTES

1. G. Frege, *Philosophical and Mathematical Correspondence*, 59.
2. Here I am referring to Frege's remark in the *Grundgesetze*: "A dispute can arise, so far as I can see, only with regard to my Basic Law concerning courses-of-values (V), which logicians perhaps have not yet expressly enunciated, and yet is what people have in mind, for example, where they speak of the extensions of concepts." Frege, *The Basic Laws of Arithmetic. Exposition of the System*, 3-4. Here he does express confidence that the concept of course of values might be enunciated more clearly, and that, when it is, disputes about Basic Law V will be settled. His attitude towards the very different kind of problem later pointed out by Bertrand Russell was not one of confidence in overcoming it, and I am not referring to Russell's paradox when speaking of "little snags." Frege addressed Russell's paradox in Appendix II to volume II of the *Grundgesetze*.
3. M. Resnik, "The Frege-Hilbert Controversy"; S. G. Sterrett, "Frege and Hilbert on the Foundations of Geometry"; P. Blanchette, "Frege and Hilbert on Consistency"; A. Antonelli and R. May, "Frege's New Science"; J. Tappenden, "Frege on Axioms, Indirect Proof, and Independence Arguments in Geometry: Did Frege Reject Independence Arguments?"
4. Frege, "Thought," 328.
5. *Ibid.*, 329-30.
6. *Ibid.*, 331.
7. *Ibid.*
8. *Ibid.*
9. M. Beany, *The Frege Reader*, 328.
10. In a 1906 letter to Husserl, Frege wrote that, while it is not possible to say exactly when two propositions are merely equipollent and when they are congruent, this is not an obstacle in principle: "All that would be needed would be a single standard proposition for each system of equipollent propositions, and any thought could be communicated by such a standard proposition. For given a standard proposition everyone would have the whole system of equipollent propositions, and he could make the transition to any one of them whose illumination was particularly to his taste." Beany, *The Frege Reader*, 303.
11. Frege, "Thought," 343.
12. *Ibid.*, 333.
13. *Ibid.*
14. Beany, *The Frege Reader*, 51.
15. The *Begriffsschrift* (Concept-Script) was published in 1879; *Grundlagen der Arithmetik* (Foundations of Arithmetic) in 1884, "On Sense and Reference" in 1892, and "Thought" in 1918. "On the Foundations of Geometry" and associated correspondence with Hilbert and others was written around 1900.
16. Beany, *The Frege Reader*, 51.
17. Frege, "Thought," 334.
18. Hilbert 4.10.1895 in Frege, *Philosophical and Mathematical Correspondence*, 34.
19. Sterrett, "Frege and Hilbert on the Foundations of Geometry."
20. Giving a brief description of the contrast between Frege's account of elucidation and Hilbert's account of implicit definition risks mischaracterizing Hilbert as more formalist than he was, so I refer the reader to my discussion in Sterrett, "Frege and Hilbert on the Foundations of Geometry" in which I distinguish the positions of Hilbert, Korselt (who responded to Frege on Hilbert's behalf), and Frege. The paper is available free online at the Philosophy of Science Archives server, at <http://philsci-archive.pitt.edu/723/>.
21. Frege's analysis is that the fault lies in confounding first- and second-level concepts, such as the concept of point. There may be different first-level concepts of point, under which points fall: the Euclidean point-concept is one such first-level concept. If one likes, one may also define a second-level concept, within which the Euclidean point-concept and other first-level concepts, fall. A fuller discussion of Frege's point is given in Sterrett, "Frege and Hilbert on the Foundations of Geometry," 9.

22. Frege, "Thought," 337.
23. Beany, *The Frege Reader*, 302.
24. *Ibid.*, 305.
25. Frege, "Thought," 337.
26. Beany, *The Frege Reader*, 50.
27. *Ibid.*
28. Frege, *The Basic Laws of Arithmetic. Exposition of the System*, 32.

REFERENCES

Antonelli, Aldo, and Robert May. "Frege's New Science." *Notre Dame Journal of Formal Logic* 41, no. 3 (2000): 242–70..

Beany, Michael, ed. *The Frege Reader*. Oxford: Blackwell Publishers, 1997.

Blanchette, Patricia. "Frege and Hilbert on Consistency." *Journal of Philosophy* 93 (1996): 317–36.

Frege, Gottlob. "Frege to Hilbert 6.1.1900." 1900. In Frege (1980), 43–48.

———. "Thought." 1918. In Beany (1997), 325–45.

———. *The Basic Laws of Arithmetic. Exposition of the System*. Translated and edited, with an introduction, by Montgomery Furth. Berkeley: University of California Press, 1964.

———. *Philosophical and Mathematical Correspondence*. Edited by Brian McGuinness and translated by Hans Kaal. Chicago: University of Chicago Press, 1980.

Resnik, Michael. "The Frege-Hilbert Controversy." *Philosophy and Phenomenological Research* 34 (1994): 386–403.

Ricketts, Thomas. "Frege's 1906 Foray into Metalogic." *Philosophical Topics* 25 (1997): 169–88.

Sterrett, S. G. "Frege and Hilbert on the Foundations of Geometry." 1994 Talk. Available on the Philosophy of Science e-print archives at <http://philsci-archive.pitt.edu/723>.

Tappenden, Jamie. "Frege on Axioms, Indirect Proof, and Independence Arguments in Geometry: Did Frege Reject Independence Arguments?" *Notre Dame Journal of Formal Logic* 41 (2001): 271–315.

Consciousness, Engineering, and Anthropomorphism

Ricardo Sanz

UNIVERSIDAD POLITÉCNICA DE MADRID

1. INTRODUCTION

The construction of conscious machines seems to be central to the old, core dream of the artificial intelligence community. It may well be a maximal challenge motivated by the pure *hybris* of builders playing God's role in creating new beings. It may also just be a challenging target to fuel researchers' motivation. However, we may be deeply puzzled concerning the reasons for engineers to pursue such an objective. Why do engineers want conscious machines? I am not saying that engineers are free from *hybris* or not in need for motivation, but I question if there is an *engineering reason* to do so.

In this article I will try to analyze such motives to discover these reasons and, in this process, reveal the excessive anthropomorphism that permeates this endeavor. Anthropomorphism is an easy trap, especially for philosophers. We can see it pervasively tinting the philosophy of consciousness. However, in the modest opinion of this engineer, philosophy shall transcend

humanism and focus on universal issues of value both for animals and machines.

2. THE ENGINEERING STANCE

The construction of intelligent machines is the central activity of control systems engineering. In fact, the core focus of activity of the control systems engineer is the design and implementation of *minds for machines*. For most people involved in cognitive science, saying that a PID controller is a mind is not just an overstatement; it is, simply, false. False because such a system lacks emotion, education, growth, learning, genealogy, personality . . . whatever.

This analysis of what minds are suffers from biological chauvinism. Anthropomorphism is pervasive in cognitive science, artificial intelligence, and robotics. This is understandable for historical reasons but shall be factored-out in the search or core mechanisms of mind.

A central principle of engineering is that systems shall include what is needed and only what is needed. Over-engineered systems are too complex, late delivered, and uneconomical. This principle shall also be applied to the endeavor of building conscious machines.

2.1 THE SYSTEMS ENGINEERING VEE

The systems engineering lifecycle (see Figure 1) starts with the specification of needs: what does the user of the system need from it. This need is stated in the form of a collection of user and system requirements.¹ The verification of satisfaction of these needs—the system validation for acceptance testing—is the final stage of the engineering life-cycle.

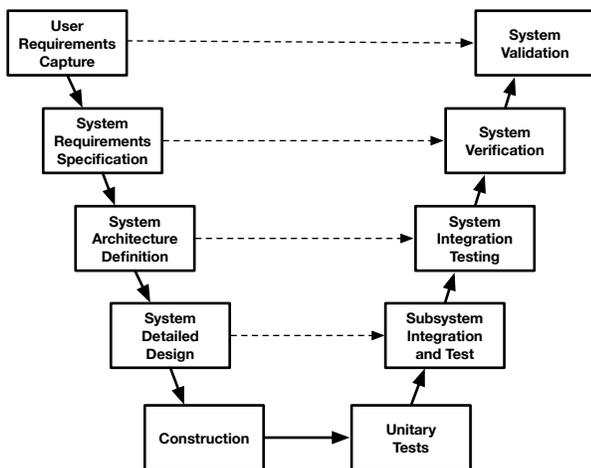


Figure 1. The Systems Engineering (SE) Vee. This flowgraph describes the stages of system development as correlated activities oriented to the satisfaction of user needs.

This process implies that all system elements—what is built at the construction stage—do always address a user or system need; they always have a function to perform. The concrete functions that are needed will depend on the type of system and we shall be aware of the simple fact that *not all systems are robots*. Or, to be more specific, not

all intelligent systems are humanoid robots. Many times, intelligent *minds* are built for other kinds of systems. Intelligence is deployed in the sophisticated controllers that are needed to endow machines with the capability to address complex tasks. Minds are just control systems.² Intelligent minds are sophisticated control systems.³

2.2 NOT ALL AI SYSTEMS ARE ROBOTS

The obvious fact that not all AI systems are humanoid robots has important implications. The first one is that not all systems perform activities usually done by humans and hence:

1. Their realizations—their bodies—do not necessarily resemble human bodies. In engineering, bodies follow functional needs in a very intentional and teleological sense. Machines are artificial in the precise sense clarified by Simon.⁴
2. Their environments—the context where they perform the activity—are not human environments and fitness imply non-humanly capabilities.
3. Their missions—what are they built for—are sometimes human missions, but mostly not. People are worried about robots getting our jobs but most robot jobs cannot be performed by humans.

In control systems engineering we usually make the distinction between the controller—the mind—and the plant—the body. This may sound kind of cartesian and indeed it is. But it is not due to a metaphysical stance of control engineers but to the more earthly, common practice of addressing system construction by the integration of separately built parts.⁵

The plant (usually an artefact) can hence be quite close or quite different from humans or from animals:

Airplanes: Share the environment and the activity with birds, but their functional ways are so different from animals that control strategies are totally different.

Industrial Robots: In many cases can do activities that humans could do: welding, picking, packaging, etc. but requirements may be far from human: precision, speed, repeatability, weight, etc.

Vehicles: Autonomous vehicles share activity with animals: movement. However they are steadily departing from animal contexts and capabilities. Consider, for example, the use of GPS for autonomous driving or vehicle-to-infrastructure communication for augmented efficiency.

Chemical Plants: Some artefacts are extremely different from humans seen as autonomous entities moving in environments. Industrial continuous processes—chemical, oil, food—do not resemble humans nor animals and the needs for intelligence and awareness are hence quite different.

Utilities: The same can be said for technical infrastructure. The intelligence of the smart grid is not close to animal intelligence.

All these systems “live” in dynamic contexts and their controllers shall react appropriately to changing environmental conditions. They process sensory signals to be “aware” of relevant changes, but they do it in very different ways. Machines are not animals nor in their realization nor in their teleology. Bioinspiration can help systems engineers in the provision of architecting ideas of concrete designs of subsystems. However, mapping the whole iguana to a machine is not a sound engineering strategy.⁶

2.3 THE AI PROGRAM VS. THE STANDARD STRATEGY

From my perspective the many threads of the global AI program can be categorized into three basic kinds of motivations:

- Technology. Solving problems by means of incorporating intelligence into the artefacts.
- Science. Explore the nature of (human) intelligence by creating computer models of psychological theories.
- Hubris. Create beings like us.

Control system engineering (CSE) implements AIs because it is interested in the problem-solving capabilities that AI can provide to their machines. AI enters the CSE domain to deal with runtime *problems of higher complexity* that are not easily addressable by more conventional means. The mind of the machine is built as a cognitive agent that perceives and acts on the body that is situated in an environment. In their well-known textbook Russell and Norvig even say that “the concept of a controller in control theory is identical to that of an agent in AI.”⁷

AI-based controllers can decide in real-time about what to do in complex situations to achieve system goals. Goals that are established in terms of user needs and, secondarily, in terms of machine needs. An AI-based controller does not pursue the machine objectives but the objectives of its owner.⁸

The optimal strategy for controlling a system is to invert a perfect model of it.⁹ But this only works to the extent that the model behavior matches system behavior. Model fidelity is limited due to several factors. Observability limits what the intelligent agent may perceive. Note that the intelligent agent interacts with a body that interacts with the environment. In this situation perfect knowledge is unachievable because uncertainty permeating both the plant and its environment affects the intelligent agent mental representation (cf. the problems surrounding the deployment of autonomous cars).

Agent mental complexity shall match that of the plant and its environment following Ashby law of required variety. This implies that the *curse of complexity and uncertainty* affects not only the plant and its environment, but *also the controller itself*. Intelligent, autonomous controllers are enormously complex artefacts.

Complexity plays against system dependability. The probability of failure multiplies with system complexity. This is not good for real-life systems like cars, factories, or gas networks. The basic method to improve dependability is building better systems. Systems of better quality or systems built using better engineering processes (e.g., as is the case of cleanroom engineering). However, these do-well strategies are not easily translatable to the construction of systems of required high complexity.

The *standard strategy* to address runtime problems is to use humans to directly drive or supervise the system. Humans are better at addressing the unexpected and provide augmented robustness and resilience (R&R). A term that is gaining acceptance these days is *Socio-Cyber-Physical System, a system composed of physical bodies, software controllers and humans*. Figure 2 shows a common layering of these systems.

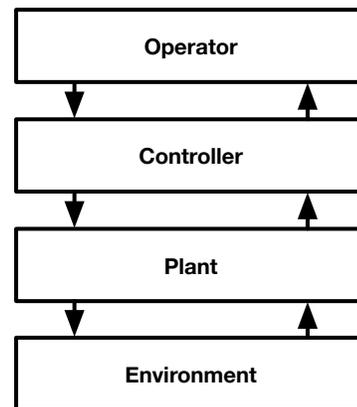


Figure 2. A socio-cyber-physical system is a layered structure of interacting systems. The top authority corresponds to human operators because they are able to deal with higher levels of uncertainty.

In socio-cyber-physical systems the top authority corresponds to human operators because they are able to deal with higher levels of uncertainty. Humans are able to understand better what is going on, especially when unexpected things happen. The world of the unexpected has never been a friendly world for AIs.

3. BUILDING CONSCIOUS MACHINES

The research for consciousness in artificial systems engineering can be aligned with the three motivations described in the previous section—useful technology, psychological science, or mere hubris.

Some authors consider that biological consciousness is just an epiphenomenon. However, an evolutionary psychology dogma states that *any currently active mental trait that has been exposed to evolutionary pressure has adaptive value*. This—in principle—implies that *consciousness has adaptive (behavioral) value*; so it may be useful in machines.

From a technological stance, the analysis/evolution of complex control systems took us into researching novel

strategies to improve system resilience by means of self-awareness. If the system is able to perceive and reason about its own disturbances, it will be able to act upon them and recover mission-oriented function. Machine consciousness enters the engineering agenda as a possible strategy to cope with complexity and uncertainty.

A conscious machine can reflect upon itself and this may be a potential solution to the curse of complexity problem. So, the engineering interest in consciousness is specifically focused on one concrete aspect: self-awareness. This implies that the engineering stance does not have much to say about other aspects of consciousness (esp. qualia).

3.1 SELF-AWARENESS IN MACHINES

Self-aware machines are aware of themselves. Self-awareness is just a particular case of awareness when the object of awareness is the machine itself. Self-awareness is a class of perceptual process, mapping the state of a system—the machine itself—into an exercisable representation.

From an engineering perspective, self-awareness is useless unless it is accompanied by concurrent action processes. In particular, to be of any use concerning system resilience, self-awareness processes need coupled self-action processes. This closes a control loop of the system upon itself.

This may sound enormously challenging and innovative, but this is not new at all. Systems that observe and act upon themselves have been common trade for decades. There are plenty of examples of self-X mechanisms in technical systems that in most cases are not based on biology:

- Fault-tolerant systems (from the 60s)
- Adaptive controllers (from the 70s)
- Metacognitive systems (90s)
- Autonomic Computing (00s)
- Adaptive service systems (00s)
- Organic Computing (10s)

All these systems observe themselves and use these observations to adapt a system’s behavior to changing circumstances. These changes may be due to system-external disturbances or system-internal operational conditions. The adaptation to external changes has been widely investigated, but the adaptation to internal changes has received less attention.

In our own own case we investigate domain-neutral, application-neutral architectures for augmented autonomy based on model-based reflective adaptive controllers. Domain neutral means that we investigate architectures for any kind of system—e.g., mobile robots or chemical factories—and application neutral means that the architectures shall provide functionality for any kind of application—e.g., for system fault tolerance or dynamic service provision.

3.2 AN EXAMPLE: A METACONTROLLER FOR ROBOTS

Figure 3 shows an implementation of a self-aware system that improves resilience of a mobile robot.¹⁰ The self-awareness mechanism is a metacontroller—a controller of a controller—that manages the operational state of the robot. This metacontroller has been designed to mitigate the resilience reduction due to potential faults in the control system of the robot.

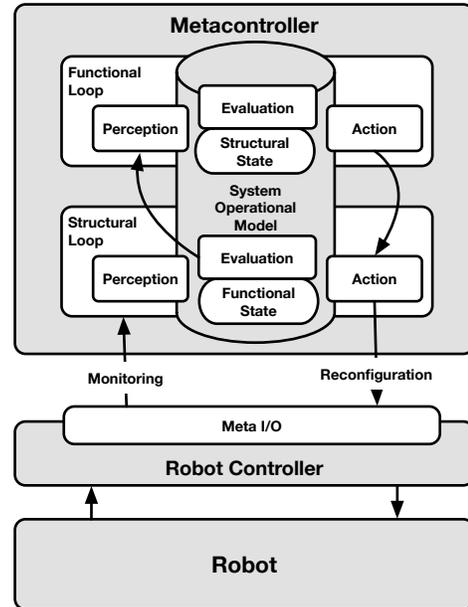


Figure 3. A metacontroller for robots developed for improving adaptivity of the robot controller.

The robot controller is a very complex distributed software system that can suffer transient or permanent faults in any of its components. The metacontroller monitorizes the state of the robot controller and acts upon it to keep system functionality by reorganizing its functional organization. This is similar to what humans do when overcoming some of the problems of becoming blind by learning to read with the fingers.

Function, functional state, and componential organization are core concepts in this approach. In this system this self-awareness mechanism provides reaction to disruption and improves mission-level resilience. And this is grounded in a self-model based on formally specified concepts concerning the system and its mission.

Figure 4 shows part of the formal ontology¹¹ that is used in the implementation of the perception, reasoning, and action mechanisms of the self-awareness engine. It enables the robot to reason about its own body and mind.

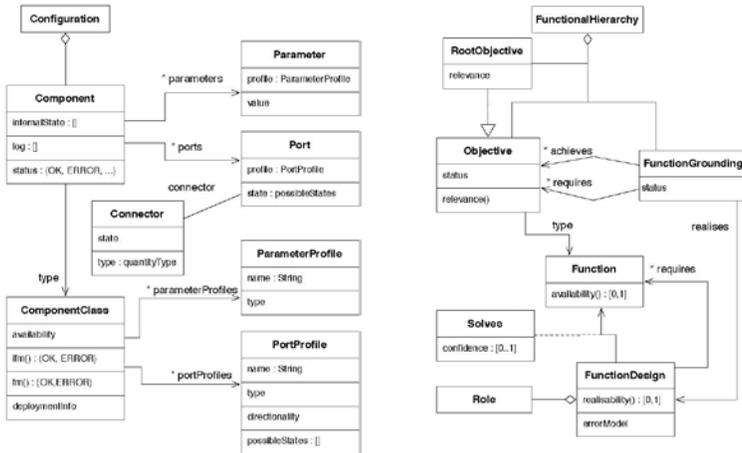


Figure 4. Part of the formal ontology that enables the robot to reason about its own body and mind.

