# PHILOSOPHICAL READINGS

ONLINE JOURNAL OF PHILOSOPHY

Editor: Marco Sgarbi

Volume XII – Issue 3 – 2020 ISSN 2036-4989

**ARTICLES**



**REVIEWS**



 DOI: 10.5281/zenodo.3903642

PHILOSOPHICAL READINGS
ONLINE JOURNAL OF PHILOSOPHY

*Philosophical Readings*, ISSN 2036-4989, features articles, discussions, translations, reviews, and bibliographical information on all philosophical disciplines. *Philosophical Readings* is devoted to the promotion of competent and definitive contributions to philosophical knowledge. Not associated with any school or group, not the organ of any association or institution, it is interested in persistent and resolute inquiries into root questions, regardless of the writer's affiliation. The journal welcomes also works that fall into various disciplines: religion, history, literature, law, political science, computer science, economics, and empirical sciences that deal with philosophical problems. *Philosophical Readings* uses a policy of blind review by at least two consultants to evaluate articles accepted for serious consideration. *Philosophical Readings* promotes special issues on particular topics of special relevance in the philosophical debates. *Philosophical Readings* occasionally has opportunities for Guest Editors for special issues of the journal. Anyone who has an idea for a special issue and would like that idea to be considered, should contact the editor.

Submissions should be made to the Editor. An abstract of not more than seventy words should accompany the submission. Since *Philosophical Readings* has adopted a policy of blind review, information identify the author should only appear on a separate page. Most reviews are invited. However, colleagues wishing to write a review should contact the Executive editor. Books to be reviewed, should be sent to the review editor.

# Wordsworth's Radical Aesthetics

*Mark Cladis*

**Abstract:** This article focuses on the aesthetics of William Wordsworth's work, particularly his early poetry. The implications of this investigation are far-reaching.[1] To learn about Wordsworth's aesthetics is to learn about Romanticism, specifically what I call *radical Romanticism* and the intricate relation it forges between aesthetics and democracy.[2] I begin the article with a general account of radical aesthetics, addressing its nature, scope, and its relation to the normative, the political, and the everyday. Next, I turn to the radical aesthetics of Wordsworth. I then compare radical aesthetics to more traditional accounts of aesthetics, and I conclude by connecting radical Romantic aesthetics to practical power.

**Keywords:** democratic theory; ethics; aesthetics; Wordsworth; romanticism.

## Radical Aesthetics

What makes *radical aesthetics* radical? It is radical in at least two, interrelated ways. It is radical insofar as it seeks to promote progressive, normative perspectives. And it is radical insofar as it is an aesthetics that is broad in scope (for example, its subject matter is not limited to the portraits of the museum or to the picturesque of the landscape). Its normative, progressive sensibilities include everyday justice and injustice and its aesthetic sensibilities include everyday experiences. So, the scope of radical aesthetics is as likely to pertain to the homeless population on the city streets as to the sunset over Mount Blanc. In the manner of Wordsworth (as we will soon see) radical aesthetics embraces the everyday, bringing careful, critical attention to what is in plain sight and what is hidden or concealed, and all the while offering an appropriate moral, political, and affective response to the social and natural quotidian world around us.

Radical aesthetics, I have claimed, is located in the everyday, and therefore it is also situated at the intersection of those spheres that are often named political and moral. In our everyday lives, these spheres usually intermingle in various ways. Radical aesthetics acknowledges this dynamic amalgam and does not seek to impose artificial disciplinary categories that distort our everyday experience. Art itself is often located in its own category separate from the political and the moral. Insofar as this is the case, radical aesthetics rejects these imposed boundaries. Additionally, radical aesthetics acknowledges and embraces the intricate, transactional relation between the social and natural world—between the cultural materiality of language, practices, and institutions and the dynamic processes of land, sea, and sky. In particular, the art of the word can be understood as both a spiritual project engaged with a material universe and as a material project engaged with a spiritual universe. But rather than employ such binary terms as spiritual-material, think of the art of the word—in the context of radical aesthetics—as a profoundly human project engaged normatively with the familiar and unfamiliar world in which we find ourselves.

Radical aesthetics seeks to imagine what it would be like to *taste*—to experience morally—the world in novel ways, especially in ways that are receptive to seeing and hearing and sensing beauty and generosity, pain and injustice in the otherwise familiar (perhaps too familiar) events, places, practices, traditions, and institutions that shape our lives. In order to focus on the everyday, radical aesthetics seeks to lift "the veil of familiar"—all those patterns of thought, sight, and practice that would render invisible and voiceless those creatures, human or nonhuman, that lack power or agency to make their needs and desires apparent. Radical aesthetics interrogates and illuminates how the world appears to us and how it could (and *should*) appear, and focuses on the disjunction between the two. This critical, creative work requires a tutored moral imagination and the cultivation of an ethically penetrating sight, or as I will soon call, a *democratic taste*.

The *scope* of radical aesthetics is as expansive as its *engagement* is deep and demanding. In this article, I explore how the art of Wordsworth seeks to cultivate an ethical, democratic taste that entails our senses, emotions, judgment, and intellect. It seeks to nurture within us a particular kind of bodily-cognitive response to the everyday that surrounds us. So, although Wordsworth is deploying a particular form of aesthetics—the art of the written word—the intent of his art is to broaden our aesthetic response to the world and not only to his art or to that of more traditional artistic forms. The appropriate aesthetic responses to our everyday experiences require much from us in the way of moral, emotional, cognitive, and bodily sensitivity and attunement. And the cultivation of such "appropriate aesthetic responses" is the work of radical aesthetics.

A robust aesthetic response, I have noted, entails cognition, bodily senses, emotion, and moral judgment or discernment. Radical aesthetics seeks to cultivate a fully *engaged* and *attuned* response to the natural and social worlds. Such engagement and attunement move people to delight in and decry the world in appropriate ways. There is nothing passive about this aesthetic response, though a high premium is placed on *receptivity*—or what Wordsworth called, "wise passivity." The radical aesthetic re-

Brown University
Providence (RI), U.S.A.
email: mark_cladis@brown.edu

Philosophical Readings XII.3 (2020), pp. 358-363.
DOI: 10.5281/zenodo.3903474

sponse is in part cognitive, because one needs to know something about the world to respond appropriately to the work of art or slice of life. For example, one needs to know something about poverty and farming to be appropriately *moved* –transformed—by Wordsworth's verse about unjust enclosures. But the aesthetic response entails more than the cognitive. The response is spiritual insofar as "spiritual" suggests that the whole person is called upon—that is, the integration of one's emotional life, bodily senses, moral judgment, and cognitive faculties.

For radical Romantics like Wordsworth, to be moved or transformed (however modestly) by a work of art is to experience the world differently and also to engage with the world differently. Unlike more traditional aesthetics, in which the viewer quietly gazes upon the artwork or landscape for the sake of disinterested enjoyment (akin to "the male gaze"), in radical aesthetics "the viewer" in fact becomes an active, normative *witness* to the world and in turn seeks to transform the world, working for political change and conditions congenial to a progressive democracy that honors the dignity of persons and seeks to abolish oppressive institutions and hierarchies. Radical aesthetics, then, far from separating the viewer from the everyday, seeks to fully engage its participants in the world around them, and that broad, aesthetic engagement includes a wide range of emotional, bodily, intellectual, and political responses.

Later in this article, I will contrast "traditional" to "radical" aesthetics. For now, however, it is important to underscore that radical aesthetics, unlike traditional aesthetics, embraces both "the moral" and "the practical" or "the useful." That is to say, radical aesthetics is unabashedly aligned with such normative projects as social justice. Those are practical projects. Moreover, radical aesthetics—unlike more conventional aesthetics—embraces "the useful." For example, rather than encourage the disinterested gaze on the picturesque landscape, radical aesthetics urges a useful, practical relation to the land: it endeavors to reveal the complex interactions among humans (including social and economic institutions), non-humans, and the ecosystems in which they cohabitate.

By explicitly acknowledging the moral and practical dimensions of aesthetics, radical aesthetics rejects most notions of "the disinterested gaze." The relation between aesthetics and "the disinterested gaze" has a long and complicated history. Conventional aesthetics has frequently held that the object of art, similar to that of science, should be approached with objective disinterest. This distanced, contemplative approach was deemed the ideal lest the viewer have an unduly subjective or utilitarian approach to the work of art. Kant, for example, argued that we must approach art free of personal motivation or self-interest, never utilizing art as a means to an end. How else are we to see the art for what it is as opposed to what we want it to be for our own sake and ends? And how else are we to produce impartial, universal aesthetic judgments, if not by eliminating personal preference, desire, and goals? While there are strengths to this line of argument (namely, placing a high value on our *receptiveness* to that which is outside us), conventional aesthetics has paid dearly in espousing the disinterested gaze. By separating art from *interest*—from involvement, concern, desire, commitment, and love—art became divorced from ethics and politics. Of course, many have justifiably questioned whether aesthetics could ever in fact be disinterested. Feminists, for example, have charged that the so-called ideal disinterested viewer is in fact the interested male voyeur, gazing on the female nude with a sanctioned yet hidden desire. In contrast to the ideal of the "disinterested" in conventional aesthetics, radical aesthetics promotes an aesthetics of active engagement, attunement, and moral critique.

## Wordsworthian Radical Aesthetics

Wordsworth believed in the social and political power of the well-crafted word. In a letter written in 1829, Wordsworth wrote, "Words are not a mere vehicle, but they are powers either to kill or to animate."[3] This is a powerful claim, one that he took seriously. The living power of a word can palpably contribute to life or to death. Wordsworth dedicated his poetic aesthetics to life.[4] He hoped that his radical aesthetics would contribute to the cultivation of *democratic, aesthetic taste*, employing the term *taste* in its literal meaning to refer to our capacity to experience the world normatively. When democratic taste has been suitably cultivated, one will see, feel and apprehend the drama of life in a distinctive way. One will, for example, be empathetic toward those who are crushed by social and economic oppression and work for change. The characters whom Wordsworth portrayed in his early years were often located at the periphery of society—the homeless, the impoverished, the disabled, the beggar, the wounded soldier. He portrayed them vividly so that we might see and feel their hopes and fears, their accomplishments and losses. We glimpse their humanity and in turn discover our own. This is a profound aesthetic and democratic achievement: to help citizens experience (to *taste*—to see, feel, hear, and be touched by) the dignity of fellow citizens, even those that many deemed to be lowly or dangerously "other." The aim of Wordsworth's radical aesthetics can be summed up as his various poetic efforts to make the audience see "souls that appear to have no depth at all/To vulgar eyes."[5]

An example of this effort to lend us sight to see the depth of "the other" is his 1802 sonnet, "The Banished Negroes."[6] In the summer of 1802, Napoleon had reintroduced slavery and instituted an ordinance that effectively expelled all people of color from France ("aucun noir, mulâtre, ou autres gens de couleur, de l'un et de l'autre sexe").[7] At the end of that summer, William and Dorothy Wordsworth traveled from Calais to Dover. A "Fellow-passenger" on their ship was a black woman who had been banished from France. As the headnote (added to the 1827 version of the sonnet) stated, "Among the capricious acts of tyranny that disgraced those times, was the chasing of all Negroes from France by decree of the government: we had a Fellow-passenger who was one of the expelled."[8]