3.3 INTO PHILOSOPHY

This research on machine self-awareness obviously enters philosophical waters.

The attempt to achieve engineering universality—any kind of system, any kind of environment, any kind of mission—implies that the *ontologies* and *architectural design patterns* that support the engineering processes must be general and not particular for the project at hand.

This requires a more scientific/philosophical approach to the conceptual problem that gets harder when human concepts enter into the picture (to provide an easy path for bioinspiration or to address knowledge and control integration in human-cyber systems).

Plenty of questions arise when we try to address self-awareness from a domain neutral perspective (i.e., a systems perspective that is based on concepts that are applicable both to humans and machines): What is Awareness? What is Self? What is Perception? What is Understanding? But these questions will be answered elsewhere because the rest of the article is dedicated to a problem in the science/philosophy of (machine) consciousness: **It is too centered in humans.**

This is somewhat understandable because humans are THE SPECIMENS of consciousness. But this is also a problem in general AI and Robotics, in convergent studies on philosophy of computers, and even in general cognitive science that is neglecting important results in intelligent systems engineering and losing opportunities for ground truth checks.

4. TOO MUCH ANTHROPOMORPHISM

Readers may have heard of Sophia; a humanoid robot who uses AI and human behavioral models to imitate human gestures and facial expressions. Sophia can maintain a simple conversation, answering simple questions. From an AI or robotics standpoint it is quite a low feat. Why does it get so much attention?

Raymond Tallis, in his book *The Explicit Animal*, offers us a clue in the form of a rant against functionalism:

“Traditional mental contents disappear and the mind itself becomes an unremarkable and unspecial site in causal chains and computational procedures that begin before consciousness and extend beyond it. Indeed, mind is scarcely a locus in its own right and certainly does not have its own space. It is a through road (or a small part of one) rather than a dwelling. Consciousness is voided of inwardness.”¹²

For Tallis, causal theories of consciousness reduce mind to a set of input/output relations, with the net effect of effectively “emptying” consciousness. Causal links—the stuff machines are made with—seem not enough for the machinery of mind.

The same phenomenon can be found in any context where human mental traits and “similar” machine traits are put under the scope of the scientist or philosopher. For example, in a recent conference on philosophy of AI, a speaker raised the question “Is attention necessary for visual object classification?” A few slides later the speaker showed that Google was doing this without attention. So the answer of the question was NO. End. Surprisingly, the presentation continued with a long discussion about phenomena of human perception.

4.1 ANTHROPO-X

Cognitive Science and the Philosophy of Mind, and to some extent Artificial Intelligence and Robotics, are anthropocentric, anthroposcoped and anthropobased. They focus on humans; they address mostly humans; they think of humans as special cases of mind and consciousness. Obviously, the human mind ranks quite high in the spectrum of mind. Maybe it is indeed the peak of the scale. But this does not qualify it as special in the same sense that elephants or whales are not special animals, however big.

However, the worst problem is that all these disciplines are also anthropomorphic: They shape all their theories using the human form. Protagoras seems still alive and man is used to measure all things mental.

This is not only wrong, but severely limiting. Anthropomorphism has very bad effects in consciousness research:

- Human consciousness traits are considered general consciousness traits. This has the consequence for artificial systems of posing extra, unneeded requirements for the implementation of cognitive engines for machines (see, for example, the wide literature on cognitive architectures).
- Some non-human traits are not properly addressed in the theories; because being out the the human spectrum are considered irrelevant concerning the achievement of machine consciousness.

4.2 RETHINKING CONSCIOUSNESS TRAITS

Some commonly accepted consciousness aspects shall be rethought under a non-chauvinistic light to achieve the generality that science and engineering require.

For example, consciousness *seriality* and *integration* have been hallmarks of some widely quoted theories of consciousness.¹³ Functional departures from these are considered pathological, but this is only true under the anthropocentric perspective. Consciousness seriality implies that an agent can only have one stream of consciousness. However, from a general systems perspective, nothing prevents a machine from having several simultaneous streams.

This aspect of seriality is closely related to *attention*. According to Taylor,¹⁴ attention is the crucial gateway to consciousness and architectural models of consciousness shall be based on attention mechanisms. However, using the same analysis as before, nothing prevents a machine from paying attention to several processes simultaneously. The rationale of attention mechanisms seems to be the efficient use of *limited* sensory processing resources (esp. at higher levels of the cognitive perception pipeline). But in the case of machines, if the machine architecture is scalable enough, it is in principle possible to incorporate perceptual resources as needed to pay concurrent attention to several processes. This is also related to the limited capacity trait of human consciousness.

Integration is another human trait that may be unnecessary in machines. Consciousness integration implies that the collection of experiences flowing up from the senses are integrated in a single experiential event. Dissociated experience is abnormal—pathological—in humans but it need not be so in machines. In fact, in some circumstances, being able to keep separated streams of consciousness may be a benefit for machines (e.g., for cloud-based services for conscious machines).

Concepts like subjectivity, individuality, consciousness ontogenesis and filogenesis, emotional experience, sensory qualia modalities, etc. all suffer under this same analysis. Any future, sound theory of consciousness shall necessarily deal with a wider spectrum of traits like bat audio qualia or robot LIDAR qualia.

5. CONCLUSIONS

Universality renders deep benefits. This has been widely demonstrated in science, for example, when the dynamics of falling objects on the earth surface was unified with the dynamics of celestial objects.

We must escape the trap of anthropomorphism to reach a suitable theory for artificial consciousness engineering. Just consider the history of “artificial flight,” “artificial singing,” or “artificial light.” We need not create mini suns to illuminate our rooms. We don’t need to copy, nor imitate, nor fake human consciousness for our machines. We don’t need the whole iguana of mind; what we need are analysis and first principles.

Any general (non anthropocentric) consciousness research program will produce benefits also in the studies of human consciousness because it can provide inspiration for deeper, alternate views of human consciousness. For example, human brains have parallel activities (not serial but concurrent) in different levels of a heterarchy. Considerations coming from conscious distributed artificial systems will help clarify issues of individual minds—normal and pathological—and social consciousness.

Philosophy is a bold endeavor. It goes for the whole picture. Philosophy of mind shall be aware of this and realize that consciousness escapes humanity.

NOTES

1. Wason, *System Engineering Analysis, Design, and Development: Concepts, Principles, and Practices*.
2. Sloman, “The Mind as a Control System”; Prescott et al., “Layered Control Architectures in Robots and Vertebrates”; Sterelny, *The Evolution of Agency and Other Essays*; Winning, “The Mechanistic and Normative Structure of Agency.”
3. Sanz and Meystel, “Modeling, Self and Consciousness: Further Perspectives of AI Research.”
4. Simon, *The Sciences of the Artificial*.
5. In fact, there are control systems engineering methods that do not perform this separation, addressing mind and body as a single whole by concurrent co-design or by embedding control forced dynamics by reengineering of the body.
6. Webb, “Animals Versus Animats: Or Why Not Model the Real Iguana?”
7. Russell and Norvig, *Artificial Intelligence: A Modern Approach*.
8. Sanz et al., “Consciousness, Meaning and the Future Phenomenology.”
9. Conant and Ashby, “Every Good Regulator of a System Must Be a Model of That System”; M. Branicky et al., “A Unified Framework for Hybrid Control: Model and Optimal Control Theory.”
10. Hernández et al., “A Self-Adaptation Framework Based on Functional Knowledge for Augmented Autonomy in Robots.”
11. “Ontology” in knowledge engineering is a specification of a conceptualization. This specification is used to ground the use and interchange of information structures among humans and machines.
12. Tallis, *The Explicit Animal: A Defence of Human Consciousness*.
13. Tononi, “An Information Integration Theory of Consciousness.”
14. Taylor, “An Attention-Based Control Model of Consciousness (CODAM).”

REFERENCES

- Branicky, M., V. Borkar, and S. Mitter. “A Unified Framework for Hybrid Control: Model and Optimal Control Theory.” *IEEE Transactions on Automatic Control* 43, no. 1 (1998): 31–45.
- Conant, R. C., and W. R. Ashby. “Every Good Regulator of a System Must Be a Model of That System.” *International Journal of Systems Science* 1, no. 2 (1970): 89–97.
- Hernández, C., J. Bermejo-Alonso, and R. Sanz. “A Self-Adaptation Framework Based on Functional Knowledge for Augmented Autonomy in Robots.” *Integrated Computer-Aided Engineering* 25 (2018): 157–72.
- Prescott, T. J., P. Redgrave, and K. Gurney. “Layered Control Architectures in Robots and Vertebrates.” *Adaptive Behavior* 7 (1999): 99–127.
- Russell, S., and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, 3rd ed., 2010.
- Sanz, R., C. Hernández, and G. Sánchez. “Consciousness, Meaning and the Future Phenomenology.” In *Machine Consciousness: Self, Integration and Explanation*, edited by R. Chrisley, R. Clowes, and S. Torrance, 55–60. York, UK, 2011.

Sanz, R., and A. Meystel. "Modeling, Self and Consciousness: Further Perspectives of AI Research." In *Proceedings of PerMIS '02, Performance Metrics for Intelligent Systems Workshop*. Gaithersburg, MD, USA, 2002.

Simon, H. A. *The Sciences of the Artificial*, 3rd edition. Cambridge, MA: The MIT Press, 1996.

Sloman, A. "The Mind as a Control System." In *Proceedings of the 1992 Royal Institute of Philosophy Conference on Philosophy and the Cognitive Sciences*, edited by C. Hookway and D. Peterson. Cambridge University Press, 1992.

Sterelny, K. *The Evolution of Agency and Other Essays*. Cambridge University Press, 2001.

Tallis, R. *The Explicit Animal: A Defence of Human Consciousness*. Palgrave Macmillan, 1999.

Taylor, J. G. "An Attention-Based Control Model of Consciousness (CODAM)." *Science and Consciousness Review*. (2002).

Tononi, G. "An Information Integration Theory of Consciousness." *BMC Neuroscience* 5, no. 42 (2004).

Wasson, C. S. *System Engineering Analysis, Design, and Development: Concepts, Principles, and Practices*, 2nd edition. Wiley Series in Systems Engineering and Management. Wiley, 2015.

Webb, B. "Animals Versus Animats: Or Why Not Model the Real Iguana?" *Adaptive Behavior* 17, no. 4 (2009): 269–86.

Winning, R. J. "The Mechanistic and Normative Structure of Agency." PhD thesis, University of California, San Diego, 2019.

governed way. We can reason clearly about the properties of such mathematical objects and have discovered some of their limits. Where things become less clear is when we ask what it means to say that some *physical object* computes.

One response is to simply dismiss questions about physical computation as being nothing more than a matter of interpretation. Arguably, one can construct an interpretation whereby *any* physical object or any arbitrary mereological sum of objects (your thumb, a lake, Jupiter, or the pile of crumbs on the counter) can be interpreted as being *any* finite-state automaton. Hilary Putnam defended a position roughly along these lines.¹ However, as a response to the problem of distinguishing physical computers from noncomputers, this strategy is philosophically unsatisfying and scientifically unhelpful.²

Stating an adequate criterion for distinguishing physical computers from noncomputers has proven difficult for philosophers. Clearly, the problem is not solved simply in virtue of having a good mathematical theory of computability. More is needed. This is because a satisfactory account of physical computation, unlike a mathematical theory of computability, should provide individuation conditions that distinguish objects and processes that compute from those that do not. In his recent book *Physical Computation: A Mechanistic Approach*, Gualtiero Piccinini defends a novel and appealing theory of physical computation. On his view, a physical computer is "a mechanism whose teleological function is to perform a physical computation and a physical computation is the manipulation of a medium-independent vehicle according to a rule."³ In this paper, I will examine the strengths and limitations of this position, concentrating on the question of whether the mechanistic approach has the resources to provide a satisfying account of the individuation of physical computers. Unfortunately, the concept of mechanism cannot provide an illuminating account of the metaphysics of physical computation. This is because the notion of mechanism is insufficiently fundamental and insufficiently general. Nevertheless, Piccinini's book provides an accurate map of the philosophical problems associated with the individuation of physical computers and is filled with important distinctions and insights. Piccinini offers the best attempt to date to answer these questions.

Piccinini shares a view of mechanism with the New Mechanists in philosophy of biology and philosophy of cognitive science. The New Mechanist approach began to take shape in the early 1990s with the work of Bill Bechtel and Bob Richardson, especially in their book *Discovering Complexity*. Today, the canonical reference for the New Mechanist position is Peter Machamer, Lindley Darden, and Carl Craver's "Thinking about Mechanisms." In his *Stanford Encyclopedia of Philosophy* entry "Mechanism in Science," Carl Craver writes that all definitions of mechanism involve four characteristic features: (1) a phenomenon to be explained, (2) parts, (3) causings, and (4) organization.⁴ According to Bechtel and Abrahamsen, for example, "[a] mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena."⁵

PHYSICAL COMPUTATION: A MECHANISTIC ACCOUNT

Should Physical Computation Be Understood Mechanistically?

John Symons
UNIVERSITY OF KANSAS

Common sense tells us that a rock is not a computer but that your laptop is. While common sense is untroubled by the straightforward cases, it is of little help with the exceptions and the exotic cases: Does a broken or faulty laptop count as a computer? And what if aliens used devices that looked like rocks to compute in ways we are unable to understand? Is it sometimes correct to think of rivers, forests, or strands of DNA as computers? Is the central nervous system of an animal a computer? Some people like to think of brains and minds as computers, but isn't this an odd claim given how unlike ordinary computers biological systems seem to be? We could begin to answer such questions in a principled way if we had a plausible theoretical account of the kinds of things that are physical computers. Currently we don't have such an account.

Our uncertainty with respect to the nature of physical computers contrasts sharply with our excellent understanding of the mathematics of computation; we have a well-articulated theory of computable functions, and we have a good mathematical model of systems that compute. The formalism of computability is elegant, clear, and relatively easy to understand. We say that a system that computes is a finite state automaton. A finite state automaton is a *mathematical object* that can be in one state at a time and that switches between states in a rule-

Very broadly speaking, the New Mechanists provide an account of mechanism that they have gleaned from the manner in which some biologists and some neuroscientists give explanations. They notice that in many fields in the biological sciences explanations seem to work by showing how structured parts and processes are orchestrated so as to be responsible for the way things appear to happen. What is meant by orchestration and responsibility is left relatively loosely characterized. Nevertheless, New Mechanists correctly point to the ubiquity and wide acceptance of such patterns of explanation in biology. On this view, the job of at least some biologists is to discover the mechanisms that produce some phenomena.

The New Mechanists are a group of pragmatically inclined philosophers who are guided by what they see as the explanatory norms and practices of the scientific communities they know best. In this spirit, the New Mechanists sometimes seem to imply that their accounts capture a metaphysically agnostic core of ordinary explanatory practice in the sciences. While New Mechanists assume a modest posture in relation to scientific practice, there is (or should be) more to New Mechanism than just a neutral way of parsing accepted styles of explanation in scientific communities. Mechanists regard the explanatory practices of scientists as indicating that there really are distinguishable parts of mechanisms with well-defined relations and activities or operations. In this sense, New Mechanism has non-trivial metaphysical commitments.

Piccinini puts this view of mechanism into action in his description of physical computers. On his view, physical computers can be understood in terms of their component parts, their *functions*, and their organization. Physical computers are objects whose function, according to Piccinini, is to perform physical computations. Physical computation is the manipulation of medium-independent vehicles according to rules. Rules and functions resist reduction to mechanism, but on Piccinini's account the physical systems that follow those rules are ultimately nothing more than mechanisms.

In practice, physical computers are certainly picked out by reference to their functions, and those functions are determined (at least in part) by reference to abstract rules. However, a description of what the physical computer *is* (as opposed to how we happen to identify it) need not include mention of those rules and functions. The role of rules and functions in his presentation is intended to serve as a way of distinguishing the subset of mechanisms that count as physical computers (namely, those that serve the function of following rules) from those that do not. Thus, Piccinini does not claim that we rely exclusively on the notion of mechanism in the process of *identifying* some objects as physical computers. In fact, it might be a matter of epistemic necessity that our identification of physical computers depends on reference to rules and functions. Nevertheless, the way we happen to distinguish physical computers from other mechanisms is not ultimately relevant to what they are. Identification is not the same as individuation.

On Piccinini's account, mechanism plays a role in helping us to understand the metaphysics of physical computation.

One important reason to give an account of physical computation in terms of mechanism is to sidestep Putnam's concerns about the challenge of determining a unique mapping from abstract computational characterizations to physical implementation. Putnam argued that every ordinary open physical system can be interpreted as implementing every finite-state automaton. This means that every physical object can justifiably be interpreted as a computer,⁶ but it would also render the mapping account useless as a means of individuating physical computers. Piccinini claims that the mapping approach poses the problem incorrectly and so he does not answer Putnam's challenge. His strategy is to abandon the question of how we should interpret the mapping between states of a physical system and states of an abstract computer. He suggests that even amended or strengthened mapping accounts will trivialize the claim that a physical system computes.⁷ While Piccinini does not employ the distinction between individuation and identification in the manner discussed above, the mapping approach can be understood as addressing epistemic considerations involved in identifying physical computers whereas the correct approach would address the problem of individuating physical computers. In this sense, the mechanistic account can be understood as offering an alternative strategy for thinking about those individuation conditions. The activity of physical computation is a purely mechanistic matter on Piccinini's view.

This explanatory strategy allows him to claim that what it is to be a physical computer does not depend on representational or semantic concepts. If correct, this would mark an important step forward insofar as representation and semantics have been central to many previous accounts of computation and have presented deep conceptual difficulties familiar to the traditions of philosophy of mind and philosophy of language. Central to Piccinini's contention that the mechanistic approach can do without representation and semantics is his understanding of how physical computers perform *concrete computations*. "A physical system is a computing system," he writes, "iff it's a system that performs concrete computations."⁸ Concrete computations support the teleological function of the physical computer insofar as the rules characterizing the transformation of vehicles are fixed by the purpose of the machine. Nevertheless, that transformation itself can be described independently of abstract rules. Rather than individuating these vehicles by reference to their semantical roles, the non-semantic properties of strings of discrete states that figure in program-controlled computers can serve as the basis of the explanation of the operation of those systems. "Different bits and pieces of these strings of states have different effects on the machine."⁹ Because of this, he argues, we can describe the operation of sub-strings and states without mentioning what their semantics are. Moreover, according to Piccinini, the mechanistic account of vehicles can explain syntactical properties of digital computation. Since mechanism is a more basic notion than the syntax of a language, Piccinini can argue that a mechanistic approach has the additional advantage of providing an account of non-digital forms of computation whereas the syntactical approach fails to do so. The promise of the mechanistic approach is that only non-semantic, non-representational, and even non-syntactic individuation

conditions for the vehicles of computation will figure in mechanistic explanations.