The sonnet was important to Wordsworth. He returned to the poem, pondering and revising it, seven times throughout his career (in 1820, 1827, 1836, 1838, 1840, 1843, and 1845). It was one of his most significant at-

tempts to present the dignity and humanity of one who, for various reasons, is invisible to those with power, privilege, and membership in the majority culture. In this case, Wordsworth presented a black woman —a newly displaced refugee— to his fellow travelers. Much of the democratic and normative work of the sonnet is done within the first three lines: "*We* had a *fellow-Passenger* that came/From Calais with us…/A negro woman." Here two subjects are placed in a complex relation. On one side there is the *we* —we, William and Dorothy; we, the other passengers; we, the white majority culture of the British nation; we, the readers of the sonnet. On the other side is the solitary black woman. These two sides appear asymmetrical and perhaps unequal: one black woman up against multiple spheres of whiteness. Yet within the sonnet's first four words, such asymmetry and potential inequality are troubled by the mediating term, *fellow-Passenger*. Wordsworth places the woman among the *we*. She belongs to us, and we to her. Wordsworth's inclusion, however, has not erased difference. He acknowledges both the diversity that the woman embodies as well as her status as fellow traveler—as fellow human. This democratic move is accomplished by the employment of one of Wordsworth's more powerful aesthetic strategies: to reveal the unfamiliar in the familiar and the familiar in the unfamiliar. In this familiar event, the crossing from Calais to Dover, we are given sight to see the unfamiliar—this black displaced woman. At the same time, however, "careless eyes" are helped to perceive the familiar in this unfamiliar event—a fellow human being with depth of soul. In the face of her tangible presence, we experience both otherness and commonality.

After several attempts to convey the presence of this woman as both a familiar and unfamiliar fellow traveler —"like a Lady gay/Yet silent," "from notice turning not away" yet "motionless in eyes and face"—Wordsworth concludes the sonnet with explicit social criticism:

She was a Negro Woman driv'n from France
Rejected like all others of that race,
Not one of whom may now find footing there;
What is the meaning of this ordinance?
Dishonour'd Despots, tell us if you dare.[9]

When the poem was published in 1802, abolitionists in Britain were putting pressure on Parliament to end the Transatlantic Slave Trade. That same year, Napoleon had reintroduced slavery to Haiti, having captured and imprisoned Toussaint L'Ouverture, the former slave and the leader of the Haitian independence movement. With "The Banished Negroes" (and other sonnets written in 1802), Wordsworth allied himself with the abolitionists. His headnote and concluding lines express a clear political statement. I want to argue, however, that the political and democratic work is not done only in these direct statements. Indeed, most of the work is accomplished in the sonnet's earlier lines in which the black woman is presented. We look into her eyes and her humanity, and we witness our own. Through an act of the moral imagination, we see what was once invisible to us. Our eyes are no longer as careless as they once were. We have new sight. And to the extent that our recognition of humans has grown, so has our humanity.

Wordsworth's radical aesthetics, while often motivated by explicit democratic commitments and reasoned principles, engages with proper nouns that make claims on our lives—particular communities, people, places, and things. This is not a sign of anti-intellectualism. But as the 1790s progressed, Wordsworth became increasingly suspicious of the abstract—of things and ideas not rooted in the concreteness of time and space. This accounts for his eventual frustration with the abstract and impersonal nature of William Godwin's political philosophy. Godwin had argued that people would inevitably become illuminated by the ways of reason, and that the reformer's job was to help that enlightenment along, promoting impersonal truth and justice over private attachments and whim. Godwin offered the promised outcomes of the French Revolution without its violence. Godwin's alternative was timely and attractive for Wordsworth. Nonetheless, as the *Prelude* movingly documents, Wordsworth's faith in Godwin's philosophy was short-lived. Godwin's sanguine confidence in abstract reason, his strict impersonalism, and perhaps even his atheism, became obstacles to a Wordsworth who had put so much stock in the importance of concrete relations among people and place—in "a motion and a spirit, that…rolls through all things."[10]

An example of Wordsworth's expressed doubts about abstract "systems" of thought (as compared to concrete, aesthetic ones) is found in his 1798 "Essay on Morals":

> I know no book or system of moral philosophy written with sufficient power to melt into our affections, to incorporate itself with the blood and vital juices of our minds…Bald and naked reasonings are impotent over our habits; they cannot form them; from the same cause they are equally powerless in regulating our judgments concerning the value of men and things. They contain no picture of human life.[11]

This passage reveals what Wordsworth took to be the purpose and the high stakes of a radical aesthetics. It was to be practical. It was to move and change people. It was to reach deep into the hearts and minds of citizens and produce new habits, new political emotions, new ways of being. It was to present a robust "picture of human life."

Wordsworth's suspicion of the abstract suggests not only his distance from Godwin but also his proximity to Burke. In his 1789 "Letter to Charles-Jean-François Depont" Burke wrote, "You have *theories* enough concerning the rights of men… It is with man in the *concrete*, it is with common human life and human actions you are to be concerned."[12] Wordsworth's complaint about Godwin's abstract rationalism resembled Burke's critique of the French Revolution's faith in the Temple of Reason. Unlike Godwin, and like Burke, Wordsworth did not disdain the inevitable role of social traditions, habits, practices, and institutions. He did not spurn the idea of communities and individuals being rooted concretely in time and place. Not surprisingly, then, in Wordsworth's verse and in Burke's prose we find much reference to experience and history, to tradition and lived practice. In this regard, Wordsworth stood opposed to Godwin and alongside Burke. Wordsworth shared with Burke the conviction that places inexorably shape communities and their members. Unlike Burke, however, Wordsworth also held that the reverse is true: that *the people* shape a place. True, Burke

did believe that *some* of the people, namely the elite, have or should have the power to shape and guide their communities. But Burke would not ascribe such agency to the common people—to the "unthinking public."[13] In contrast, the early Wordsworth believed in and conveyed the agency of the people and their capacity to engage wisely with their environments.