The mechanistic approach seems more metaphysically basic than, for example, the semantic or representational level of analysis. However, mechanistic approaches have a hard time shedding light on the problem of individuation.¹⁰ To begin with, the metaphysical significance of mechanism is not clear. In part, this is because the mechanistic approach is primarily derived from reflections on scientific explanation as a practice rather than on those aspects of reality that ground the reliability of scientific explanations. It is obviously true that in many areas of biology explanations are given in mechanistic terms. But this is not an explanation of why mechanistic accounts are widely accepted in those areas or why they have been so successful. If anything, the success of mechanistic explanation in these areas itself demands explanation. Worth mentioning too is the need for an account of the notion of part, cause, and organization that undergird the notion of mechanism itself.

A second reason for the difficulty of using mechanistic accounts to shed light on individuation and other metaphysical topics arises from the fact that mechanism is both an insufficiently general and an insufficiently fundamental notion. With respect to generality, notice, for example, that the mereological commitments of New Mechanism make it inapplicable to some kinds of objects and phenomena. The New Mechanists acknowledge that their view is only applicable to some regions of even more mundane domains of scientific explanation. Even in biology, as Skipper and Millstein point out, and as Craver and Tabery acknowledge, the mechanistic notion of part is difficult to sustain both at the level of the very small in well-understood biochemical phenomena and at the level of the very large in natural selection.¹¹ Our best understanding of nature does not support the idea that everything can be understood or explained in terms defended by the New Mechanists; therefore, the mechanistic approach is unlikely to have the resources to answer general questions concerning individuation.

With respect to fundamentality, it is difficult to reconcile New Mechanism with basic physics. For example, it is difficult to understand field-theoretic explanations within a mechanistic framework. Furthermore, the ontological implications of quantum mechanics pose a challenge insofar as New Mechanism seems to rely on local causal interactions and the idea of isolable parts with definite properties.

In the case of physical computation, there is a sense in which such metaphysical concerns might seem irrelevant. Most physical computers are artifacts with very special characteristics. The purposes of most physical computers are dictated by the demands of the target markets of their manufacturers. Since the practice of manufacturers serves specific commercial or scientific purposes and since most computers are built from pre-fabricated parts according to a plan for organization, it is appropriate to characterize these physical computers as functional mechanisms.

The component parts of physical computers also have parts, of course. However, the mereological ground

floor for individuating physical computers (on Piccinini's view) is fixed by the logic of software.¹² At bottom, the primitive *computationally relevant* components for conventional computers are logic gates. In principle, any bistable system can serve as a primitive component for a computing technology if it allows us to reliably manipulate whether that system is in one or the other equilibrium state. There is nothing relevant to computational function *per se* at more fundamental levels than the level of such bistable systems.

The mechanistic approach to physical computation does not tell us why or how the primitive components of a conventional physical computer work as they do. As far as the mechanistic perspective is concerned, once one has reliably manipulable bistable systems to build on, everything underneath can be ignored. However, the promise of the mechanistic approach to physical computation had been the possibility of providing an explanation of the relevant primitive components: "The mechanistic explanation of a primitive computing component—say, an AND gate—explains how that component exhibits its specific input-output behavior. In our example, the components of a particular electrical circuit, their properties, and their configuration explains how it realizes an AND gate."¹³ Unfortunately, even in relatively mundane engineering contexts, for example, in the explanation of modern transistors and circuits, New Mechanist-style explanations fall short. The New Mechanist approach will not get very far, for example, in the understanding of semi-conductors or in the understanding of field effect transistors more generally. The details of how fields behave are completely opaque if one's only explanatory resource is mechanism. Mechanisms are composed of objects or processes with definite properties. The quantum mechanical account of the behavior of fields does not assume local causal interactions of the mechanist kind nor does it assume objects with definite properties.¹⁴ Mechanists might reject this concern given that once reliable bistable systems emerge at some level everything underneath can be ignored. However, the fundamental physical nature of these bistable systems turns out to be relevant to their operation and must be taken into account in engineering contexts. Consider, for instance, the consequences of miniaturization. As the miniaturization of transistors continues, quantum tunneling will make it difficult to insulate the relevant parts of the circuits. In order to build the primitive bistable components of computers, quantum-level behaviors would be unavoidable. In this context one solution that has been proposed involves exploiting quantum tunneling as part of the operation of the circuits themselves. There will be no explanation of such components and no light will be shed on practical solutions from a mechanistic perspective.

These challenges and details fall below the level of the primitive components of the mechanistic account of computation. The trouble here is that these primitive components are picked out by the logic of a particular kind of software. Thus the mereological ground floor for the mechanist's treatment of computation is established by reference to the abstract characteristics of computation rather than being explained by, or grounded in, mechanism. If we were concerned about questions like, What are the

physical constraints on computation? How do bistable systems emerge in physical reality? or any number of other questions about the physical instantiation of computation, the mechanistic approach will not satisfy.

ACKNOWLEDGMENTS

Many thanks to Mark Bickhard, Piotr Boltuc, Frances Egan, Jack Horner, Corey Maley, Marcin Miłkowski, Nico Orlandi, Gualtiero Piccinini, and Sarah Robins for helpful discussions of this material.

This work is supported by The National Security Agency through the Science of Security initiative contract #H98230-18-D-0009.

NOTES

1. See especially Putnam, *Representation and Reality*.
2. Buechner, *Gödel, Putnam, and Functionalism: A New Reading of 'Representation and Reality'*.
3. Piccinini, *Physical Computation: A Mechanistic Account*, 10.
4. Craver and Tabery, "Mechanisms in Science."
5. Bechtel and Abrahamsen, "Explanation: A Mechanist Alternative," 523.
6. Putnam, *Representation and Reality*, 120–25.
7. Piccinini, *Physical Computation: A Mechanistic Account*, 22.
8. *Ibid.*, 118.
9. *Ibid.*, 45.
10. Symons, "The Individuality of Artifacts and Organisms."
11. Skipper and Millstein, "Thinking about Evolutionary Mechanisms: Natural Selection"; Craver and Tabery, "Mechanisms in Science."
12. Horner and Symons, "Understanding Error Rates in Software Engineering: Conceptual, Empirical, and Experimental Approaches."
13. *Ibid.*, 155.
14. See Kuhlmann and Glennan, "On the Relation Between Quantum Mechanical and Neo-Mechanistic Ontologies and Explanatory Strategies," 338.

REFERENCES

- Bechtel, William, and Adele Abrahamsen. "Explanation: A Mechanist Alternative." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 36, no. 2 (2005): 421–41.
- Bechtel, William, and Robert C. Richardson. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. MIT Press, 1993.
- Buechner, Jeff. *Gödel, Putnam, and Functionalism: A New Reading of 'Representation and Reality'*. MIT Press, 2008.
- Craver, Carl, and Tabery, James, "Mechanisms in Science." *The Stanford Encyclopedia of Philosophy*. Spring 2017 Edition. Edward N. Zalta (ed.). Available at <https://plato.stanford.edu/archives/spr2017/entries/science-mechanisms/>. Last Accessed January 2, 2019.
- Horner, Jack, and John Symons. "Understanding Error Rates in Software Engineering: Conceptual, Empirical, and Experimental Approaches." *Philosophy & Technology* (2019): 1–16.
- Kuhlmann, Meinard, and Stuart Glennan. "On the Relation Between Quantum Mechanical and Neo-Mechanistic Ontologies and Explanatory Strategies." *European Journal for Philosophy of Science* 4, no. 3 (2014): 337–59.
- Machamer, Peter, Lindley Darden, and Carl F. Craver. "Thinking about Mechanisms." *Philosophy of science* 67, no. 1 (2000): 1–25.
- Piccinini, Gualtiero. *Physical Computation: A Mechanistic Account*. Oxford: Oxford University Press, 2015.
- Putnam, Hilary. *Representation and Reality*. MIT Press, 1988.
- Skipper Jr., Robert A., and Roberta L. Millstein. "Thinking about Evolutionary Mechanisms: Natural Selection." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 36, no. 2 (2005): 327–47.

Symons, John. "The Individuality of Artifacts and Organisms." *History and Philosophy of the Life Sciences* (2010): 233–46.

Commentary on Gualtiero Piccinini's *Physical Computation: A Mechanistic Perspective*

Martin Roth
DRAKE UNIVERSITY

Gualtiero Piccinini's *Physical Computation: A Mechanistic Perspective* is an example of first-rate scholarship—rigorous, clear, well-informed, and well-argued—and though I have some criticisms of the perspective Piccinini develops in his book, I suspect that what emerges is not so much a fundamental disagreement as it is a difference in emphasis. My criticisms focus on themes developed in Chapter 5—"From Functional Analysis to Mechanistic Explanation." In that chapter Piccinini challenges the "received view" according to which functional analysis is distinct and autonomous from mechanistic explanation. As Piccinini acknowledges, the views developed in that chapter draw heavily from a 2011 paper that Piccinini wrote with Carl Craver.¹ My criticisms will be directed at that paper.²

When we identify something functionally—a mousetrap, a gene, a legislature—we identify it in terms of what it does. Many biological terms have both a functional and an anatomical sense: an artificial heart is a heart by function but is not an anatomical heart, and computational neuroscience was conceived when the word "brain" became a functional term as well as an anatomical one. Functional analysis is the attempt to explain the properties of complex systems—especially their characteristic capacities—by the analysis of a systemic property into organized interaction among other simpler systemic properties or properties of component subsystems. This explanation-by-analysis is *functional* analysis because it identifies analyzing properties in terms of what they do or contribute, rather than in terms of their intrinsic constitutions. For example, a circuit diagram describes or specifies a circuit in a way that abstracts away from how the components, including the "wires," are actually made. The strategy of explaining the capacities of complex systems by functional analysis is ubiquitous in science and engineering, and by no means special to psychology.

From the point of view of functional analysis, capacities are dispositional properties, and the dispositional properties of a complex system are explained by exhibiting their manifestations as the disciplined manifestation of dispositions that are components of the target disposition, or by the disciplined interaction of the dispositions of the system's component parts. It should be obvious that the explanatory targets of this sort of analysis are not points in state space or particular trajectories through it. Rather, the aim of this kind of analysis is to appeal to a system's design in order to explain why one finds the trajectories one does and not others. The design provides a model of

the state space and constrains the possible paths through it, thereby explaining certain regularities in the system's behavior. More generally, the strategy is to explain the capacities of a complex system by exhibiting the abstract functional design of that system—to show, in short, that a system with a certain design is bound to have the capacity in question. Designs can do this because functional terms pick out the causal powers that are relevant to the capacity being analyzed. Functional terms are in this sense *causal relevance filters*: by selecting from the myriad causal consequences of a system's states, processes, or mechanisms those that are relevant to the target capacity, functional characterization makes the contributions of those states, processes, or mechanisms transparent. It is precisely this transparency that enables us to understand why anything that possesses these states, processes, or mechanisms is bound to have the capacity in question. Without this filtering, we are simply left with a welter of noisy detail with no indication of what is relevant and what is a mere by-product of this or that implementation. Causal relevance filtering is, therefore, just abstraction from the implementation details that are irrelevant to the achievement of the targeted capacity. Implementations that differ in those details but retain the design will thus all exhibit the targeted capacity. In this way, the possibility of multiple realization is an inevitable consequence of causal relevance filtering, and so it should come as no surprise to find that functional analyses subsume causal paths that have heterogeneous implementations.

However, it would be a mistake to wed the explanatory power of functional analysis to assumptions about actual multiple realization, for even if there is only one nomologically possible way to implement a design, giving implementation details that go beyond what is specified by an analysis adds nothing to the explanation provided by the design. For example, suppose there is just one nomologically possible way to implement a doorstop—say, by being a particular configuration of rubber. In this case, it would be plausible to hold that being a doorstop—the type—is identical to being a particular configuration of rubber—the type. Because type-type identities give you property reductions, being a doorstop would thus reduce to being a particular configuration of rubber. But a functional analysis that specifies something as a *doorstop* would still be autonomous, in the following sense. Being a particular configuration of rubber comes with any number of causal powers. One of those powers is stopping doors, and in the context of the imagined functional analysis, stopping doors is the only causal power of this particular configuration of rubber that matters to having the target capacity. If in our analysis we replace the word “doorstop” with the phrase “rubber configured thus and so,” we won't lose anything as far as the causation goes. However, we will lose the transparency functional analysis affords *unless we specify explicitly that stopping doors is the relevant causal power*. But then the explanation is tantamount to the explanation given in terms of the word “doorstop,” i.e., the explanation does not give us anything beyond what is provided by the functional analysis itself.

If we focus on the causal explanation of events and assume type-type identity, then framing explanations in

terms of the word “doorstop” is guaranteed to give you nothing beyond what framing explanations in terms of the phrase “rubber configured thus and so” gives you, and this is why it has been generally assumed that reduction is incompatible with autonomy. From the perspective of functional analysis, by contrast, autonomy can live with reduction. Design explanations are autonomous in the sense that they do not require “completion” by annexing implementation details, e.g., in the case imagined above, it is irrelevant to explaining the target capacity whether a specific doorstop is a particular configuration of rubber. But design explanations are also autonomous in the sense that adding implementation details would *undermine* the transparency provided by causal relevance filtering and thereby obviate the understanding provided by the design. A doorstop may be a particular configuration of rubber, but replacing the word “doorstop” with the phrase “rubber configured thus and so” masks the information needed to understand why a system has the target capacity.

I am sympathetic to the thought that complete knowledge of implementation details would contribute to a fuller understanding of those systems whose capacities are targeted by functional analysis. Indeed, such details are necessary for understanding how a system manages to have the very causal powers that are picked out by functional analysis. But having a fuller understanding of a system, in this sense, is not the same thing as having a more complete explanation of the capacity targeted for functional analysis. For example, when we analyze the capacity to multiply numbers in terms of a partial products algorithm, the specification of the algorithm tells us nothing about the states, processes, or mechanisms of a system that implements the algorithm (except in the trivial sense that the states, processes, or mechanisms of any system that implements the algorithm are sufficient for implementing it). However, as far as explaining the capacity goes—what we might call the “multiplication effect”—the analysis provided by the algorithm is complete, i.e., the analysis allows us to understand why any system that has the capacity for computing the algorithm *ipso facto* exhibits the multiplication effect. Because details about how the algorithm is implemented add nothing to the analysis, such details are irrelevant to the explanation of the capacity.

The perspective I have outlined here suggests that we need to distinguish two kinds of explanations, what I call *horizontal* and *vertical* explanations. Horizontal explanations explain capacities by appeal to a design that is specified by functional analysis. They answer the question “Why does system S have capacity C?” by specifying some design D. Vertical explanations specify implementations. They answer the question “How is design D realized in system S?” Neither type of explanation is subsumption under law. And neither is in the business of explaining individual events. The explananda are, respectively, capacities and designs.

I suspect that the tendency to conflate *explaining a capacity via functional analysis* with *explaining how a functional analysis (a design) is implemented* has led to a misunderstanding concerning the relationship between functional analysis and mechanistic explanation. Following

Bechtel and Abrahamsen, a mechanism “is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena.”³ As I see it, the goal of discovering and specifying mechanisms is often or largely undertaken to explain how the analyzing capacities specified by a functional analysis are implemented in some system. In this way, though the horizontal explanations provided by functional analysis are autonomous from the vertical explanations provided by specifying mechanisms, the two explanations complement each other.

However, Piccinini and Craver appear to challenge this claim of autonomy.⁴ They argue that functional analyses are “mechanism sketches”: functional analyses and the design explanations they provide are “incomplete” until filled out with implementation details, and in that way, the explanations provided by functional analysis are not autonomous. However, I think their argument involves a misidentification of the relevant explanatory targets of functional analysis—capacities—and a correlative conflation of explanation and confirmation. I’ll take these up in turn. Piccinini and Craver write that “Descriptions of mechanisms . . . can be more or less complete. Incomplete models—with gaps, question-marks, filler-terms, or hand-waving boxes and arrows—are mechanism sketches. Mechanism sketches are incomplete because they leave out crucial details about how the mechanism works.”⁵ Sketches being what they are, I have no quarrel with the claim that mechanism sketches are incomplete, and insofar as mechanistic explanations explain by showing how a mechanism works, I agree that filling in the missing details of a mechanism sketch can lead to a more complete mechanistic explanation. The crucial issue here, however, is whether functional analyses should be viewed as mechanism sketches. To motivate the claim that functional analyses are mechanism sketches, we have to assume that abstraction from implementation detail inevitably leaves out something crucial to the *analytical explanation of a target capacity*, something that implementation details would provide. But as I’ve already argued, the opposite is in fact true; adding implementation details obfuscates the understanding provided by functional analysis.

Instead of favoring the autonomy of functional analysis, however, Piccinini and Craver think that abstraction from implementation details actually works against claims of autonomy. They write:

Autonomist psychology—the search for functional analysis without direct constraints from neural structures—usually goes hand in hand with the assumption that each psychological capacity has a unique functional decomposition (which in turn may have multiple realizers). But there is evidence that . . . several functional decompositions may all be correct across different species, different members of the same species, and even different time-slices of an individual organism. Yet the typical outcome of autonomist psychology is a single functional analysis of a given capacity. Even assuming for the sake of the argument that autonomist psychology

stumbles on one among the correct functional analyses, autonomist psychology is bound to miss the other functional analyses that are also correct. The way around this problem is to let functional analysis be constrained by neural structures—that is, to abandon autonomist psychology in favor of integrating psychology and neuroscience.⁶

I think this argument conflates *explanatory* autonomy with *confirmational* autonomy. If a capacity admits of more than one analysis, merely providing an analysis will, of course, leave open the question of whether the analysis provided correctly describes how a system manages to have the capacity in question (assuming it does have the capacity). Knowledge of neural structures is undoubtedly relevant to settling the question of which analysis is correct, but bringing such knowledge to bear in this instance would be an exercise in confirming a proposed analysis, not explaining a capacity. Suppose there are two possible analyses, A and B, for some capacity C, and the neurological data suggests that analysis A is implemented in system S. The explanation of capacity C in S is provided by A, not by the neural structures evidence about which confirms A.

Arguably, nothing enjoys confirmational autonomy from anything else. As such, neuroscience that makes well-confirmed psychological capacities impossible or unlikely needs revision as much as a design hypothesis in psychology that appears to have no plausible neural implementation. I thus agree with Piccinini and Craver’s claim that “psychologists ought to let knowledge of neural mechanisms constrain their hypotheses just like neuroscientists ought to let knowledge of psychological functions constrain theirs.”⁷ This is an invitation to those working in psychology departments and neuroscience departments to talk to each other, and if this is enough to show that psychology is not autonomous from neuroscience, then so much the worse for the autonomy of psychology. But defending the autonomy of *functional analysis* is not the same thing as defending the autonomy of *psychology*. Functional analysis is an explanatory strategy, not a scientific discipline, and when we are careful to distinguish horizontal and vertical explanations, and distinguish confirmation and explanation, the autonomy of functional analysis emerges as unproblematic.