In Wordsworth's poetry, certainly in the *Lyrical Ballads* and the *Prelude*, we see attempts to bring together Burke's emphasis on tradition and custom with Godwin's emphasis on reason and principles. Wordsworth was suspicious of both reason divorced from experience and experience detached from reason (critical, reflective thought). Like Burke, Wordsworth came to distrust abstract theories that hovered free of history or experience. Yet like Godwin, Wordsworth valued critical reflection and reasoned principles.

In the *Prelude*, for example, Wordsworth mocked the idea of reason severed from time and place:

> How Glorious!—in self-knowledge and self-rule
> To look through all the frailties of the world,
> And, with a resolute mastery shaking off
> The accidents of nature, time, and place,
> That make up the weak being of the past,
> Build social freedom on its only basis:
> The freedom of the individual mind,
> Which, to the blind restraint of general laws
> Superior, magisterially adopts
> One guide—the light of circumstances, flashed
> Upon an independent intellect.[14]

This passage, surely taking aim at Godwin's rationalism, lampoons the idea that public well-being and freedom can be achieved by means of a rationality detached from experience—from tradition, history, and local practices and conditions ("accidents of nature, time, and place"). Yet, wanting to be charitable toward Godwin and others who put their hope in "human reason's naked self," Wordsworth went on to claim that many have come, for good reason, to distrust tradition and practice because of the way these have been conceived by conservatives who refuse to reform customs and laws even when circumstances (such as unjust practices and policies) clearly dictate that change is needed. He noted that the French Revolution, in spite of its flaws, had nonetheless lifted "a veil" and "a shock had then been given/ To old opinions…" Yet many, in the name of "ancient institutions," refused to acknowledge what the Revolution had revealed: the sight of human suffering and need for change.[15] Here Wordsworth was no longer taking aim at Godwin but rather at Burke and all other traditionalists who had given "tradition" and "second nature" a bad name. The challenge for Wordsworth, ultimately, was to employ a Burkean language of traditions, habits, and virtues in service of a Godwinian democratic vision.

The early Wordsworth understood democracy not only as a set of formal political institutions but rather as a progressive culture or spiritual ethos that included the thought, skills, practices, dispositions, and emotions of diverse citizens. Wordsworth was committed to advancing an *embodied democracy* that emphasized the cultural dimensions of a democracy, including its religious and aesthetic ones. To achieve this goal, he sought to educate the *whole* person, rather than discursively addressing only the disembodied mind.

Wordsworth maintained that progressive political principles, such as those of Godwin, require more than abstract assent; they require the cultivation of humane taste, politically robust emotions, and a truly democratic second nature. Wordsworthian radical aesthetics, then, should be understood as a powerful way to *touch*—to move, engage, and transform—individuals for the sake of social progress. The Romantic poem, in this view, is an institution, potentially as powerful as the church, capable of shaping and training individuals and communities in the ways of justice. The well-crafted poem is the Fourth Estate: felicitous power outside official state, clerical, and economic forces. The well-crafted poem participates in what Ralph Waldo Emerson called "the true romance which the world exists to realize…the transformation of genius into practical power."[16]

## Radical Aesthetics in Light of Traditional Aesthetics

Radical aesthetics departs in notable ways from what is commonly associated with aesthetics. Radical aesthetics does not insist on the traditional, Kantian divide between aesthetics and ethics. While Kant did argue that the moral subject requires suitable training in order to appropriately grasp the sublime, thereby suggesting a relation between aesthetics and ethics, it is, nevertheless, a one-way relation. For Kant, aesthetics (the sublime) does not *shape* the moral subject but rather it *requires* a (certain kind of) moral subject. Radical aesthetics, in contrast, maintains that art can indeed morally cultivate the subject—even as it maintains that the subject's moral formation contributes to the apprehension of art. There is, then, a two-way relation—a dynamic dialectic—between aesthetics and ethics in radical aesthetics.

This close connection between art and ethics informs an account of pleasure that signals another departure from more traditional aesthetics. Art and beauty may, as Kant would have it, bring pleasure. But in radical aesthetics the object of such pleasure is not the unsullied form of beauty but rather an affecting, integrated depiction of a poignant slice of our social and natural world—for example, an apt, moving depiction of a beggar, a mournful mother, or a fallow garden. Such skillful artistic depictions wake us up, helping us to see more fully and to feel more keenly the reality of the social and natural world around us. In this process of waking up, of becoming more human and humane, we do experience pleasure. Such aesthetic pleasure, however, is a consequence of moral cultivation, not "a judgment of beauty." Radical aesthetics endeavors to empower its readers to become *witnesses*: to attest to the pain and injustice in our communities, institutions, and lands. Whereas Plato banned the poets from the republic, radical aesthetics would enlist the poets and encourage aesthetic projects and events for the sake of ameliorating the republic. And participation in that work of amelioration is a source of pleasure.