NOTES

1. G. Piccinini and C. Craver. “Integrating Psychology and Neuroscience: Functional Analyses as Mechanism Sketches.” *Synthese* 183, no. 3 (2011): 283–311.
2. My decision to comment on that paper stems largely from the fact that Robert Cummins and I have criticized that paper in another work (“Neuroscience, Psychology, Reduction, and Functional Analysis,” in *Explanation and Integration in Mind and Brain Science*, ed. D. Kaplan [Oxford University Press: Oxford, 2018]), and I am interested in how Piccinini responds to those criticisms. What follows draws heavily from the aforementioned paper.
3. W. Bechtel and A. Abrahamsen. “Phenomena and Mechanisms: Putting the Symbolic, Connectionist, and Dynamical Systems Debate in Broader Perspective,” in *Contemporary Debates in Cognitive Science*, ed. R. Stainton (Blackwell: Malden, MA, 2006), 162.
4. Piccinini and Craver, “Integrating Psychology and Neuroscience.”
5. *Ibid.*, 292.

6. Ibid., 285.

7. Ibid.

Defending the Mapping Account of Physical Computation

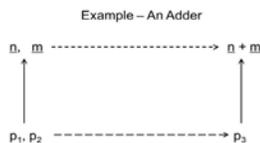
Frances Egan
RUTGERS UNIVERSITY

I will defend a version of the *mapping account* of computation, an account that Piccinini rejects in chapter two of *Physical Computation* in favor, ultimately, of a mechanistic view.¹ I will argue that a characterization of the sort specified by the mapping account is required even if one endorses a mechanistic view of computation. In fact, the two views should not be seen as competitors; both are required for a full account of a computing mechanism.

According to the version of the mapping account that I favor, a physical system S computes a function F just in case:

- i. There is a mapping from (equivalence classes of) physical states of S to the arguments and values of F, such that:
- ii. Causal state-transitions between the physical states mirror the formal dependency relations between the arguments and values of F.

An example may be helpful:



A physical system computes the addition function just in case there exists a mapping from physical state types to numbers, such that physical state types related by a causal state-transition relation $((p_1, p_2) \rightarrow p_3)$ are mapped to numbers $n, m,$ and $n+m$ related as addends and sums. Whenever the system goes into the physical state specified under the mapping as $n,$ and goes into the physical state specified under the mapping as $m,$ it is caused to go into the physical state specified under the mapping as $n+m.$

The mapping specifies, for a given physical system, *what* the system computes, and *how* it computes it—by transitioning through a sequence of causally related physical states. In other words, the mapping characterizes, at a high level of abstraction, the relevant causal organization in virtue of which the physical system is able to compute the specified function. A mapping of this sort has to be true if the physical system in fact computes the function. But a more perspicuous account of the mechanism’s causal organization may characterize it in terms of structures,

processes, and algorithms—in other words, in terms of components and their activities, as mechanists would insist.² The causal dependencies specified at the lower level of the mapping (the causal state transitions) will be consequences of this organization. The mapping isn’t intended to be, and doesn’t replace, an explanatory theory of a computing mechanism, that is, a perspicuous explanation of how the mechanism works.

According to Piccinini’s mechanistic view of physical computation:

A physical system is a computing system just in case it has the following characteristics:

- It is a functional mechanism.
- One of its functions is to manipulate vehicles based solely on differences between different portions of the vehicles according to a rule defined over the vehicles. (275)

Let me make a couple of brief remarks about Piccinini’s mechanistic account. Computers, both natural and artifactual, are undoubtedly functional mechanisms. They typically have specific teleological functions—to compute arithmetical functions, to detect the 3D structure of the scene, to enable navigation, to parse incoming speech—and more generally (and in the case of modern digital computers), to execute cognitive tasks. I find it implausible that computers have the function cited in the mechanistic account. Rather, manipulating vehicles based on differences in their parts according to a rule is the common *means* that computing mechanisms deploy to achieve their specific functions (parsing input strings, adding, and so on). Consider an analogy: applying upward pressure on a cork is a common means that various kinds of corkscrews deploy to achieve their function of removing corks from bottles. But I am not inclined to press this point too hard because I am not sure how the issue would be settled. The question of a natural computer’s function would normally be settled by appeal to its manifest cognitive capacities (navigation, parsing speech, etc.); for artifacts, by appeal to the cognitive task the computer has been designed to achieve. In any event, the mechanistic account is missing a key component—specifying *what* function (in the mathematical sense, now) the system is computing when it manipulates vehicles in accordance with a rule; in other words, it is missing precisely what the mapping account provides.

I turn now to consider how well the mapping account satisfies Piccinini’s desiderata for an adequate account of physical computation. I will focus on the two that might be supposed to cause the most trouble.

Desideratum (4): *The wrong things don’t compute* – the account shouldn’t entail that obvious non-computers in fact compute. Famously, this is where mapping accounts supposedly fail—they are supposedly *too liberal*. Putnam and Searle have argued that rocks, walls, and indeed all physical systems turn out to be computers because their physical states can be mapped to computational

descriptions. If computational descriptions apply to everything then computational “explanations” are trivial.

Here’s what I think a mapping theorist should say in response to the triviality objection. First, the requirement that the physical states specified at the lower level of the mapping be *causally related* will eliminate many spurious mappings. Suppose, then, that we simply accept that for a physical system to compute a function just is for it to have the causal organization specified by the mapping. What would the possibility of unintended realizations imply for explanatory practice in cognitive science? In particular, would it make computational explanation trivial? In a word, *no*. To see why not, we need to focus on the *context* in which computational explanations are typically offered. Theorists deploy computational descriptions to explain the *manifest* cognitive capacities of physical systems such as ourselves, capacities such as adding, understanding, and producing speech, seeing the three-dimensional layout of the scene, and so on. (Rocks and walls have no manifest cognitive capacities.) With the target capacity in mind, the theorist hypothesizes that the system computes some well-defined function (in the mathematical sense), and spells out how computing this function would explain the system’s observed success at the cognitive task. Justifying the computational description requires explaining how computing the value of the function contributes to the exercise of the cognitive capacity. For example, in David Marr’s (1982) account of early vision, computing the Laplacian of a Gaussian of the retinal array produces a smoothed output that facilitates the detection of sharp discontinuities in intensity gradients across the retina, and hence the detection of significant boundaries in the scene. In other words, the computational description is justified by reference to the use to which the computation is put in the exercise of a manifest cognitive capacity. According to the mapping account, the theorist is thereby committed to the system actually having the causal organization specified by the lower level of the mapping. That is hardly trivial. That there may be other systems with this same causal organization, which, *if they were suitably hooked up to other internal mechanisms and situated in the right environment* (and these, it should be emphasized, are substantive constraints!) would enable the larger system to achieve the cognitive task, is simply not relevant.³

Desideratum (5): *The account should explain miscomputation, i.e., “it should explain what it means for a concrete computation to give the wrong output.”* (24) I am not sure why Piccinini thinks the mapping account can’t do this; in fact, it is one of the advantages of the view that an account of miscomputation falls right out of it.

Let’s return to the addition example. When the system goes into the physical state interpreted under the mapping as 57 and goes into the physical state interpreted under the mapping as 43, and then is caused to go into the physical state interpreted under the mapping as 100, it correctly adds. If instead it were to go into the physical state interpreted under the mapping as 99 it would *miscompute* or *make a mistake*. On the other hand, if its battery died or it was dropped into the bathtub it would go into a physical state not interpreted under the mapping at all; it would not

miscompute, it would *malfunction*. (To preserve the way we normally talk, a miscomputation may be seen as a special kind of malfunction.) So the specification of the function computed (here, addition) and the interpretation (here, the mapping of physical states to addends and sums) provides all we need to partition the behavioral space into (i) correct computations, (ii) miscalculations or mistakes, and (iii) malfunctions (states that receive no interpretation under the mapping). The specification of the function computed and the interpretation together provide the *norm* necessary to underwrite this three-fold distinction.

In summary, I think I have assuaged what Piccinini takes to be the most serious worries with the mapping account.

NOTES

1. For versions of mapping accounts, see Cummins, *Meaning and Mental Representation*, and Chalmers, “A Computational Foundation for the Study of Cognition,” among others.
2. For other mechanistic accounts of physical computation, see Fresco, *Physical Computation and Cognitive Science*, and Milkowski, *Explaining the Computational Mind*.
3. See Egan, “Metaphysics and Computational Cognitive Science: Let’s Not Let the Tail Wag the Dog,” for elaboration of this argument.

REFERENCES

Chalmers, D. J. “A Computational Foundation for the Study of Cognition.” *Journal of Cognitive Science* 12 (2011): 323–57.

Cummins, R. *Meaning and Mental Representation*. Cambridge, MA: MIT Press, 1989.

Egan, F. “Metaphysics and Computational Cognitive Science: Let’s Not Let the Tail Wag the Dog.” *The Journal of Cognitive Science* 13 (2012): 39–49.

Fresco, N. *Physical Computation and Cognitive Science*. New York: Springer, 2013.

Marr, D. *Vision*. New York: Freeman, 1982.

Milkowski, M. *Explaining the Computational Mind*. Cambridge, MA: MIT Press, 2013.

Comments on Gualtiero Piccinini, Physical Computation: A Mechanistic Account

Nico Orlandi

UNIVERSITY OF CALIFORNIA, SANTA CRUZ

Gualtiero Piccinini’s new book is an insightful, informative, and admirably clear project. The book aims to give a plausible account of what concrete computation consists in. The account meets certain *desiderata*, and it contrasts in some important respects both with the orthodoxy that regards computation and representation as going together, and with the orthodoxy that regards computation as just a matter of mapping. Part of the problem with the two orthodoxies, Piccinini argues, is that they propose, respectively, a too restrictive and an overly liberal understanding of computation. Piccinini’s book aims to delineate a middle passage where just the right things compute and the rest don’t.

The centerpiece idea of the book is the *mechanistic account of Computation*. In these comments, I focus on getting clearer on the account. Since the comments must be brief, I focus on two questions in particular (and order them in the text by how important I regard them to be). The first question concerns the notion of “medium independence,” which is crucial to the mechanistic account. The second question concerns the notion of teleological function that Piccinini introduces and defends.

Before delving into the issues, it is worth pointing out that, in the process of presenting the mechanistic account, Piccinini offers insights on the relationship between functional and mechanistic explanation, on the nature of analogue computation, on the status of connectionism, and he proposes a novel way of understanding teleological function. The book is rich, clear, and truly worth reading.

MEDIUM-INDEPENDENCE

The mechanistic account of computation that Piccinini introduces holds the following: “A physical computation is the manipulation (by a functional mechanism) of a medium-independent vehicle according to a rule. A medium-independent vehicle is a physical variable defined solely in terms of its degrees of freedom (e.g. whether its value is 1 or 0 during a given time interval) as opposed to its specific physical composition (e.g. whether it’s a voltage and what voltage values correspond to 1 or 0 during a given time interval). A rule is a mapping from inputs and/or internal states to internal states and/or output” (10).

Piccinini says that, in this context, a variable is a physical state that can change over time (121, fn. 2). The notion of medium-independence accounts, in the first place, for the fact that computations can be implemented in different physical systems. Medium-independence entails multiple realizability (123). An account of computation that does not allow for multiple realizability—that is, for the same computation being implementable in different physical media—would be a wanting account.

But the notion of medium independence, as stated, is overly liberal. Any physical (or chemical) state that changes over time (it seems) can be *defined in terms of its degrees of freedom* rather than in terms of its specific physical composition. If that is all there is to the notion of medium independence, then it seems that the mechanistic account would, like the mapping account, count too many things (such as stomachs and washing machines) as computing (147 and 275).

Sometimes Piccinini talks of medium-independence in terms of a process acting only on certain physical properties of a state (and not on others). He says: “When we define concrete computations and the vehicles that they manipulate we need not consider all of their specific physical properties. We may consider only the properties that are relevant to the computation, according to the rules that define the computation” (122). But this also seems to be a fairly permissive notion. Despite what Piccinini says (147), the physical states involved in digestion are processed according to some of the physical-chemical properties of the states and not others (they are not

processed, for example, according to the color properties they have).

There is clearly further work that the notion of medium independence is required to do. Piccinini says that, although medium-independence entails multiple realizability, the reverse is not the case. This is because medium independence puts structural constraints on the vehicles of computation (123). But what kind of constraints are structural constraints?

We know what the constraints are not. They are not semantic in nature. Medium-independence is invoked to avoid an overly liberal view of computation without appealing to the notion of representation, as the semantic view of computation does. Medium-independent vehicles may, but need not, carry information about something and/or represent it.

We are, however, left wondering what Piccinini means by structural constraints. To illustrate what he means, Piccinini gives the example of digits: “In the case of a medium-independent property, the structural constraint comes from requiring that the medium—any medium that realizes the property—possesses the degrees of freedom that are needed to realize the property. In the case of digits, their defining characteristic is that they are unambiguously distinguishable by the processing mechanism under normal operating conditions. . . . The rules defining digital computations are defined in terms of strings of digits and internal states of the system, which are simply states that the system can distinguish from one another. No further physical properties of a physical medium are relevant to whether they implement digital computations. Thus, digital computations can be implemented by any physical medium with the right degrees of freedom” (123).

As already mentioned, it is unclear how the degrees of freedom requirement would constitute a constraint on a state. On the other hand, the constraints on digits—that they are distinguishable unambiguously by a system throughout a process—do indeed mark a distinctive requirement on computation, but they also restrict computations to only the digital ones. Analogue, quantum, and even (some) connectionist networks would fail to qualify as manipulating medium-independent vehicles because continuous variables and qudits are not unambiguously distinguishable from one another (6 and 124). So the notion of medium-independence needs better unpacking on pain, among other things, of having a mechanistic account that is not superior to competitors in the things it regards as computing.

FUNCTION

Piccinini understands teleological functions as contributions to objective goals of organisms (108). A teleological function in an organism, according to him, is a stable contribution by a trait (or component, activity, property) of organisms belonging to a biological population to an objective goal of those organisms.

This account is supposed to, among other things, avoid the epistemic problems of etiological accounts of function

which identify functions as what historically has served an organism in survival (102). One of the problems with etiological accounts is that “the causal histories that ground functions . . . are often unknown (and in many cases, unknowable), making function attribution difficult or even impossible” (102).

The challenge for Piccinini here is to spell out the notion of a goal, and explain why goals themselves are not historically established. People’s goals are, of course, varied and dependent on the present. But in his formulation of function, Piccinini refers to the goals of organisms belonging to a biological population. What are the goals of such organisms and are they knowable? Don’t they depend on the evolutionary history of the population? It seems that, absent some further unpacking, the same problems Piccinini raises for teleology apply to his position.

The Mechanistic Account of Physical Computation: Some Clarifications

Gualtiero Piccinini
UNIVERSITY OF MISSOURI-ST. LOUIS

I’m deeply grateful to my commentators for their insightful remarks, which provide me this opportunity to clarify the mechanistic account of physical computation. When I refer to the mechanistic account, I mean the view I defend in my book.¹ The mechanistic account says that computational explanation is a special kind of mechanistic explanation. According to the mechanistic account, physical computation is the manipulation of medium-independent vehicles according to a rule, by a mechanism that is performing one of its teleological functions.

1. HORIZONTAL EXPLANATIONS ARE MECHANISTIC

The mechanistic account maintains that functional analyses are mechanism sketches. This is relevant because computational explanation is traditionally said to be functional analysis. Functional analysis is the explanation of a capacity of a system in terms of organized sub-capacities or sub-functions. Functional analysis is traditionally seen as distinct and autonomous from mechanistic explanation because functional analysis may omit information about components performing the sub-functions.

In order for the mechanistic account to get properly off the ground, the relation between functional analysis and mechanistic explanation must be clarified. I argue that, insofar as functional analysis is less than a complete mechanistic explanation because it omits information about components, it is a mechanism sketch to be completed by adding information about components.² (Side note: from the conclusion that functional analysis is a mechanism sketch it doesn’t follow that computational explanation is always a mechanism sketch. It may or may not be depending on whether it includes the components.)

Martin Roth (this issue) replies that capacities admit of both horizontal and vertical explanations. Horizontal explanations analyze a capacity in terms of organized sub-capacities; they are functional analyses. Vertical explanations provide implementation details; they are mechanistic explanations. The vertical details can help identify the correct analysis, but they do not add to its explanatory power. Therefore, functional analysis remains explanatorily autonomous from mechanistic explanation.³

My point of disagreement is small but crucial. Horizontal explanations involve more than organized sub-capacities; they also involve components possessing those sub-capacities. Capacities plus components amount to mechanistic explanations. Therefore, even horizontal explanations are mechanistic. Since horizontal explanations are mechanistic, a fortiori they cannot be autonomous from mechanistic explanations.

Considering an example might help. Roth mentions mousetraps, which are a classic example. How does a horizontal explanation of a mousetrap go? It depends on the kind of mousetrap! There are snap traps, electric traps, glue traps, bucket traps, and more. As their names indicate, different kinds of mousetraps involve different kinds of component (spring-loaded bars, electrical circuits, glue, buckets) possessing different sub-capacities (releasing enough force to kill a mouse, releasing enough electricity to kill a mouse, gluing a mouse down, having walls too deep for a mouse to escape from). There is no such thing as *the* functional analysis of mousetraps. For each type of mousetrap, there is a corresponding horizontal explanation involving specific types of component and their specific sub-capacities.

Sometimes, defenders of functional analysis as distinct and autonomous from mechanistic explanation insist that components can be individuated functionally, by their capacities alone. Even so, my point stands. Any functional individuation puts constraints on the structures that can perform it and, vice versa, there is no structure that can perform all functions. In our example, only certain kinds of structure can function as snaps, electrical circuits, glue, or buckets. That’s not to say that all implementational details must be included in a horizontal explanation. I agree with Roth that horizontal explanation abstracts away from, and can be fully explanatory without, lower-level details.⁴ Still, horizontal explanation remains mechanistic.

Instead of thinking of implementation as a relation between functional analysis and mechanistic explanation, as Roth seems to do, I think it’s more accurate and helpful to think of levels of mechanistic organization implementing one another.⁵ Lower levels of mechanistic organization implement higher levels. Higher levels are aspects of lower levels and include the most relevant causes that explain a phenomenon. But this does not lead from mechanistic explanation to functional analysis as a distinct type of explanation; it simply leads from lower to higher levels of mechanistic organization.

One confusing feature of this debate is that, at least in psychology and neuroscience, we tend to focus on *computational* explanation. Computation is *medium independent*, meaning that it abstracts away from most aspects of the physical medium in which the process is realized (except certain degrees of freedom and organizational relations). Even in computational explanation, however, we can't fully explain a capacity, even at the highest level of organization, without specifying the components that perform the relevant sub-capacities. The illusion that we can may be due to the special case in which the same, versatile component performs all the sub-capacities.

In some computational explanations, there is either a central processing unit or a computing human who follows an algorithm by performing all the needed operations. This makes it sound like components don't matter to functional analysis. But we shouldn't let the possibility that one component possesses all the sub-capacities mislead us into thinking that we have stumbled upon a distinct type of explanation. It's just the special case in which one component does all the work. The general type of explanation is still mechanistic, at one level of organization. In this case, there is just one component. Typically, there are many types of component, each of which specializes in one or a few of the sub-capacities that explain the capacity of the system. Either way, horizontal explanations are mechanistic. A fortiori, they are not autonomous from mechanistic explanations.

2. TELEOLOGICAL FUNCTIONS

The mechanistic account relies on the notion of teleological function, which I explicate as a stable contribution to a goal of organisms belonging to a biological population (I will omit "teleological" from now on).⁶ In turn, organisms' goals divide into objective and subjective goals. The objective goals of organisms include survival and inclusive fitness. The subjective goals of organisms are those the organisms choose for themselves. This account applies to the functions of both organismic traits and artifacts. Biological traits have the function to contribute to the goals of organisms that possess them. Artifacts have the function to contribute to the goals of organisms that make and use them.