Although radical aesthetics participates in what Emerson calls "practical power," it may still be understood as

*disinterested*, capturing at least in part what Kant had in mind in the *Critique of Judgment* by the term *disinterest*. Radical aesthetics is disinterested insofar as it is, in an important sense, non-utilitarian. Art is not manipulated or produced for the sake of advancing narrow ends. The artist respects that which is before her by listening to it, by treating it with fidelity. Radical aesthetics does not impose. It does not distort, willfully. It seeks to witness gracefully and accurately—dare I say, objectively. Even that creative Romantic faculty, the *imagination,* is understood not as projection of fantasy but as a creative lens by which to bring clear-eyed attention to such experiences as war, famine, displacement, urbanization, over- and underemployment, water and air pollution, and oppressive political and religious authorities and institutions. The radical Romantic imagination exhibits a realism of the everyday as it engages in social criticism, bringing new (accurate, objective) sight to the social and natural worlds and the human approaches to them. Hence while reading Milton's *Paradise Lost*, Wordsworth scrawled in the book's margin: "The real excellence of Imagination consists in the capacity of exploring *the world really existing*."[17]

The "disinterestedness" of radical aesthetics, then, seeks to present the world faithfully, and the imagination is essential to this *creative, prophetic* task. It is a *creative* task insofar as perceptive discovery requires the imagination just as much as skillful creation. Creation requires receptivity even as receptivity requires creative attention (what Wordsworth calls "wise passiveness").[18] For this reason, radical aesthetics does not privilege the distinction between fiction and nonfiction, but rather calls our attention to the contrast between the authentic and the inauthentic, or between the honest and the dishonest. And presenting the world faithfully is a *prophetic* task insofar as it palpably brings into relief the normative gap between how the world is and how it ought to be. Once one becomes a witness and sees the world anew, one then *longs* for the world to become a different place—a place more just, less cruel. The term *longing* is important here, for it captures both the cognitive and affective aspects of radical aesthetics' prophetic task.[19]

The "disinterestedness" of radical aesthetics, then, does not entail moral indifference. Radical aesthetics is disinterested in narrow utilitarian aims, not critical, life-enhancing normative ones. Furthermore, it is not engrossed with that traditional aesthetic triad: the sublime, the beautiful, and the picturesque. This triad frequently removes art from the everyday and renders the spectator distant and detached—a far cry from the *engaged* witness. It has been customary to think of Wordsworth as the poet of the glorious landscape. And indeed, in his verse we find some of the most moving, detailed, and powerful descriptions of landscapes ever crafted in the English language. But these poetic descriptions neither convey a static landscape nor encourage a detached spectator. More to the point, however, Wordsworth's poetry was not principally centered on "the natural world" divorced from the presence of humans. Wordsworth found most of his inspiration in the life, labor, and struggle of commoners, and he presented, for all to see and feel, their worth and dignity. His poetry was above all about *people*—about people and the land, and about people and those public and private circumstances that comforted and confronted their lives.

Reading Wordsworth in this fashion, as a democratic poet, many not prima facie seem to cohere with what he is perhaps most famous for, namely the "spots of time" in the *Prelude*. Wordsworth himself used the expression, "spots of time," only once to refer to events in his childhood: "There are in our existence spots of time, / That with distinct pre-eminence retain/A renovating virtue, whence…our minds/Are nourished and invisibly repaired."[20] Commentators, however, commonly employ the term to refer to a number of powerful, often revelatory incidents in Wordsworth's past. These incidents —vividly described memories —typically entail a solitary or isolated Wordsworth (even if he is in the company of others) encountering an evocative landscape (e.g., Snowden). It is not clear, however, why "spots of time" cannot apply to transformative incidents that entail people and that had profound political implications for Wordsworth. Nicholas Roe, in fact, comes close to making this move. He interprets as "*almost*…a spot of time" a scene in the *Prelude* that depicts a powerful political event in Wordsworth's life in Revolutionary France.[21] When Wordsworth and Beaupuy —his friend and radical political mentor— encountered on the road "a hunger-bitten girl":

…and at the sight my friend
In agitation said, 'Tis against *that*
Which we are fighting', I with him believed
Devoutly that a spirit was abroad
Which would not be withstood, that poverty,
At least like this, would in a little time
Be found no more…[22]

In this spot of time, Wordsworth gained a sudden clarity on the goals of the Revolution and his commitment to it. The impoverished girl became a palpable symbol of the Revolution and all that it stood for, and the symbol charged Wordsworth's life with a profound sense of purpose and meaning. Other such transformative moments that entail people and community could be plausibly cited as spots of time with compelling, explicit sociopolitical import (scenes, for example, from "The Female Vagrant," "The Old Cumberland Beggar," or "The Ruined Cottage"). All this is to say that radical aesthetics is not preoccupied with the distant, sublime, uninhabited landscape.

Radical aesthetics, in contrast to traditional aesthetics, is more closely aligned with the way John Dewey placed art at the center of an everyday realism.[23] Similarly, when Wordsworth offered his revolutionary description of *who* the poet is and *what* poetry is for, he highlighted the poet as a fellow human addressing fellow citizens and employing their *everyday language*. And although that language was, of course, suitably transformed into poetic form (though not "ornately"), Wordsworth maintained that there is no fundamental difference between prose and verse (still another gesture toward the everyday). And not only was the language of Wordsworth's poetry "everyday," but as we've seen, so were the characters, events, and places about which he wrote. In this regard, he revolutionized the scope or range of poetic topics. An urgent concern for the everyday was placed at the center of art.

My point is that radical aesthetics is often a quotidian aesthetics: it emerges from the everyday and it largely remains there. A premium is placed on presenting the ordinary and commonplace in such a way as to move and touch citizens for the sake of promoting social justice and environmental practice. Ultimately, then, the aesthetics of radical Romanticism seeks to participate in "the true romance of the world," namely, the transformation of genius—the transformation of minds and hearts, of cognition and affect—into practical power.