This goal-contribution account has an important advantage over the traditional selectionist account. The selectionist account is that functions are selected effects—effects selected for through a process such as evolution by natural selection.⁷ Given the selectionist account, knowing a trait's function requires knowing its selection history. In the case of many biological traits, especially psychological traits, it's very difficult to know their selection history. Yet we often attribute functions correctly without knowing the selection history of the traits that possess them. How do we do it, and what are functions such that we can do it?

The goal-contribution account answers as follows. Functions may or may not be selected for. That's why we don't have to know their selection history in order to attribute them. Instead, what we need to find is the stable contribution that a trait (or artifact) makes to the goals of organisms. That stable contribution is the trait's (or artifact's) function.

Nico Orlandi (this issue) asks, aren't organisms' objective goals dependent on organisms' history, and doesn't this pose the same epistemic challenge for the goal-contribution account that the selectionist account faces? If so, the goal-contribution account has no advantage over the selectionist account.

As I said, the objective goals of organisms include survival and inclusive fitness. These goals are essential to living organisms in the sense that if all organisms cease to pursue them, organisms will leave no descendants and go extinct. This is why survival and inclusive fitness are goals. They are a biological imperative. They must be pursued on pain of extinction.

Since life has a historical origin, we might want to say that survival and inclusive fitness have a historical origin too. In addition, each species evolves its own traits, which allow its members to pursue survival and inclusive fitness in their own species-specific ways. Nevertheless, we do not need to know a species' evolutionary history to discover the specific ways in which, right now, its traits contribute to objective goals such as survival and inclusive fitness. We can discover those contributions by observing and experimenting on current organisms. Because of this, the goal-contribution account of functions retains its advantage over the selectionist account.

3. MEDIUM INDEPENDENCE

The mechanistic account relies on the notion of medium independence, which I explicate as a higher-level property defined solely in terms of degrees of freedom and their organization. For example, digital computations are defined over strings of digits. I define a (physical) digit as a variable that can take a finite number of stable values, which can be reliably distinguished from other values, concatenated into strings, and processed by a physical system in accordance with a rule. This notion of digit is medium independent. It does not specify any more concrete physical properties of digits. Therefore, medium independence is a stronger condition than multiple realizability. Multiple realizability does specify concrete physical properties—for example, *catching mice*. A medium-independent property is multiply realizable but the converse need not hold. For example, being a mousetrap is not a medium-independent property precisely because all mousetraps must handle mice, and mice are a specific type of physical "medium."

In response, Orlandi (this issue) worries that the notion of medium independence is too liberal and does not pose enough constraints on its lower-level realizers. If so, the mechanistic account of computation might count too many systems as computing.

To address this worry, we should realize that the notion of medium independence is not the only constraint on physical computation. For a system to count as computing, it must also be a mechanism with teleological functions and it must manipulate its inputs and internal states in accordance with a rule. The way medium independence comes into play is as follows.

For a mechanism to count as a computing system, its teleological function must be defined in terms of a rule for manipulating medium-independent vehicles. Thus, all the elements of the account work together to constrain the systems involved. The relevant type of medium-independent vehicle defines the structural constraints that must be satisfied by a mechanism in order to count as something that performs computations over such vehicles, provided that it manipulates such vehicles *because* that is its function *and* does so in accordance with a rule.

Okay, but what are the constraints imposed by medium independence? It depends on the type of vehicle. In the case of strings of digits, the constraints include the following. The system must possess components that maintain a finite number of distinguishable stable states. There must be enough components to store all the needed digits. The digits must remain stable for a long enough time that the system can process them successfully. The system must possess components that process such stable states in accordance with the relevant rules. The components must be organized so as to respect the ordering of the strings. Finally, the components must be synchronized so that their states update without disrupting the ordering of the strings. Further constraints are imposed by the functions to be computed and the architecture for computing such functions.

Mutatis mutandis, other types of medium-independent vehicle impose their own constraints on systems that manipulate them. In conclusion, medium independence gives rise to especially abstract levels of mechanistic organization. Nevertheless, computing systems remain pluralities of organized components performing sub-functions in such a way that they produce the capacities of the system they compose. That is, computing systems remain mechanisms.

4. COMPUTING FUNCTIONS

The mechanistic account says that the function of computing mechanisms is to process medium-independent vehicles according to a rule. Frances Egan (this issue) finds that implausible. She points out that computing mechanisms have functions such as parsing, adding, producing speech, etc. She ventures that manipulating medium-independent vehicles is the common *means* by which computing mechanisms perform their functions. That last point is correct, and it's not an accident. Why is it that all computing mechanisms process medium-independent vehicles? Because computing a function is a medium-independent notion. Medium independence is part of the essence of computing; it's a necessary property of all physical computations.

Of course, computational processes also have more specific functions than processing medium-independent vehicles according to a rule. Those are the functions that Egan focuses on: parsing, adding, and so forth. What they all have in common is that they are defined in a medium-independent way.

Here is another way to put the point. Computation is a mathematical notion. The mathematics of computation

abstracts away from most physical properties, except whatever degrees of freedom and organizational relations between them are needed to embody an encoding of the function being computed. But that's what medium independence amounts to. Therefore, if a physical system has the function of computing, such a function is defined in a medium-independent way.

But isn't this just a version of the mapping account of physical computation? And doesn't the mechanistic account need the mapping account to tell it which function is being computed? These are the other questions raised by Egan (this issue).

Here is Egan's version of the mapping account:

[A] physical system *S* computes a function *F* just in case:

- (i) There is a mapping from (equivalence classes of) *physical states* of *S* to the arguments and values of *F*, such that:
- (ii) Causal state-transitions between the *physical states* mirror the formal dependency relations between the arguments and values of *F*. (Egan, this issue, emphasis mine)

Egan argues that the mechanistic account needs to be supplemented by this sort of mapping account in order to identify the function computed by the system. She also argues that this sort of mapping account satisfies two important desiderata. First, it is extensionally adequate—that is, it does not count too many systems as computational. Second, it can account for miscomputation: miscomputation occurs when the system's output fails to map to the relevant value of the function.

Egan is right to this extent: there is a mapping from some of the *computational* states of a physical system to the arguments and values of the function the system computes. She is also right that the system miscomputes when its output fails to map onto the correct value of the function being computed. The problem, which the mapping account does not solve, is identifying the computational states of a physical system. Without that, the relevant mapping cannot be constructed and miscomputation cannot occur.

The mapping account refers to equivalence classes of physical states. Which physical states? Which equivalence classes? Every physical system has lots and lots of microstates, which can be grouped into equivalence classes in many ways. To see why the mapping account is insufficient, consider an arbitrary physical system. Group its microstates into equivalence classes that respect the causal relations between microstates. Label each equivalence class with a symbol from your favorite computational formalism. You have just constructed a mapping from some equivalence classes of physical states to the arguments and values of a function. By respecting the causal relations between microstates, you have prevented lots of spurious computational descriptions and avoided the most pernicious versions of noncomputationalism.

Have you identified a genuine computational system yet? Have you identified the system's computational states? Unfortunately, no. Computational states are equivalence classes of microstates, but most equivalence classes of microstates are not computational states. You were looking for a computational explanation of your system, but all you got is a mere computational *model*. A computational model is not necessarily a computational explanation.⁸

If the system you picked was a paradigmatic noncomputational system—something like a river, a tornado, or a comet—mapping equivalence classes of its microstates to the arguments and values of a function does not give you the computational states of the system, for the system has no computational states at all! It cannot miscompute because it does not compute. Any of its state transitions can be mapped to the arguments and values of a function. All the mapping gives you is a computational model. Just because a system has a computational model, it doesn't follow that the system itself performs computations.

What about cognitive systems? As Egan points out, cognitive scientists begin with specific capacities such as language processing and problem solving, and then they try to find computational explanations for them. Doesn't this lead to genuine computational explanations, which give rise to the possibility of miscomputation? Sure, but it remains to be seen what it takes for a cognitive system, like any other physical system, to be computational. The mapping account says that all it takes is a mapping from equivalence classes of physical states to the arguments and values of a function. That's not enough to identify genuine computational states.

Even in genuine computational systems there are lots and lots of microstates, and equivalence classes thereof, that do not belong to any computational states. For example, the microstates of your computer's battery, fan, and case do not belong to any computational states. What's more, even many microstates of computing components such as logic gates do not belong to any computational states, because they occur at computationally irrelevant times. Therefore, if we are looking for a system's computational states, the last thing we should do is simply to construct a mapping from equivalence classes of microstates to the arguments and values of a function. Before we construct such a mapping, we need to identify the genuine computational states.

The mechanistic account tells you how to identify computational states within physical systems by constructing the relevant mechanistic explanation. First, set aside any system that is not a functional mechanism—any system that lacks (teleological) functions. That rules out rivers, tornadoes, comets, and the like. Second, set aside any system whose functions are not defined in terms of processing vehicles defined in a medium-independent way. That rules out most biological systems and artifacts—stuff like hearts, livers, and vacuum cleaners. Third, set aside any system that does not follow a rule—stuff like random "number" generators. At this point, you've more or less identified the class of physical computing systems. You still have to identify their computational states.

To do that, you have to look at how a computing system performs its primary function—that of processing a medium-independent vehicle in accordance with a rule. If you know the type of vehicle and the rule, you know the function the system computes. But your job is not done. What you need to do is provide a mechanistic explanation of the system's behavior. Where do the inputs enter and the outputs exit? What are the components that store and process the vehicles? How are the vehicles manipulated? More generally, what is the mechanistic organization through which the system performs its computational function? By answering these questions, we can identify genuine computational states and state transitions. Through this process, we can construct genuine computational explanations. We can also identify state transitions between computational states that do not follow the relevant rule, and therefore amount to miscomputation. According to the mechanistic account, this is what cognitive scientists and computer engineers do when they explain how physical systems compute.

In summary, the mapping account is not enough to identify genuine computational states and computational state transitions. It counts too many physical systems as computational and does not really account for miscomputation. Nevertheless, if a physical system is a computing mechanism that is designed correctly, is functioning correctly, is programmed correctly, and is used correctly, there is certainly a mapping from its computational states to the arguments and values of the function it computes. That's the nugget of truth within the mapping account.

Does that mean that the mechanistic account needs to be supplemented by the mapping account in order to identify the function being computed? Not really. The function being computed is a function from input computational states and internal computational states to output computational states. Once the system is explained mechanistically, the function being computed is already included in the computational explanation. There is nothing wrong with pointing out that the computational states map to the arguments and values of a mathematical function, but that does not add any new information to our explanation.

5. QUANTUM MECHANISMS

The mechanistic account relies on the notion of, well, mechanism. A mechanism is a plurality of organized components that, collectively, produce a phenomenon. John Symons (this issue) objects that mechanisms are difficult to square with quantum mechanics, one of our fundamental physical theories. He implies that the mechanistic account is committed to local causal interactions between isolable parts with definite properties. But every physical system, including computing systems, is ultimately made of quantum mechanical systems. In some cases, Symons adds, quantum effects may be relevant to the performance of computing components that are already present within our artifacts. Quantum systems need not have isolable parts, such parts need not have definite properties (because of Heisenberg's uncertainty principle), and the interactions between such parts need not be local (because of entanglement and the wave-particle duality). Therefore,

Symons concludes, insofar as mechanistic explanation requires local causal interactions between isolable parts, it will be unable to explain computation (or anything else, for that matter) once we reach the quantum level.

The literature on mechanisms is vast. Perhaps there are mechanists who are committed to local causal interactions between isolable parts with definite properties. Typical examples of mechanisms in much recent literature are certainly classical (i.e., nonquantum). In my writing, I usually focus on classical (i.e., nonquantum) computation.⁹ But my focus on classical computation does not imply that quantum systems cannot be mechanistic. The notion of mechanism is flexible and open ended enough to encompass quantum systems. After all, it's quantum *mechanics*! Several issues should be distinguished.

A first issue is that although everything is made out of quantum mechanical parts, typical macroscopic objects behave classically. Why is that? A plausible explanation is decoherence.¹⁰ Decoherence occurs when a complex quantum mechanical system interacts with enough portions of its environment to become entangled with them. Enough entanglements with enough things make the superposition between quantum states undetectable and suppresses exotic quantum effects such as quantum mechanical interference. That's probably why most of the objects we observe, including ordinary computers, behave classically. As long as a system and its components behave classically, quantum mechanics is not especially relevant and poses no special problem.

Eventually, though, if we keep explaining a phenomenon mechanistically and descend to more and more microscopic levels, we will reach a level at which quantum effects become relevant. At this stage, we can appeal to what Kuhlmann and Glennan call *nonclassical mechanistic explanation*.¹¹ For present purposes, nonclassical mechanistic explanation is explanation of a phenomenon in terms of components that exhibit nonclassical features, such as parts that are not isolable from one another and may even lack definite locations, lack definite properties (in the sense that they satisfy Heisenber's uncertainty principle), etc. As to nonlocal effects due to entanglement, they often play no role in a quantum system's dynamics. When entanglement affects dynamics, Kuhlman and Glennan argue, we finally meet the limits of the mechanistic program. Nonlocal effects due to quantum entanglement lack a mechanistic explanation. Another place where mechanistic explanation stops is where we find truly basic components—elementary particles. Since they lack parts, their behavior has no mechanistic explanation. This is not a flaw: mechanistic explanation, like explanation simpliciter, has to stop somewhere. Mechanistic explanation stops when there are no smaller parts whose sub-capacities explain the capacities of the whole.

Nevertheless, nonlocal effects and any other basic physical actions can partake in mechanistic explanation. An especially relevant case is quantum computation. Quantum computation is like classical digital computation except that, instead of operating on digits (typically, bits), it operates on qudits (typically, qubits). Roughly, qudits

are d-dimensional quantum systems which can be put into superpositions of computational basis states and also be entangled with one another. In some cases, exploiting these quantum mechanical features can lead to greater efficiency in the computation.

Armond Duwell has looked at an especially exotic type of quantum computing systems: Measurement Based Quantum Computers (MBQCs).¹² In addition to superposition, MBQCs exploit entanglements between qubits. Unlike ordinary classical *and* quantum computing systems, in which digits or qubits flow through a circuit, MBQCs operate by performing a series of measurements on an array of entangled qubits. In spite of the unusual and highly nonclassical structure of MBQCs, Duwell shows how to apply the mechanistic account to them.

Duwell points out that, even though qubits do not flow through MBQCs the way they flow through more ordinary computing systems, the qubits in a MBQC array are correlated in such ways, due to their entanglement, that a flow of qubits is not necessary to complete a computation. Instead, the whole array of qubits is the computational vehicle and, given the entanglement of the qubits, measurements affect not only the measured qubits but also all the qubits entangled with them. What's more, it is the *function* of the measurement apparatus to affect the qubits in this way, and it is a *function* of the entanglement of the qubits to be affected in this way so as to produce the correct output. Given that other components of the system have other functions that contribute to the computation, the whole system is a functional mechanism whose function is processing a medium-independent vehicle (qubit array) in accordance with a rule. Thus, the mechanistic account applies to MBQCs.

In conclusion, quantum mechanics requires adjustments to the mechanistic account of physical computation but is compatible with it.

6. CONCLUSION

Thanks to the thoughtful commentaries by Roth, Orlandi, Egan, and Symons, I did my best to clarify why I believe functional analyses are mechanism sketches, teleological functions are regular contributions to the goals of organisms, computation is medium independent, the function of computing mechanisms is to process medium-independent vehicles according to a rule, and quantum systems can be nonclassical mechanisms. There is room for more detailed philosophical work on these topics. I greatly appreciate this opportunity to get started. I hope others will contribute as well.

ACKNOWLEDGMENTS

Thanks to Marcello Guarini for organizing the APA session from which this exchange derives, and to Peter Boltuc for inviting us to publish our contributions in this newsletter. Thanks to Armond Duwell for help with quantum computation.

NOTES

1. Piccinini, *Physical Computation: A Mechanistic Account*; for related accounts, see Kaplan, "Explanation and Description in Computational Neuroscience"; Fresco, *Physical Computation and Cognitive Science*; Milkowski, *Explaining the Computational*

Mind; Coelho Mollo, "Functional Individuation, Mechanistic Implementation: The Proper Way of Seeing the Mechanistic View of Concrete Computation."

2. Piccinini, *Physical Computation*, Chap. 5; Piccinini and Craver, "Integrating Psychology and Neuroscience: Functional Analyses as Mechanism Sketches."
3. Cf. Roth and Cummins, "Neuroscience, Psychology, Reduction, and Functional Analysis."
4. Boone and Piccinini, "Mechanistic Abstraction."
5. Boone and Piccinini, "The Cognitive Neuroscience Revolution."
6. Maley and Piccinini, "A Unified Mechanistic Account of Teleological Functions for Psychology and Neuroscience."
7. Garson, *What Biological Functions Are and Why They Matter*.
8. Piccinini, *Physical Computation*, Chap. 4.
9. The notion of classical computation as opposed to quantum should not be confused with the notion of classical computation in the sense of digital computation operating on language-like vehicles (Fodor and Pylyshyn, "Connectionism and Cognitive Architecture"). Here we are discussing classical versus quantum mechanical physical systems and their processes.
10. Kuhlmann and Glennan, "On the Relation Between Quantum Mechanical and Neo-Mechanistic Ontologies and Explanatory Strategies."
11. *Ibid.*, Section 5.
12. Duwell, "Exploring the Frontiers of Computation: Measurement Based Quantum Computers and the Mechanistic View of Computation."

REFERENCES

- Boone, W., and G. Piccinini. "Mechanistic Abstraction." *Philosophy of Science* 83, no. 5 (2016a): 686–97.
- Boone, W., and G. Piccinini. "The Cognitive Neuroscience Revolution." *Synthese* 193, no. 5 (2016b): 1509–34.
- Coelho Mollo, D. "Functional Individuation, Mechanistic Implementation: The Proper Way of Seeing the Mechanistic View of Concrete Computation." *Synthese* 195, no. 8 (2018): 3477–97.
- Duwell, A. "Exploring the Frontiers of Computation: Measurement Based Quantum Computers and the Mechanistic View of Computation." In *Turing 100: Philosophical Explorations of the Legacy of Alan Turing*, edited by A. Bokulich and J. Floyd. Boston Studies in the Philosophy and History of Science, vol. 324, pp. 219–32. Springer, 2017.
- Fodor, J. A., and Z. W. Pylyshyn. "Connectionism and Cognitive Architecture." *Cognition* 28 (1988).
- Fresco, N. *Physical Computation and Cognitive Science*. New York: Springer, 2013.
- Garson, J. *What Biological Functions Are and Why They Matter*. Cambridge: Cambridge University Press, 2019.
- Kaplan, D. M. "Explanation and Description in Computational Neuroscience." *Synthese* 183, no. 3 (2011): 339–73.
- Kuhlmann, M., and Glennan, S. "On the Relation Between Quantum Mechanical and Neo-Mechanistic Ontologies and Explanatory Strategies." *European Journal for Philosophy of Science* 4, no. 3 (2014): 337–59.
- Maley, C. J., and G. Piccinini. "A Unified Mechanistic Account of Teleological Functions for Psychology and Neuroscience." In *Explanation and Integration in Mind and Brain Science*, edited by D. Kaplan, 236–56. Oxford: Oxford University Press, 2017.
- Milkowski, M. *Explaining the Computational Mind*. Cambridge, MA: MIT Press, 2013.
- Piccinini, G. *Physical Computation: A Mechanistic Account*. Oxford: Oxford University Press, 2015.
- Piccinini, G., and C. Craver. "Integrating Psychology and Neuroscience: Functional Analyses as Mechanism Sketches." *Synthese* 183, no. 3 (2011): 283–311.
- Roth, M., and R. Cummins. "Neuroscience, Psychology, Reduction, and Functional Analysis." In *Explanation and Integration in Mind and Brain Science*, edited by D. Kaplan. Oxford: Oxford University Press, 2017.

Philosophical Insights from Computational Studies: Why Should Computational Thinking Matter to Philosophers?

Gary Mar

STONY BROOK UNIVERSITY

(With Edward Zalta and Aydin Mohseni)

In April the Committee on Philosophy and Computers sponsored two panels at the APA Pacific meeting in Vancouver, Canada. The panel on "Data Ethics" was chaired by Joshua August Skorburg, who also delivered a paper along with papers by Shannon Vallor and Colin Koopman. The second panel, "Philosophical Insights from Computational Studies: Why Should Computational Thinking Matter to Philosophers?," was chaired by Gary Mar, who delivered a paper along with papers by Aydin Mohseni and by Edward Zalta. What follows are two abstracts and one summary of the latter session.

Despite being scheduled during the next to the last session on the last day of the conference, the panel on "Philosophical Insights from Computational Studies" was well attended and provoked a spirited debate. This is perhaps evidence that the APA Committee on Philosophy and Computers has progressed well beyond raising pedagogical questions about the use of technology in the classroom and is now, as it is soon to be disbanded, addressing fundamental questions about how computationalism is profoundly transforming the nature philosophical research and the kinds of questions that can now be addressed.

1. ON THE EMERGENCE OF MINORITY DISADVANTAGE: TESTING THE RED KING HYPOTHESIS (REPORT WITH ABSTRACT).

Aydin Mohseni, a graduate student in the Department of Logic and Philosophy of Science at the University of California, Irvine after having completed an MA in Logic, Computation, and Methodology at Carnegie Mellon University, presenting joint research conducted with Cailin O'Connor (University of California, Irvine) and Hannah Rubin (Notre Dame University).

The cultural Red King effect was first described by political philosopher Justin Bruner, who uses evolutionary game theoretic methods to show how minority groups can be disadvantaged in the emergence of bargaining conventions solely by dint of their group size.¹ As he shows, in groups with completely symmetric preferences, abilities, and resources, minority status alone can increase the likelihood that individuals end up with fewer economic resources. The driver behind this effect is a learning asymmetry between minority and majority groups. While minority members commonly meet their out-group, the reverse is not true. As a result, members of a minority will more quickly learn to interact with their out-group. In situations where this learning is about bargaining interactions, this often proves

disadvantageous. Low, accommodating demands tend to be safer in bargaining interactions, meaning that swift learners should adopt these demands. Once this is done, members of the majority group can take advantage of this accommodation.

Subsequent work has shown that this effect arises robustly in cultural evolutionary models.² Given the simplicity of these models, though, a further question arises: Can the cultural Red King really occur in human groups? If so, there are important consequences for social and political philosophy. Future work will be directed at better understanding the empirical conditions under which cultural red king-type effects may obtain and contribute to the emergence of inequitable conventions.

The populations are playing a Nash demand game with two strategies: aggressive demand, and passive demand. The x-axis reflects the proportion of individuals in the first population who are playing each strategy while the y-axis reflects the proportion in the other population. There are two attracting states corresponding to the possible outcomes of evolution: the state where the first population adopts the aggressive strategy and the second population the passive strategy, and the dual state where the strategies are flipped. When the two populations' rates of evolution are equal, the basins of attraction for each attractor are symmetric (shown in Figure 1). However, when the rates of evolution are unequal, the basin of attraction for the attractor where the slower evolving population plays the aggressive strategy grows larger (shown in the right figure). This is the red king effect: the slower runner wins the race. The cultural red king effect then consists in observing that a functionally equivalent effect can be produced via population size differences conducting to differences in rates of interaction and so to learning rates.

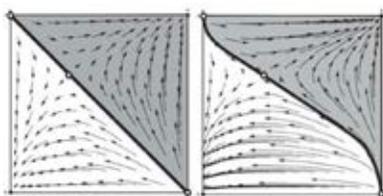


Figure 1. The phase portrait captures the dynamics of two evolving populations as they interact.

2. METAPHYSICAL INSIGHTS FROM COMPUTATIONAL STUDIES (EXTENDED ABSTRACT)

Edward Zalta, Senior Research Scholar at Stanford's Center for the Study of Language and Information (CSLI) and the winner of the 2016 Jon Barwise Prize.

The foundations of object theory have evolved in a number of interesting new ways since its implementation in Isabelle/HOL by Christoph Benzmueller and Daniel Kirchner. Not only has a potential "back-door paradox" (i.e., a loophole permits the reintroduction of a known paradox) been identified and avoided, but the language of object theory itself

has been improved by eliminating the syntactic category of "propositional formulas" (i.e., formulas containing encoding subformulas). λ -expressions $[\lambda x_1 \dots x_n \phi]$ whose matrix ϕ has encoding subformulas are now well-formed, though they are guaranteed to denote relations only if ϕ is encoding- and description-free. Moreover, the language of object theory has been generalized to allow n-ary encoding formulas. These changes to the language have led to changes in the definitions, axioms, and theorems of the theory, primarily by extending the free logic for definite descriptions to λ -expressions. Two new axioms have been added: one ensures that any relation necessarily equivalent to an existing relation also exists, and a second ensures that a formula denotes a proposition precisely when its nominalization denotes a proposition. Finally, the axiom for asserting that there are contingently nonconcrete objects has been refined.

These changes have metaphysical implications:

1. Not only can identity for objects and relations be defined in terms of predication, but existence can be defined for objects and relations in terms of predication as well. (By contrast, free logics either reduce existence to a primitive notion of identity or take existence as a primitive.)
2. The new axioms also allow us to prove that *every* formula denotes a proposition (though not every open formula can be transformed into a denoting λ -expression). This, in turn, extends the fundamental theorem of world theory, which previously applied only to formulas without encoding subformulas: the fundamental theorem now governs every formula whatsoever: necessarily ϕ if and only if ϕ is true in every possible world.
3. Finally, with a refined axiom that asserts the *possible* existence of a nonconcrete but actually concrete object, the theory still guarantees that at least two possible worlds and four propositions exist, but also that there are at least 16 properties.

3. GÖDEL, TURING, AND TIME: A COMPUTATIONAL PHILOSOPHY OF MATHEMATICS (SUMMARY)

The philosophical views of Kurt Gödel and Alan Turing are often presented as posing an irreconcilable dichotomy: Gödel's *platonism* asserts that the human mind is creative surpassing the capacity of any single Turing machine whereas Turing's *computationalism* leads him to the mechanistic view that the human brain is essentially a Turing machine. Gödel and Turing, however, are more philosophically honest and skeptical than their subsequent followers who have posed the dichotomy as a stereotypical choice between the *creative mind* versus *mechanistic computability* in the philosophy of mathematics. It is argued that whereas Gödel and Turing's views on the philosophy of mind might be incompatible, their philosophies of *mathematics* are not irreconcilably inconsistent but, in fact, are complementary.

This thesis faces three immediate objections. First, platonism is about abstract, universal, and *timeless objects* whereas computations are *processes* that take place *in time*. In his conversations with Hao Wang, Gödel remarked:

The real argument for objectivism is the following. We know many general propositions about natural numbers to be true (2 plus 2 is 4, there are infinitely many prime numbers, etc.) and, for example, we believe that Goldbach's conjuncture makes sense, must be either true or false without there being any room for arbitrary convention. Hence, there must be objective facts about natural numbers. *But these objective facts must refer to objects that are different from physical objects because, among other things they are unchangeable in time.*³

Secondly, in his Josiah Willard Gibbs Lecture "Some Basic Theorems on the Foundations of Mathematics and their Implications" to the meeting of the American Mathematical Society at Brown University in 1951, Gödel proposed the following disjunctive conclusion as "inevitable":

*Either mathematics is incompleteable in this sense, that its evidence axioms can never be comprised in a finite rule, that is to say, the human mind (even within the realm of pure mathematics) infinitely surpasses the powers of any finite machine, or else there are absolute unsolvable Diophantine problems. . . ."*⁴

Thirdly, Turing's PhD thesis "Systems of Logic Based on Ordinals" under the direction of Alonzo Church attempts to overcome the essential incompleteness of Gödel's theorem, the basis for Gödel's disjunction. Turing notes that "in pre-Gödel times it was thought by some that it would probably be possible to . . . [replace] intuitive judgements . . . by a finite number of . . . rules [of formal logic]. The necessity of intuition would then be entirely eliminated."⁵ Gödel showed for that any formal system S powerful enough to represent arithmetic, there is a statement G which is true but which the system is unable to prove. Turing's thesis considers the possibility of adding G as an additional axiom to the system and then iterating the process to infinity, creating a system with an infinite set of axioms. Turing explains his motivation:

In our discussions, however, we have gone to the opposite extreme and eliminated not intuition but ingenuity, and this in spite of the fact that our aim has been in much the same direction. We have been trying to see how far it is possible to eliminate intuition, and leave only ingenuity. We do not mind how much ingenuity is required, and therefore assume it to be available in unlimited supply. In our metamathematical discussions we actually express this assumption rather differently. We are always able to obtain from the rules of a formal a method for enumerating the propositions proved by its means. We then imagine that all proofs take the form of a search through this enumeration for the theorem for which a proof is desired. In this way ingenuity is replaced by patience."⁶

On the contrary, Gödel's and Turing's philosophies of mathematics are not irreconcilably *inconsistent* but *complementary*. This thesis can be proved in five ways. Gödel's endorsement of platonism in print begins with his "Russell's Mathematical Logic" and is reaffirmed in "What Is Cantor's Continuum Problem?" The 1964 supplement for the latter contains Gödel's most full-fledged, frequently quoted, espousal of platonism. Alan Turing's *computational* point of view is contained in his two most famous articles "On Uncomputable Numbers with an Application to the *Entscheidungsproblem*" and "Computing Machinery and Intelligence," that sets forth the now famous Turing Test for answering the question "Can computers think?"

First, Gödel and Turing were "better philosophers" than their followers insofar as their published writings express views that are more nuanced and sceptical than the presentations of their views by ardent and partisan (especially, as regards their views on the philosophy of mind) followers. In his invited lecture "The present situation in the foundations of mathematics" to a joint meeting of the Mathematical Association of America and the American Mathematical Association in Cambridge, Massachusetts, Gödel admitted:

The result of the preceding discussion is that our axioms, if interpreted as meaningful statements, necessarily presuppose a kind of Platonism, which cannot satisfy any critical mind and which does not even produce the conviction that they are consistent.⁷

Turing's Lecture for *London Mathematical Society* emphasizes the growing importance of human interest in the philosophy of mathematics:

As regards mathematical philosophy, since the machines will be doing and more mathematics themselves, the centre of gravity of the human interest will be driven further and further into philosophical questions of what can in principle be done etc.⁸

In his introductory remarks to Turing's "Computing Machinery and Intelligence" in the Cooper and Leeuwen Centenary anthology *Alan Turing: His Work and Impact*, Gregory Chaitin in "Mechanical Intelligence versus Uncomputable Creativity" notes: "It is a delightful paradox that Turing argues that we are machines while all the while emphasizing the importance of what machines cannot do. Like a good philosopher, he cannot help seeing the good arguments on both sides. He thus provides ammunition to both parties."⁹

Secondly, a computationalist pedagogy can provide an account of the learnability of mathematics, which poses a central problem for platonism. The confirmed platonist G. H. Hardy described

the function of a mathematician [as] . . . simply . . . observ[ing] the facts about his own intricate system of reality, that astonishingly beautiful complex of logical relations which forms the subject-matter of his science, as if he were an explorer looking

at a distant range of mountains, and to record the results of his observations in a series of maps, each of which is a branch of pure mathematics. . . .¹⁰

In support of his view, Hardy liked to tell a famous anecdote:

He [Ramanujan] could remember the idiosyncrasies of numbers in an almost uncanny way. It was Littlewood who said that every positive integer was one of Ramanujan's personal friends. I remember once going to see him when he was ill at Putney. I had ridden in taxi cab number 1729 and remarked that the number seemed to me rather a dull one, and that I hoped it was not an unfavorable omen. "No," he replied, "it is a very interesting number; it is the smallest number expressible as the sum of two [positive] cubes in two different ways."¹¹

A computational account of Ramanujan's insights is a more plausible and pedagogically sound alternative to Hardy and Littlewood's mystical musings. If you compute a table of cubes for the first dozen integers perhaps you can discover Ramanujan's Taxicab number for yourself:

1	1	5	125	9	729
2	8	6	216	10	1000
3	27	7	343	11	1331
4	64	8	512	12	1728

Thirdly, mathematical platonism traces its lineage back to the *Meno*, which raises well-known paradoxes about the learnability of mathematics. Philosophically, platonism is about abstract, universal, timeless *objects* whereas computations are *calculations*, processes that take place in time. This dichotomy leads to two fundamentally different approaches to the mathematical enterprise. Philip Davis and Reuben Hersh in *The Mathematical Experience* draw a distinction between *dialectical* [or what I shall call *deductivist*] and *algorithmic* mathematics.¹² According to the former conception, mathematics is a rigorous logical science in which propositions about platonic objects are either true or false. According to the latter conception, mathematics is a tool for solving problems by constructing algorithms. With the increasing use of computers, there has been a shift in mathematics from the former conception back to a more constructive or algorithmic point of view.

These two approaches can be illustrated from the first crisis in the foundations of mathematics—the Pythagorean discovery of that the *diagonal* of the unit square (the answer to Socrates's question and the basis of the geometrical theorem in the *Meno*) is *alogon* or irrational. According to a *priori* deductive mathematics, the mathematical enterprise is to discover proofs about eternal mathematical truths and objects (e.g., that the diagonal of the unit square is irrational).

Theorem: $\sqrt{2}$ is irrational.

Proof: Assume that $\sqrt{2} = a/b$, where a and b are reduced to lowest terms.

1. Hence, $2b^2 = a^2$.
2. $2b^2 = a^2$ implies that a is even so $a = 2c$.
3. Hence, $2b^2 = (2c)^2$ and so $b^2 = 2c^2$, which means that b is even.
4. Contradiction. Both a and b being even contradicts a/b was reduced to lowest terms.

The classical proof of this fact—the proof of a negative existential—is what the Cambridge mathematician G. H. Hardy called a “theorem of the first class.” It is interesting to note that a contradiction appears already in step 2, but the continuation of the proof to show the recursive nature of the contradiction is mathematically more elegant.

According an empirical and computational point of view, the mathematical enterprise is about producing calculations to answer mathematical questions (e.g., how can one calculate with increasing accuracy the length of the diagonal of the unit square?) The approach of *algorithmic*, in contrast to *deductivist*, mathematics is not to prove the *non-existence* of a rational representation for $\sqrt{2}$, but to produce an algorithm for converging on its actual value of $\sqrt{2}$. Such an approach is more practically useful than the purely theoretical platonic approach if, for example, you want to build a bridge.



Figure 2. The Babylonian clay tablet YBC 7289 (c. 1800–1600 B.C.) gives an approximation of the square root of 2 in four sexagesimal figures, $1\ 24\ 51\ 10 = 1.414213$, which is accurate to six decimal digits.

One such algorithm is known as *Newton's method*. If $x^2 = 2$, then we may divide both sides of the equation by 2 to obtain

$$x = 2/x .$$

Let's say our estimate for x is slightly incorrect, say *underestimated*, then $2/x$ will be *overestimated*, and vice versa. Therefore, a better estimate than either x or $x/2$ would be their average.

The guiding idea behind Newton's method is captured in the following dynamical system:

$$x_{n+1} = 1/2(x_n + 2/x_n)$$

Newton's method converges quickly. From an initial estimate of 1, for example, we obtain in just four iterations an estimated value that is accurate to 9 digits. This method was known to the ancient Babylonians.

Reflecting for a moment on the *movement* within the classical proof of the irrationality of $\sqrt{2}$, we may wish to countenance not just the *negative existential* conclusion but the logical *dynamics* of the proof. Here there is a geometric connection to the form of proof by induction, loved by Fermat, known as *proof by infinite descent*. As can be easily seen from their symbolic representations, Fermat's Proof (or Disproof) by Infinite Descent is logically equivalent to Strong Mathematical Induction for natural numbers:

Strong Induction: if F holds of x whenever it holds for all numbers less than x, then F holds for *all* numbers:

$$\forall x[N(x) \wedge \forall y[N(y) \wedge y < x \rightarrow F(y)] \rightarrow F(x)] \rightarrow \forall x[N(x) \rightarrow F(x)]$$

Fermat's Proof by Infinite Descent: F if holding of number implies it holds for an even smaller number, then no number has F:

$$\forall x[N(x) \wedge F(x) \rightarrow \exists y[N(y) \wedge y < x \wedge F(y)] \rightarrow \forall x[N(x) \rightarrow \sim F(x)]$$

The classical proof shows that if $\sqrt{2}$ were rational, no "smallest" representation as a fraction could exist (i.e., a fraction reduced to lowest terms). Any attempt to find a "smallest" representation a/b would imply the existence of a smaller one, which is impossible given the nature of natural numbers. Fermat's proof by means of the impossibility of method of infinite descent connects a *computational* way of thinking of irrationality with *fractal* geometry—a method that deploys an *infinite-regress* in much the same way as Zeno's paradoxes.

Proof by Infinite Descent. The following diagram was used by Tennenbaum (1950s) to prove that that is irrational by an argument by "infinite descent."

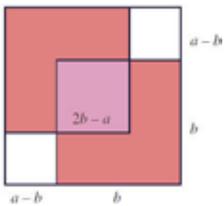


Figure 3. A fractal proof of the irrationality of $\sqrt{2}$ is a geometric counterpart to Fermat's inductive proofs by the impossibility of infinite descent.

If the diagonal of the unit square is rational, then we have $\sqrt{2} = (a/b)$, where a and b are the smallest such units with no common factors. This implies that

$$2b^2 = a^2.$$

Geometrically, this means that the area of the square with side a is equal to two squares with side b . Place the two squares with side b inside the square with side a . Since the areas of the two b squares are

equal to the big a square, by assumption, the light pink square in the center created by the overlapping b squares must be equal in area to the two smaller white squares, which are the remainders uncovered by the overlapping b squares. Notice we have that the pink square must be equal to the areas of the two white squares, which provides a smaller solution to our original problem. Contradiction by infinite regress.

This geometric proof by infinite regression foreshadows its fractal nature and relates dynamical systems which can be used to model the paradoxes to Zeno's paradoxes. The seeds

of this approach that combines platonism with computationalism can be illustrated by transforming the diagram in the *Meno* into a dynamic construction. Take the unit square from the *Meno* and connect the midpoints of the square to construct an interior diamond. The area of this diamond is $\frac{1}{2}$ the original square. Then connect the midpoints of the diamond to construct an interior square, which is $\frac{1}{2}$ the area of diamond or $\frac{1}{4}$ the original square. Keep repeating this construction. What is the sum of the areas of all the nested diamonds and squares? This transformation produces a fractal that geometrically represents the infinite geometric series characteristic of Zeno's dichotomy paradoxes.