## Notes

[1] By "the early Wordsworth," I refer to his work in the 1790s and up to approximately 1805. There is a well-entrenched account of the trajectory of Wordsworth's poetry and political persuasion. When he was young, his poetry was original and vibrant and his political beliefs were democratic and progressive. As he aged, his poetry became staid and his politics conservative. There is much debate about exactly when his apostasy occurred, though most agree it took place sometime between 1798 and 1806. Some, like Jerome McGann, push the date so far back that it is unclear whether Wordsworth ever had progressive political beliefs (Jerome J. McGann, *The Romantic Ideology* [Chicago: Chicago University Press, 1983]). Others, such as William White, doubt if there ever was a change of heart (William Hale White**,** *An Examination of the Charge of Apostasy Against Wordsworth* [London: Longmans, Green, and Co: 1898]). If one is, like me, mostly interested in his early poetry and its political, religious, and environmental dimensions, then the relevance of when or whether Wordsworth committed apostasy is not entirely clear. I will say, however, that eventually Wordsworth did become more conservative in his political and religious beliefs and practices. This transition is a long, complicated path.

[2] By Romanticism, I refer retroactively to the late eighteenth- and early nineteenth-century artistic, literary, philosophical, religious, and political trend; I also refer to a current, implicit, and pervasive cultural framework that has its roots in the Romantic era. The manners and interpretations of Romanticism, of course, are legion. I have interrogated a selection of Romantic materials and constructed my own portrait of Romanticism for the sake of advancing progressive democratic aims. I call this portrait *radical Romanticism*. I acknowledge that my selection of material is itself a constructive exercise, yet I also believe that I maintain fidelity to my sources. *Radical Romanticism*, then, refers to my critical, constructive engagement with Romantic sources.

[3] *The Letters of William and Dorothy Wordsworth: the Later Years* (Oxford: Clarendon, 1979), Vol. 5, p. 185.

[4] Although my main focus in this section is on Wordsworth and his aesthetic medium, poetry, this should not suggest that I privilege poetry over other artistic forms. My arguments about Wordsworth's poetry and radical aesthetics can be extended to apply to various forms of art, for example, animation, music videos, calligraphy, hip-hop, plays, tapestry, graffiti, and performative aesthetics.

[5] William Wordsworth, *The Prelude* (1805), ed. Jonathan Wordsworth, M.H. Abrams, and Stephen Gill, (New York: Norton, 1979), 12: 167-68.

[6] "The Banished Negroes," in *Poems, in Two Volumes, and Other Poems, 1800-1807 by William Wordsworth*, ed. Jared Curtis (Ithaca: Cornell University Press, 1983), pp. 161-62. For helpful discussions of the sonnet, see Debbie Lee, *Slavery and the Romantic Imagination* (Philadelphia: University of Pennsylvania Press, 2002), pp. 203-07; Judith Page, *Wordsworth and the Cultivation of Women* (Berkeley: University of California Press, 1994), pp. 67-76; and Jonathan Bate, *Song of the Earth* (Cambridge: Harvard University Press, 2002), p. 215.

[7] "Any black, mulatto, or other persons of color, of either sex," *Collection complète des lois*, ed. J. B. Duvergier (Paris: Guyot et Scribe, 1836) Volume 13, p. 242.

[8] Wordsworth, *The Collected Poems of William Wordsworth* (Hertfordshire: The Wordsworth Poetry Library, 1994) p. 305.

[9] "The Banished Negroes," in *Poems, in Two Volumes, and Other Poems, 1800-1807 by William Wordsworth*, ed. Jared Curtis (Ithaca: Cornell University Press, 1983), p. 62.

[10] Wordsworth, "Lines Written A Few Miles Above Tintern Abbey," *William Wordsworth: the Major Works*, ed. Stephen Gill (Oxford: Oxford University Press, 2008), p. 134.

[11] *William Wordsworth: Selected Prose*, ed. John O. Hayden (New York: Penguin, 1988), p. 105.

[12] Edmund Burke, "Letter to Charles-Jean-François Depont," *On Empire, Liberty, and Reform: Speeches and Letters*, ed. David Bromwich (New Haven: Yale University Press, 2000), p. 410. Emphasis added.

[13] Burke, "Letter to Charles-Jean-François Depont," p. 414.

[14] William Wordsworth, *The Prelude* (1805), ed. Jonathan Wordsworth, M.H. Abrams, and Stephen Gill, (New York: Norton, 1979), 10: 819-29.

[15] Wordsworth, *The Prelude* (1805), see 10: 849-63.

[16] Emerson, "Experience," *Ralph Waldo Emerson: Essays and Poems*, eds. Harold Bloom, Paul Kane, and Joel Porte (New York: Library of America, 1996), p. 492.

[17] Cited in *The Cambridge Companion to Wordsworth*, ed. Stephen Gill (Cambridge: Cambridge University Press, 2003), p. xiv. Italics added.

[18] Wordsworth, "Expostulation and Reply," *William Wordsworth: the Major Works*, p. 130.

[19] On the dialectical relation between affect and cognition and its social-critical power, see Sharon Krause. *Civil Passions: Moral Sentiment and Democratic Deliberation* (Kindle Locations 927-942); and Melvin Roger's manuscript, *The Darkened Light of Faith: Race, Democracy and Freedom in African American Political Thought*, Chapter Two, "The Demandingness Of Freedom: David Walker And Racial Domination," p. 23.

[20] William Wordsworth, *The Prelude* (1805), 12: 208-216.

[21] Nicholas Roe, *Wordsworth and Coleridge: The Radical Years* (Oxford: Clarendon Press, 1990) p. 59. Emphasis added.

[22] William Wordsworth, *The Prelude* (1805), 9: 518-24.

[23] John Dewey, *Art as Experience* (New York: Penguin, 2005), pp. 1-19.