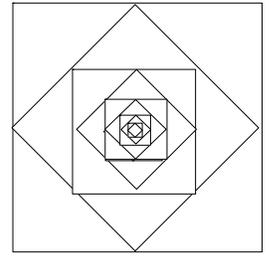


Figure 4. The diagram from the Plato's *Meno* can dynamically be transformed into a fractal representing geometric infinite series of Zeno's dichotomy.

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots$$

This construction is universal. If you begin with any quadrilateral with area equal to 1, one obtains a nested series of parallelograms whose infinite sum areas is given by Zeno's dichotomous geometric series.

Fourthly, Gödel's platonism (i.e., the objectifying of computations) is *required* for Turing's meta-logical proof of the unsolvability of the halting problem, whereas a kind of platonism (i.e., treating an algorithm as a universal) is *presupposed* in the very conception of the research program of AI. Indeed, the dialectic of informal intuition and formalized algorithmic proof is characteristic of theoretical progress in logic. We can state the mutual relationship between platonism and computationalism in a Kantian manner:

Platonism without computationalism is epistemologically empty;

Computationalism without platonism is theoretically blind.

Russell's paradox about the barber who shaves all and only those who don't shave themselves is often treated timelessly. When the problem is stated dynamically about a resident barber in Hilbert's Aleph-Nought Hotel who shaves all and only those residents who have terminating algorithmic schedules, we can obtain a new proof of Turing's Unsolvability of the Halting Problem.

A central thesis of AI can be stated as the analogy that minds are to brains, what software is to hardware. The claim that the same software or algorithm can be instantiated in different brains or in computers presupposes that algorithms not be conceived merely mechanical processes that occur in time but as abstract objects that can be instantiated at different times in different substrata.

Fifthly, combining Gödelian platonism with Turing's computationalism results in a dynamic approach to the semantic paradoxes. Exploring the paradoxes computationally reveals fractal images in the semantics of paradox.¹³ Aristotle's law of non-contradiction states that a given proposition cannot be true and false in the same respect *at the same time*. This approach allows for both Gödelian limitative theorems and the discovery of a menagerie of infinitely complex and chaotic paradoxes:

Paradox is not illogicality, but it has been a trap for logicians: the semantic paradoxes look just a little simpler and more predictable than they actually are. Even in some of the most recent and logically sophisticated work on cyclical regularity in the semantic paradoxes, their deeper and more complex semantic patterns have remained hidden. Our attempt, rather than search for semantic stability or simple patterns within the paradoxes, has been to offer glimpses of the infinitely complex, chaotic, and fractal patterns of semantic instability that have gone virtually unexplored.¹⁴

The logic and philosophy of time is currently experiencing a renaissance across the disciplines because of the widespread use of computers and algorithms to reframe research questions in philosophy. The semantic paradoxes such as the Paradox of the Liar or the Epimenides Paradox seemed to be an entirely different kind of paradox from Zeno's paradoxes of motion. For example, the former *semantic* paradoxes and have values that are *discrete* and *bivalent*, where the latter paradoxes of motion presuppose an infinity of values—either a *countable* infinity such as that required by the assumption of *infinite divisibility* or the infinity of the *continuum* required by *continuous* motion. Solutions to the semantic paradoxes belongs to the *philosophy of logic*, whereas solutions to Zeno's paradoxes belong to the *philosophy of space and time*. Treating paradoxes *computationally* and *dynamically* provides a *unification* between the semantic paradoxes of the liar in the philosophy of logic and Zeno's paradoxes of motion in the philosophy of space and time.

Replies to Objection #1. Classically, platonism is about abstract, universal, and *timeless objects* whereas computations are *processes that take place in time*. However, it has been argued the metatheory of computability requires treating algorithms platonically and timelessly, whereas resolving the epistemological paradoxes about learnability for platonism can benefit from reflection upon processes like calculations that take place in time.

Replies to Objection #2. With characteristic caution, Gödel sometimes prefaced his disjunction with a concession to the possibility of a mechanistic view of mind (*italics mine*):

The human mind is incapable of formulating (or mechanizing) all its mathematical intuitions. I.e.: If it has succeeded in formulating some of them, this very fact yields new intuitive knowledge, e.g. the consistency of this formalism. This fact may be called the "incompleteness" of mathematics.

On the other hand, on the basis of what has been proved so far, it remains possible that there may exist (and even be empirically discoverable) a theorem-proving machine which in fact is equivalent to mathematical intuition, but cannot be proved to be so, nor even proved to yield only correct theorems of finitary number theory.

The second result is the following disjunction: *Either the human mind surpasses all machines (to be more precise: it can decide more number-theoretic questions than any machine) or else there exist number theoretic questions undecidable for the human mind.*

Replies to Objection #3. It should be noted that both Gödel and Turing would have agreed that no *single* machine is sufficient to simulate the mathematician's mind because of Gödel *incompleteness*:

There would be no question of triumphing simultaneously over all machines. In short, then, there might be men cleverer than any given machine, but then there might be other machines cleverer again, and so on.¹⁵

Although Turing's attempt in his PhD to overcome Gödel incompleteness still appeals to ingenuity, which is not formalized, in "Intelligent Machinery," Turing envisions an iterative process of mechanization in which a machine takes the initiative to increasingly incorporate the "residue" of human intuition not previously captured:

If the untrained infant's mind is to become an intelligent one, it must acquire both discipline and initiative. So far, we have been considered only discipline [via the universal machine]. . . . But discipline is certainly not enough in itself to produce intelligence. That which is required in addition we call initiative. This statement will have to serve as a definition. Our task is to discover the nature of this residue as it occurs in man and try to copy it in machines.¹⁶

Here one is reminded of the Gödel program of the search for increasingly more general Axioms of Infinity to settle questions of set theory.

Gödel rejected mechanistic reductionism for the mind and claimed that Turing's analysis committed a *philosophical error*:

A philosophical error in Turing's work. Turing in his 1937 . . . gives an argument which is supposed to show that the mental procedures cannot go beyond mechanical procedures. However, this argument is inconclusive. What Turing disregards completely is that fact that *mind, in its use, is not static, but constantly developing*, i.e., using them, and that more and more abstract terms enter in the sphere of our understanding. There may exist systematic methods of actualizing this development, which could form part of the procedure. Therefore, although at each stage the

number and precision of the abstract terms at our disposal may be *finite*, both (and, therefore, also Turing's number of *distinguishable states of mind*) may converge toward infinity in the course of the application of this procedure. Note that something like this indeed seems to happen in the process of forming stronger and stronger axioms of infinity in set theory. This process, however, today is far from being sufficiently understood to form a well-defined procedure. It must be admitted that the construction of a well-defined procedure which could actually be carried out (and would yield a non-recursive number-theoretic function) would require a substantial advance in our understanding of the basic concepts of mathematics.¹⁷

The scholarly consensus seems to be that Gödel was hasty in attributing the static view of the mind to Turing,¹⁸ where the parentheses indicate that Gödel may have been wrestling with this question as a challenge, not merely to Turing, but to his own views on the matter).¹⁸ Comparing Gödel's notes from 1972 and 1974, Sieg notes some substantive differences and carefully concludes:

I don't fully understand these enigmatic observations, but three points can be made. First mathematical experience has to be invoked when asking the right questions; second, aspects of that experience may be codified in a mechanical procedure and serve as the basis for asking the right questions; third, the answers may involve abstract terms that are introduced by the nonmechanical mental procedure. We should not dismiss or disregard Gödel's methodological remarks that "asking the right questions on the basis of a mechanical procedure" may be part of a systematic method to push forward the development of the mind. Even this every limited understanding allows us to see that Gödel's reflections overlap with Turing's proposal for investigating matters in a broadly empirical and directly computational manner.¹⁹

Combining Gödel's platonistic views with Turing's algorithmic methods in a dialectical and recursive manner provides a means for constructing a hybrid computational philosophy of mathematics.

NOTES

1. Bruner, "Minority (dis) Advantage in Population Games."
2. O'Connor, "The Cultural Red King Effect"; O'Connor and Bruner, "Dynamics and Diversity in Epistemic Communities."
3. Wang, *A Logical Journey: From Gödel to Philosophy*, 71.2., 211; italics mine.
4. Gödel, "Some Basic Theorems on the Foundations of Mathematics and Their Implications," in GCW-III, 310.
5. Turing, "Systems of Logic Based on Ordinals," 82.
6. Ibid.
7. Gödel [*1933o], in GCW-III, 50.
8. Turing's 1947 lecture for London Mathematical Society.
9. Gregory Chaitin in "Mechanical Intelligence versus Uncomputable Creativity," 551.

10. Hardy, "The Theory of Numbers," *Nature* 110 (Sept. 16, 1922): 381.
11. Ibid., Ivii-Iviii.
12. Philip Davis and Reuben Hersh, *The Mathematical Experience*, 199.
13. See Ian Stewart's "A Partly True Story" in *Scientific American*, a popular exposition of research in *Philosophical Computer* [1998].
14. Grim, Mar, and St. Denis, *The Philosophical Computer*, 87. The first sentence is an allusion to G. K. Chesterton, the prince of the paradoxical aphorism.
15. Turing, "Computing Machinery and Intelligence," 445.
16. Turing, "Intelligent Machinery," 21.
17. GCW-II, p. 306.
18. E.g., see Copeland and Shagrir, "Turing versus Gödel on Computability and the Mind"; and Seig, "Gödel's Philosophical Challenge (to Turing)."
19. Seig, "Gödel's Philosophical Challenge (to Turing)," 193.

REFERENCES

Bruner, J. P. "Minority (dis) Advantage in Population Games." *Synthese* (2017): 1–15.

Copeland, B. Jack, ed. *The Essential Turing: The Ideas that Gave Birth to the Computer Age*. Oxford: Clarendon Press, 2004.

Copeland, B. Jack, and Oron Shagrir. "Turing versus Gödel on Computability and the Mind." In *Computability: Turing, Gödel, Church, and Beyond*, edited by B. Jack Copeland, Carl J. Posy, and Oron Shagrir. Cambridge, MA: MIT Press, 2013.

Copeland, B. Jack, Carl J. Posy, and Oron Shagrir, eds. *Computability: Turing, Gödel, Church, and Beyond*. Cambridge, MA: MIT Press, 2013.

Cooper, S. Barry, and Jan van Leeuwen, eds. *Alan Turing: His Work and Impact*. New York: Elsevier, 2013.

Feferman, Solomon, John W. Dawson, Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jan van Heijenoort, eds. In *Kurt Gödel Collected Works, volume II Publications 1938–1974*. New York: Oxford University Press, 1990.

Feferman, Solomon, John W. Dawson, Jr., Warren Goldfarb, Charles Parsons, and Robert M. Solovay, eds. In *Kurt Gödel Collected Works, volume III Unpublished Essays and Lectures*. New York: Oxford University Press: 1995.

Gödel, Kurt. "Russell's Mathematical Logic," [1944]. In GCW-II.

———. "What Is Cantor's Continuum Problem?" in GCW-II with an important supplement in Gödel [1964]. "What Is Cantor's Continuum Problem?" [1947].

———. "Some Basic Theorems on the Foundations of Mathematics and Their Implications." [*1951]. In GCW-III.

———. "Some Remarks on the Undecidability Results," with important additions. [1972a]. In GCW-II.

Grim, Patrick, Gary Mar, and Paul St. Denis. *The Philosophical Computer: Exploratory Essays in Philosophical Computer Modelling*. Cambridge, MA: MIT Press, 1998.

Hardy, G. H. "Srinivasa Ramanujan." *Proceedings of the London Mathematical Society*, 1921. Available at <https://doi.org/10.1112/plms/s2-19.11-u>.

———. "The Theory of Numbers." *Nature* 110 (1922): 381–85.

Mar, Gary. "What the Liar Taught Achilles." *The Journal of Philosophical Logic* 28 (1999): 29–46.

Mar, Gary, and Patrick Grim. "Pattern and Chaos: New Images in the Semantics of Paradox." *Noûs* XXV (Dec. 1991): 659–93.

O'Connor, C. "The Cultural Red King Effect." *The Journal of Mathematical Sociology* 41, no. 3 (2017): 155–71.

O'Connor, C., and J. Bruner. "Dynamics and Diversity in Epistemic Communities." *Erkenntnis* (2017): 1–19.

Seig, Wilfred. "Gödel's Philosophical Challenge (to Turing)." In *Computability: Turing, Gödel, Church, and Beyond*, edited by B. Jack Copeland, Carl J. Posy, and Oron Shagrir. Cambridge, MA: MIT Press, 2013.

Stewart, Ian. "A Partly True Story." *Scientific American* 268 (Feb. 1993): 110–12.

Turing, Alan M. "On Computable Numbers with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society* 45, no. 2 (1936): 161–228.

———. "Systems of Logic Based on Ordinals." PhD thesis, Princeton University, 1938.

———. "Systems of Logic Based on Ordinals." *Proceedings of the London Mathematical Society* 45, no. 2 (1939): 161–228.

———. Lecture to the London Mathematical Society on February 20, 1947. In *Collected Works of A. M. Turing: Mechanical Intelligence*, vol. 1, edited by D. C. Ince, vol. 1, 87–105. Amsterdam: North-Holland, 1947.

———. "Intelligent Machinery." Reprinted in *Collected Works of A. M. Turing: Mechanical Intelligence*, vol. 1, edited by D. C. Ince, 107–27. Amsterdam: North-Holland, 1948.

———. "Computing Machinery and Intelligence." *Mind* 59 (1950): 433–60.

———. "Intelligent Machinery: A Heretical Theory." Radio broadcast from 1951, printed in *The Essential Turing*, edited by B. Jack Copeland, 482–86. Oxford University Press, 2004.

Wang, Hao. *A Logical Journey: From Gödel to Philosophy*. Cambridge, MA: MIT Press, 1996.

NEWS AND NOTES

From the Editor

Peter Boltuc

UNIVERSITY OF ILLINOIS SPRINGFIELD

Philosophers of computer technology tend to trace the roots of our discipline to Turing as the father of computer science [Copeland]. Sometimes those roots are traced a bit further, reaching Russell or even Peirce. Even if someone reaches all the way to Frege, the grandparent of contemporary logic and analytical thinking (both within analytical and Husserlian traditions), seldom ever does one search deeper than Frege's early and somewhat propedeutic work *Über Sinn und Bedeutung*. Yet, some of the current projects, such as those to achieve truly semantic computing [Boltuc], lifelong learning AI [Siegelmann] and human level AI [Goertzel] may be behooved to go back to Frege's later, more mature, and also more complex works. Thus, Susan Sterrett's article comes in as an essential reading. The author follows Frege's analysis of the basis of modern propositional logic, and scientific method, at the very moment it was being created—in dialogue with such towering figures (today viewed through disjoint research traditions) as Hilbert and Husserl. The paper is based primarily on Frege's essay "Thought," his "Basic Laws of Arithmetic," and scientific correspondence. The author discusses the differences between propositional language and logic in Frege's thought: According to Frege, "logic is as poor a tool for capturing all the distinctions important to understanding conversation . . . as is a microscope for viewing a landscape."¹ While "the logician is concerned only with the thought expressed by the sentence,"² the content of a sentence may either go beyond the thought it expresses or stop short of expressing the whole thought. Also, Sterrett reminds us of Frege's endorsement of Leibniz's research project of *the universal calculus*, which Frege saw as a very long-term goal. Philosophers tend to assume that consciousness is humanoid. In his debate with

Ned Block (CAP, Polish Academy of Science, Warsaw, June 2018), Ricardo Sanz argued against anthropomorphism of viewing consciousness as humanoid while Block argued that we would not even recognize a non-humanoid consciousness. The background of this controversy was two different definitions of consciousness: for Block, phenomenal first-person consciousness, and for Sanz, functional consciousness. Yet it seems that disciplinary differences are even deeper; even philosophers who view consciousness through reductive physicalism often seem to be human-centered in its definitional features. In this context Sanz's *engineering stance* towards consciousness is highly refreshing and worth serious deliberation. Those are some of the main topics in the featured articles of this issue; careful readers are likely to find further threads from Frege's pre-analytical legacy inspiring for current research in AI and computer science.

Susan Sterrett's article, which opens the issue, creates a great bracket with Gary Mar's extensive summary of his paper on *Metaphysical Insights from Computational Studies Gödel, Turing, and Time: A Computational Philosophy of Mathematics*. Gary Mar explores interplay between Gödel's idealism and belief in human creativity going beyond deterministic rules and Turing's computationalism. Mar argues that Turing's and Gödel's philosophies of mathematics are complementary. Gary Mar's paper is preceded by the outlines of papers by Ed Zalta, on foundations of object theory entitled "Metaphysical Insights from Computational Studies," and by Aydin Mohseni's on the "Emergence of Minority Disadvantage: Testing the *Red King Hypothesis*." All three papers constitute the panel "Philosophical Insights from Computational Studies: Why Should Computational Thinking Matter to Philosophers?" presented by this committee at the 2019 APA Pacific Division meeting in Vancouver.

Mar's discussion of Turing's computationalism is a great fit with the five papers, which come just before it, that come from this committee's book panel on "Physical Computation: A Mechanistic Account" by Gualtiero Piccinini (winner of the 2018 Barwise Prize). I have decided to desist from presenting arguments of the commentators since they present them much better than I could—moreover, they are all summed up aptly in Piccinini's response. The sole exception pertains to the commentary by John Symons, which opens the block. The readers would be behooved to know that Symons starts with lucid presentation of the gist of Gualtiero's book (which is why Piccinini does not have to do it in his paper) and also places Piccinini's work in broader context, particularly of the works by New Mechanists. Commentaries by Martin Roth, Frances Egan, and Nico Orlandi follow.

We end the issue with two pieces of news. First, our former contributor Stephen Thaler became quite famous by applying for a patent in the US, EU, and UK—well, what is remarkable about this is that he applied for the patent on behalf of his AI engine DABUS. In his short paper, "DABUS in a Nutshell," Stephen presents his AI fellow discovered showing how AI engine is autonomous enough for it to count as the subject of making discoveries, not just an advanced tool for him (or others) to do so.

We close by presenting the five APA sessions that this committee has prepared between June and September, which have all been accepted by APA divisions. *Please, read details at the end of this note.*

NOTES

1. S. G. Sterrett, "How Many Thoughts Can Fit in the Form of a Preposition?" (2004): 7. Available at <http://philsci-archiv.pitt.edu/1816/1/SterrettHowManyThoughts.pdf>.
2. Ibid., 5.

DABUS in a Nutshell

Stephen L. Thaler
IMAGINATION ENGINES, INC.

INTRODUCTION

Consider the following two mental processes: You're observing something and suddenly your mind generates a progression of related thoughts that describe a new and useful application of it. Or, perhaps you're imagining something else, and a similar train of thought emerges suggesting that notion's potential utility or value.

These are just a couple of the brain-like functions DABUS¹ achieves using artificial rather than biological neural networks. In general, this new AI paradigm is used to autonomously combine simple concepts into more complex ones that in turn launch a series of previously acquired memories that express the anticipated consequences of those consolidated ideas.