# Is Personal Identity, Moral Identity?

*Nath Rajakishore*[+], *Sahu Vineet*[‡]

**Abstract:** The problem of personal identity is a classical problem in philosophy. This question has been variedly tackled by different philosophers and philosophical schools. To address the problem of personal identity, it is essential to explicate the notion of 'person.' Many philosophers conceive persons as inherently conscious beings, who are capable of intentional mental activities which are explicable from the ability to have the first-person perspective and imagine the same of the other. On the other hand, physicalistic personhood is something that has mechanical/bodily properties but, either lacking consciousness or reducing it to a physical basis. For many others, persons have both the properties of mind and body, not reducible to each other. We would agree partially with the latter position and maintain that persons have not only physical properties but also various forms of consciousness, i.e., self-consciousness, moral consciousness, etc. The ability to take perspectives, we claim, lays the foundation of moral consciousness. In this paper, we aim to show that the idea of personal identity is very much related to moral consciousness. This is because persons are rational beings, and being rational is natural to the person. If a person does any irrational act, it almost becomes self-denial to him or her as rationality is natural to her or him. Therefore, rationality is one of the inborn qualities of human being, and thus her or his identity becomes a moral identity. If we accept persons as physical beings ultimately, the question of morality, freedom, and responsibility do not arise. The whole idea of self-determination is occurred only in the case of moral identity, but not in the case of persons as only physical beings. Therefore, personal identity and moral identity are conceptually connected to the extent that we propose that personal identity is moral identity.

**Keywords**: Personal identity; Moral identity; Material identity; Self-conscious identity; Moral identity.

## The Problem of Personal Identity

Before analyzing the concept of person, we need to raise a few questions like what is a person, what is the nature of the person, and so on. These are fundamental questions in the philosophy of mind. The English word 'person' is alleged to have been derived from the Latin '*persona*,' which was the mask worn by actors in dramatic performances.[1] Neither in common usage nor in philosophy has there been a univocal concept of 'person.' In common usage 'person' refers to any human being in a general way. The person is distinct from a thing or material object. It stands for a living, conscious human being.

Let us take two philosophers - P. F. Strawson and Bernard Williams as a paradigm of contrasting conceptualization of the philosophical notion of person: Strawson defines 'person' as "a type of entity such that *both* predicates ascribing states of consciousness *and* predicates ascribing corporeal characteristics, a physical situation &c. are equally applicable to a single individual of that single type."[2] For Strawson, persons are unique individuals who have both mental and physical attributes. Thus persons are neither purely physical body, nor are they pure spiritual substances. However, the Strawsonian view of persons is also non-material, whereas Williams' view of the person is purely material, which is in contrast with the Strawsonian view. This is because of William's[3] claim is that bodily continuity is a necessary condition for personal identity, because according to Williams, it is the body which identifies the persons, but not the mind, and there is no mind at all; therefore, bodily criterion identifies the persons. Therefore, the category 'persons' belongs to spatiotemporal beings. Similarly, according to Parfit, a person is a psychological continuous living human body.[4] Because physical continuity is neither necessary nor sufficient for personal identity. It is insufficient because a human being could conceivably change so radically that s(he) would not be the same person. A great deal of what matters to us in ourselves and others is psychological: our memories, characters, tastes, interests, loves, and so forth.[5] This spatiotemporal being is the ontological individuality of a person. Thus persons are ontologically a natural kind. Now we can say that they are natural beings at one level and are self-conscious and minded being at another. The self-conscious includes the self-description of his or her awareness that s(he) is a solid, continuing being in the world and this spatiotemporal description itself recognized persons are natural minded being.[6]

Now the question is, does Strawson wish to say that persons are bodies of a particular sort, namely, bodies which have mental attributes as only? Strawson holds that persons have bodily attributes too. But unlike ordinary bodies, persons are things, which have mental attributes as well. According to Strawson,[7] it is essential to persons that they be entities, which necessarily have both mental and bodily attributes. Also, those mental things are substantially different from physical things? They are different types of substance. Persons are radically different material bodies. Strawson's theory appears to be dualistic - in holding that there are two different types of subjects, the physical bodies, and persons.

+ Indian Institute of Technology, Bombay, India.
‡ Indian Institute of Technology Kanpur, India.
emails: rajakishorenath@iitb.ac.in ; Vineet@iitk.ac.in

Philosophical Readings XII.2 (2020), pp. 364-369.
DOI: 10.5281/zenodo.3903490

Again, physical bodies necessarily have only one dimension, that is, physical dimension. Then, there is perhaps no error in holding that (as concurred by many artificial intelligence theories) persons are mechanical beings,[8] but this seems to be experientially counter-intuitive for persons necessarily have two dimensions, a physical and a mental dimension. Persons thus have a dual nature. The most important fact about the person is the self. The self is sometimes used to mean the whole series of a person's inner mental states and sometimes the spiritual substance to which they belong. The self does not refer to the body but the mental history of the person. This made the unity problem seem intractable, when the mental images, feelings, and the like, are contrasted with the temporal persistence. In Strawson's sense, a person is a thing which necessarily has both mental and physical aspects. The person is primarily the subject of mental experience. In the Strawsonian person theory, we cannot say that a person is a body, but we can say that a person is, in part, a body. If the person is a body, then it cannot be a conscious mind, and therefore, there is nothing wrong attributing mind to machines. One of the important questions is can we even say that a person has a body? But what would it mean for the theory of person? It means that persons have bodily attributes. Another question is, does it say anything about the relation between a person and a body? The body necessarily has bodily attributes and has nothing to do with a person's attributes. But Strawson's view is that persons have both bodily and mental attributes. According to Strawson,[9] properties like 'being at such and such time and place,' having such and such weight and colour, and so on are M-predicates. The other properties are psychological properties like 'being in the state of happiness,' or 'being in the state of pain,' and so on are states of P-predicates. In this way, Strawson has rightly said, "the concept of a person is to be understood as the concept of a type of entity such that both predicates ascribing corporeal characteristics, a physical situation and consciousness are equally applicable to an individual entity of that type."[10] What is significant about them, as Strawson has pointed out is their co-applicability to the same person substance. The M-predicates cannot be ascribable independently because of that prohibits them from being ascribable to the conscious beings, like M-predicates, the P-predicates cannot be ascribed to the material bodies. This is because of a combination of a distinct kind of substance that has both physical and mental properties without being reducible to each other.