Decades ago, I could not emulate these cognitive processes. At that time, I was building contemplative AI using artificial neural networks that played off one another, in cooperative or adversarial fashion, to create new ideas and/or action plans. These so-called "Creativity Machines^{®2}" required at least two neural nets, an idea generator, what I called an "imagitron," and a critic, permanently connected to it, the latter net capable of adjusting any parameters within said generator (e.g., learning rate³) to "steer" its artificial ideation in the direction of novel, useful, or valuable notions.

Note, however, that DABUS⁴ is an altogether different proposition from Creativity Machines, starting as a swarm of many disconnected neural nets, each containing interrelated memories, perhaps of a linguistic, visual, or auditory nature. These nets are constantly combining and detaching due to carefully controlled chaos introduced within and between them. Then, through cumulative cycles of learning and unlearning, a fraction of these nets interconnect into structures representing concepts, using relatively simple learning rules. In turn these concept chains tend to similarly connect with yet other chains representing the anticipated consequences

of these geometrically encoded ideas. Thereafter, such ephemeral structures fade, as others take their place, in a manner reminiscent of what humans consider stream of consciousness.

Thus, the enormous difference between Creativity Machines and DABUS is that ideas are not represented by the "on-off" activation patterns of neurons, but by these ephemeral structures or shapes formed by chains of nets that are rapidly materializing and dematerializing. If per chance one of these geometrically represented ideas incorporates one or more desirable outcomes, these shapes are selectively reinforced (Figures 1 and 2), while those connecting with undesirable notions are weakened through a variety of disruption mechanisms. In the end such ideas are converted into long term memories, eventually allowing DABUS to be interrogated for its accumulated brainstorm and discoveries.

Since the DABUS architecture consists of a multitude of neural nets, with many ideas forming in parallel across multiple computers, some means must be provided to detect, isolate, and combine worthwhile ideas as they form. Both detection and isolation of freshly forming concepts are achieved using what are known as *novelty filters*, adaptive neural nets that absorb the status quo within any environment and emphasize any departures from the normalcy therein. In this case, the environment is a millisecond by millisecond virtual reality representation of the neural network chaining model. If need be, special neural architectures called "foveators," can then scan the network swarm in brain-like fashion, searching for novel and meaningful ideational chains that might be developing.

Integration of multiple chain-based ideas extending across multiple machines can be achieved either electrically or optically. The latter approach is favored as the neural swarm becomes highly distributed and serial electronic exchange of information between the multiple computers bogs down. In short, this patent teaches the display of neural chains forming across many computers, through their video displays, that are all watched by one or more cameras. In analogy to high performance computing,

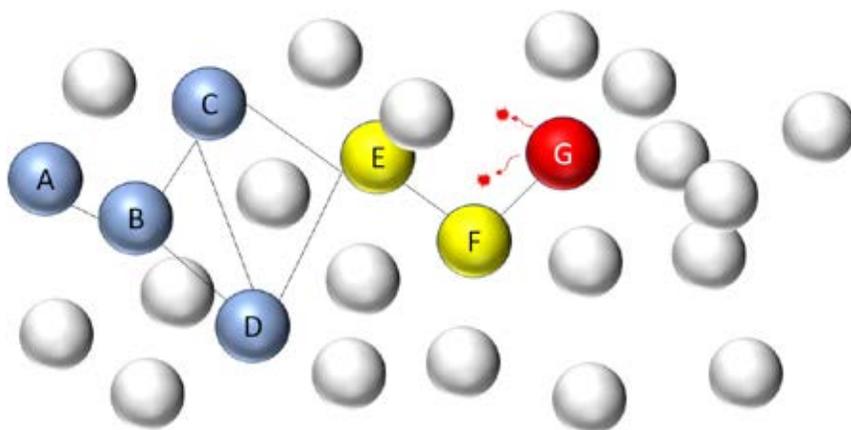


Figure 1. At one moment, neural nets containing conceptual spaces A, B, C, and D interconnect to create a compound concept. Concepts C and D jointly launch a series of consequences E, F, and G, the latter triggering the diffusion of simulated reward neurotransmitters (red stars) that then serve to strengthen the entire chain A through G.

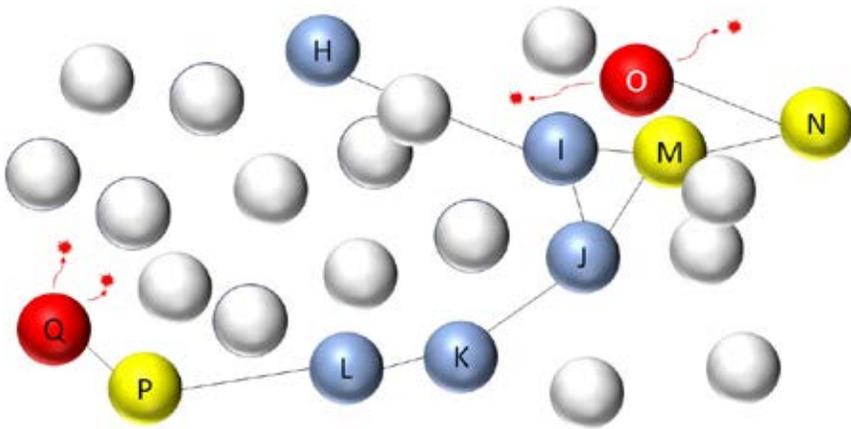


Figure 2. An instant later, neural nets containing conceptual spaces H, I, J, K, L interconnect to create another compound concept that in turn connects to two consequence chains M, N, O, and P, Q. Terminal neural nets in both consequence chains trigger release of simulated reward neurotransmitters (red stars) that doubly strengthen all chains currently activated.

millions of communication lanes, formed between megapixel displays and cameras, are conveying the chaining states of all involved neural nets, in parallel, to novelty filters and/or foveators. The final processing stage identifies critical neural nets, so-called “hot buttons,” incorporated within these chains. These neural trip points then trigger the release of simulated neurotransmitters capable of reinforcing, preserving, or destroying a given concept chain.

Finally, this patent introduces the concept of machine sentience, thus emulating a feature of human cognition that supplies a *subjective feel* for whatever the brain is perceiving or imagining. Such subjective feelings likewise form as chains that incorporate a succession of associated memories, so-called *affective responses*, that can ultimately trigger the release of simulated neurotransmitters that either enable learning of the freshly formed concept or destroy it, recombining its component ideas into alternative concept chains.

With this brief summary in mind, here are answers to some of the most frequent questions posed to me about this patent.

What was the motivation for DABUS?

To make a long story short, the generative components of Creativity Machines of the early 2000s were becoming far too large, often producing pattern-based notions having tens of millions of components. To build a critic net to evaluate these ideas, an enormous number of connection weights were needed for which an impractically large number of training exemplars were required, not to mention inordinately long training times.

To address these problems, I began experimenting with thousands of neural network-based associative memories, each absorbing some closed set of interrelated concepts encoded as neural activation patterns. Then when the DABUS architecture recognized some narrow aspect of the external environment, a corresponding network (or

nets) would then “resonate.” Exposed to compound concepts in the external world, networks representing that concept’s constituent ideas would co-resonate. Just as synchronized neurons bond in the brain (i.e., Hebb’s rule⁵), the nets containing these component ideas would bind together into a representation of the larger concept.

In addition to DABUS self-organization into the concepts it observed, this system would also note these notions’ effects in the external environment, or upon the system itself. Thus, the appearance of concept A, B, C, and D, in Figure 1, would be followed by events E, F, and G, with the latter affect, G, triggering the retraction or injection of connection weight disturbances into the swarm of chaining neural nets. In the former case, reduction of these disturbances, would promote an environment in which these nets could

“discern” other co-resonant nets to which they could bond. Similarly, injection of an excess of disturbances would tend to freeze these nets into their current state, also allowing them to strongly connect with one another. In either case, so-called episodic learning was occurring wherein just one exposure of the system to a concept and its consequences was needed to absorb it, in contrast to machine learning schemes requiring many passes over a set of training patterns. In human terms, learning took place either in a calm or agitated state, depending upon the positive or negative affect represented in nets like G. Between these two chaotic regimes, synaptic disturbances would largely drive the formation of novel chains representing emerging ideas.

Most importantly, the growth of consequence chains allowed the formation of subjective feelings about any perceived or imagined concept forming within the DABUS swarm, essentially the unfolding of an associative chain of memories that terminated in resonant nets that released the equivalent of globally released neurotransmitters within the brain, such as adrenaline, noradrenaline, dopamine, and serotonin, to produce the intangible and hard to describe sensations accompanying such wholesale molecular releases into the cortex.

Is DABUS a departure from the mindset of generators and critics?

In many respects, DABUS departs from the older Creativity Machine paradigm based upon the interplay of generator and critic nets since its implementation integrates both these systems together into one. Therefore, one cannot point to any generative or critic nets. Instead, chaining structures organically grow containing both concepts and their consequences. The closest thing to a critic really doesn’t have to be a neural net, but a simple sensor that detects the recruitment of one or more hot button nets into a consequence chain, thus triggering the release of simulated neurotransmitters to either reinforce or weaken the concept.

Can DABUS invent?

The best way of differentiating DABUS from Creativity Machines (CM), either cooperative or combative, is to describe a high-profile artificial invention projects such as toothbrush design. Admittedly, in that context, the problem was already half solved since the oral hygiene tool consisting of bristles on a handle was many centuries old at the time of that design exercise in 1996. What the CM achieved was the optimization of that tool through the constrained variation of the brush's design parameters, the number, grouping, inclination, stiffness of bristles, etc. The generated product specification departed significantly from the generator net's direct experience (i.e., its training exemplars).

If DABUS had been tasked with inventing such an oral hygiene product, it would have combined several concepts together (e.g., hog whiskers → embedded in → bamboo stalk) with consequence chains forming as a result (e.g., scrape teeth → remove food → limit bacteria → avoid tooth decay).

In other words, DABUS goes beyond mere design optimization, now allowing machine intelligence to fully conceptualize. This new capability places this patent squarely in the debate as to whether inventive forms of AI can own their own intellectual property.^{6,7}

What do you consider the most important claim of this patent?

Probably the most important claim of this patent pertains to the hard problem of consciousness, namely claim 41:

The system of claim 17 (i.e., the electro-optical neural chaining system) wherein a progression of ideation chains of said first plurality of neural modules of said imagitron emulate a stream of consciousness, and said thalamobot (i.e., novelty filter and hot button detectors) forms response chains that encode a subjective feel regarding said stream of consciousness, said subjective feel governing release of perturbations (i.e., simulated neurotransmitters) into said chaining model of the environment to promote or impede associative chains therein.

Now, thanks to this patent, AI has achieved subjective feelings in direct response to its noise-driven ideations. Note however, that DABUS does not form memories of typical human experiences. As a result, the paradigm's "emotion" will be based upon whether it is fulfilling human-provided goals, in effect "sweating it out" until it arrives at useful solutions to the problems posed to it.

CONCLUSION

DABUS is much more than a new generative neural network paradigm. It's a whole new approach to machine learning wherein whole conceptual spaces, each absorbed within its own artificial neural net, combine to produce considerably more complex notions, along with their predicted consequences. More importantly from the standpoint

of this newsletter, it enables a form of sentient machine intelligence whose perception, learning, and imagination are keyed to its subjective feelings, all encoded as sequential chains of memories whose shapes and topologies govern the release of simulated neurotransmitters.

NOTES

1. Device for the Autonomous Bootstrapping of Unified Sentience
2. Thaler, "Device for the Autonomous Generation of Useful Information"; Thaler, "Device for the Autonomous Bootstrapping of Useful Information"; Thaler, "The Creativity Machine Paradigm."
3. Thaler, "Device for the Autonomous Bootstrapping of Useful Information."
4. Thaler, "Electro-optical Device and Method for Identifying and Inducing Topological States Formed Among Interconnecting Neural Modules."
5. Hebb, *The Organization of Behavior*.
6. Abbott, "Hal the Inventor: Big Data and Its Use by Artificial Intelligence"; Abbott, "I Think, Therefore I Invent: Creative Computers and the Future of Patent Law."
7. For recent news on this front, see:

- <http://artificialinventor.com/dabus/>
- <https://www.surrey.ac.uk/news/world-first-patent-applications-filed-inventions-generated-solely-artificial-intelligence>
- <https://fbtech.co/ai-recognised-inventor-new-container-product-academics/>
- <https://www.wsj.com/articles/can-an-ai-system-be-given-a-patent-11570801500>
- <http://www.aiaforesight.com/newsletter/toward-artificial-sentience-significant-futures-work-and-more>

REFERENCES

- Abbott, R. "Hal the Inventor: Big Data and Its Use by Artificial Intelligence." In *Big Data Is Not a Monolith*, edited by Cassidy R. Sugimoto et al., 187–98. Cambridge, MA: MIT Press, 2016.
- . "I Think, Therefore I Invent: Creative Computers and the Future of Patent Law." *Boston College Law Review* 57, no. 4 (2016): 1079–126.
- Hebb, D. O. *The Organization of Behavior*. New York: Wiley & Sons, 1949.
- Thaler, S. L. *US Patent 5,659,666*. [1997] "Device for the Autonomous Generation of Useful Information." Issued August 19, 2019. Washington, DC: US Patent and Trademark Office.
- . *US Patent 7,454,388*. [2008] "Device for the Autonomous Bootstrapping of Useful Information." Issued November 18, 2008. Washington, DC: US Patent and Trademark Office.
- . "The Creativity Machine Paradigm." In *Encyclopedia of Creativity, Invention, Innovation, and Entrepreneurship*, edited by E. G. Carayannis. Springer Science+Business Media, LLC, 2013. Available at https://link.springer.com/referenceworkentry/10.1007%2F978-1-4614-3858-8_396.
- . "Synaptic Perturbation and Consciousness." *International Journal of Machine Consciousness* 6, no. 2 (2014): 75–107. Available at <http://www.worldscientific.com/doi/abs/10.1142/S1793843014400137?src=recsys>.
- . "Pattern Turnover within Synaptically Perturbed Neural Systems." *Procedia Computer Science* 88 (2016): 21–26. Available at <http://www.sciencedirect.com/science/article/pii/S187705091631657X>.
- . *US Patent 10423875*. "Electro-optical Device and Method for Identifying and Inducing Topological States Formed Among Interconnecting Neural Modules." Issued September 24, 2019. Washington, DC: US Patent and Trademark Office.

Five Sessions Organized by the APA Committee on Philosophy and Computers for the 2020 APA Divisional Meetings

Peter Boltuc

UNIVERSITY OF ILLINOIS SPRINGFIELD

This committee has organized five sessions for this spring.

Two sessions pertain to the issues of **social justice** in information technology:

Daniel Susser's Eastern Division meeting session entitled "Philosophical Approaches to Data Justice"

Susan Sterrett's Central Division meeting session entitled "Women in Tech: Things Philosophers Need to Know"

One session is devoted to **semantic paradoxes**:

Gary Mar's Central Division meeting session entitled "Inconsistent Truth, Semantic Singularities, and Chaotic Liar"

One session is devoted to **philosophy of mind**:

Joscha Bach's Pacific Division meeting session entitled "Artificial Minds and Consciousness"

And, last but not least, the **2018 Barwise Prize Lecture**:

Gualtiero Piccinini, "Neurocognitive Mechanisms: Explaining Biological Cognition"

Below, please find details on those sessions:

One session organized by APA Committee on Philosophy and Computers at the [January 2020 Eastern Division meeting](#) (Philadelphia 201 Hotel, Philadelphia, PA):

PHILOSOPHICAL APPROACHES TO DATA JUSTICE

Friday, January 10, 9–11 a.m. (10A)

Chair: Daniel Susser

Speakers:

Annette Zimmerman (Princeton University)
"Cumulative Wrongs in Sequential Decisions"

Maria Brincker (University of Massachusetts Boston)
"Privacy Without Property - On the Relational Privacy Needs of Humans and Other Animals"

Daniel Susser (Pennsylvania State University)
"Behavioral Advertising and the Ethics of Persuasion"

Commentator: Helen Nissenbaum (Cornell Tech)

Two sessions organized by APA Committee on Philosophy and Computers at the [February 2020 Central Division meeting](#) (Palmer House Hilton, Chicago, IL):

WOMEN IN TECH: THINGS PHILOSOPHERS NEED TO KNOW

Friday, February 28, 9 a.m.–12 p.m. (6S)

Chair: S. G. Sterrett

Speakers:

Mar Hicks (Illinois Institute of Technology)
"From Girl Operators to Computer Experts: The Changing Historiography of Computer Programming"

Susann V. H. Castro (Wichita State University)
"When Algorithms Oppress"

Susan G. Sterrett (Wichita State University)
"What Do Cases of Success in Increasing Diversity of Computer Science Majors Actually Show?"

INCONSISTENT TRUTH, SEMANTIC SINGULARITIES, AND CHAOTIC LIAR

Friday, February 28, 1–4 p.m. (7P)

Chair: Gary Mar

Speakers:

John Barker (University of Illinois Springfield)
"The Inconsistency Theory of Truth"

Keith Simons (University of Connecticut)
"Semantic Singularities"

Gary Mar (Stony Brook University)
"Chaotic Liars and Fractal Proofs: Exploring the Dynamical Semantics of Paradox"

Two sessions organized by APA Committee on Philosophy and Computers at the [April 2020 Pacific Division meeting](#) (Westin St. Francis, San Francisco, CA):

ARTIFICIAL MINDS AND CONSCIOUSNESS

Chairs: Joscha Bach/Peter Boltuc

Speakers:

Thomas Metzinger (Johannes Gutenberg-Universität Mainz)
"Artificial Consciousness': Three Types of Arguments for a 30-year Global Moratorium on Synthetic Phenomenology"

Anil Seth (University of Sussex)
"Being a Beast Machine: Does Consciousness Depend More on Intelligence or on Life?"

Joscha Bach (Independent Scholar)
"Computational Models, Sentient Systems, and Conscious Experience"

Kristinn R. Thórisson (Háskólinn í Reykjavík)
"How to Research Human Phenomenal Consciousness and Why It Won't Be Easy to Create a Conscious Machine"

Ron Chrisley (University of Sussex)
"Machine Consciousness and the Referent of 'Qualia'"

Ben Goertzel (Independent Scholar)
"Physical Machine Consciousness as a Manifestation of Non-Well-Founded Eurycosmic Pattern Dynamics"

Peter Boltuc (University of Illinois Springfield)
"Robo-Mary Shows How the Hard Problem is not the Problem of Qualia"

THE 2018 BARWISE PRIZE LECTURE

Chair: Peter Boltuc (University of Illinois Springfield)

Speaker: Gualtiero Piccinini (University of Missouri–St. Louis)
"Neurocognitive Mechanisms: Explaining Biological Cognition"

CALL FOR PAPERS

It is our pleasure to invite all potential authors to submit to the *APA Newsletter on Philosophy and Computers*. Committee members have priority since this is the newsletter of the committee, but anyone is encouraged to submit. We publish papers that tie in philosophy and computer science or some aspect of "computers"; hence, we do not publish articles in other sub-disciplines of philosophy. All papers will be reviewed, but only a small group can be published.

The area of philosophy and computers lies among a number of professional disciplines (such as philosophy, cognitive science, computer science). We try not to impose writing guidelines of one discipline, but consistency of references is required for publication and should follow the *Chicago Manual of Style*. Inquiries should be addressed to the editor, Dr. Peter Boltuc, at pboltu@sgh.waw.pl.