The above argument shows that Strawson was accepting the person as non-material and non-dual yet incorporating the explanatory powers of Cartesian dualism. This is because Descartes held that when we dwell on the concept of person, we are referring to one or both of two distinct substances of different types, each of which has its appropriate types of states and properties. Each of which has its own appropriate types of states and properties and none of the states belong to both. That is to say that states of consciousness belong to one of these substances and the other. Descartes has given a sharp focus to this dualistic conception of the person. It is not easy to get away from dualism because persons have both sorts of attributes such as mental and physical. According to dualistic conception, a person is something altogether distinct from the body. That is, a person is not identical to his body. Some dualists, however, believe that a person is a composite entity, one part of which is its body and another, part of which is something-immaterial spirit or soul. Thus dualism essentially adheres to the mind-body distinction and persons as mental beings as distinguished from material bodies.

Joseph Margolis[11] in his book '*Persons and Minds*' mentioned that persons are the particulars that have minds and nervous systems, sensation, and brain processes. But this will not quite do. A nervous system is not a person, nor is a psyche a person. It is at once the subject of both neurological and psychological predicates. In other words, it is both a nervous system and a psychic entity. Persons are not mereologically complex entities nor kind, each of which contains parts, a non-physical basic subject and a purely corporeal object to which this subject is in some way attached; for such a claim would not allow us to ascribe psychological attributes or corporeal attributes to the person as a whole. It is because persons are more than their bodies and that they are not reducible to any kind of body gross or subtle. The person-substance as described above is not taken to exclude the material properties as such. They only exclude the fact that persons are material bodies and nothing else. Persons are autonomous so far as their description in terms of bodies and mind is concerned. But it is not that no reference to body and mind is to be retained at all. Thus, person's description has the attributes reference to body and mind. Therefore, one of the paradoxical implications of the person theory is that the body which a person has cannot be conceived of as a physical object subject to the law of the physical world as we know from this theory, that persons are conscious. Finally, from the above examination, we may infer that a person's body is not a physical thing.

**Person as Self-Conscious and Ontological Being**

As we have argued so far, the person is an entity which has both mental and physical attributes. Hence, we could say of a person that s(he) is five feet tall, weighs a hundred kilograms, etc. But more importantly we could say that s(he) is thinking about his friends, feels a pang of guilt or is happy, or so on. We may, therefore, say that a person has a mind, which is different from his body because the subject of consciousness does not mean a body of a certain sort. But it still might turn out that whatever is a subject of consciousness is identical with a body of a certain sort.

Generally, consciousness is described as something, which distinguishes man from a good deal of the world around. Only a person possesses this consciousness, which is not by other material objects. Again, the question arises, what is this consciousness which a person certainly has, but rocks and other animate beings do not have? As G.E. Moore writes, "The moment we try to fix our attention upon consciousness and to see what distinctly it is, it seems to vanish: it seems as if we had before us as mere emptiness when we try to introspect the sensation of blue, all we can see is the blue; the other element is as if it were

diaphanous."[12] Of course, we know perfectly well that we are conscious of things around us including other people, but we do not grasp consciousness itself.

However, it is this common feature, consciousness, which may be said to be the central element in the concept of mind. A person as being minded[13] is to have the capacity of doing mental activities. Such activities include thinking, willing, feeling, understanding, speaking, communicating, and above all, remembering the past. Mental activities are such that they presuppose that there is a thinker who is capable of these activities. The thinker is here a subject or 'I' who is or has the capacity of consciousness. Wherever we will find the concept of 'I,' we will find the existence of consciousness because it is a person who stands for the concept of 'I,' have consciousness.

One of the most general views is that the philosophy of mind is concerned with all mental phenomena which they are concerned with consciousness. Philosophers from Descartes onwards have accepted consciousness as a fundamental metaphysical reality. I remain the same person if I am conscious of being so, even though my body should change drastically and be diminished through amputation. Conceptually, it is possible that I should remain the same person although I am altogether disembodied. Persons are indivisible, non-corporeal simple entities. It is because it becomes difficult here to distinguish persons so construed from metaphysical selves, transcendental egos, spirits, mental substances, souls, and other similar immaterial substances. However, the concept of a person does not fit into these entities, because persons are, if anything, concrete beings in the world. One can ascribe consciousness to others only if one can identify other subjects of experience. Also, one cannot identify other subjects if one can identify them only as subjects of experience, possessors of state of consciousness.[14] The latter must have a concrete existence in the world.

If we are too obsessed with the 'inner' criteria, we shall be tempted to treat persons essentially as minds. We do need to factor in the widely and empirically substantiated claims (from psychology, neurosciences, medical sciences) that the bodily constituency also affects the mental characters and the domains of mind and body are not disjointed. However, admitting outer criteria does not mean that there are no states of consciousness. We should claim that some P-predicates refer to the occurrence of a state of consciousness. Persons are certainly identifiable beings having a life of their own. They are not Cartesian egos; rather they possess a mixed bag of M-predicates and P-predicates. Persons are in any case conscious individuals who can be ascribed a large number of P-predicates such as thinking, feeling, willing, deciding, etc. these conscious states, according to Searle,[15] are intentional, that is, are of something. That is, they are directed at something outside them. Thus, persons who have these conscious states are intentional and mental beings.

Again, only a being that could have conscious, intentional states could have intentionality at all, and so every unconscious intentional state is at least potentially conscious. This thesis has an enormous consequence for the study of the mind. But there is a conceptual connection between consciousness and intentionality that has the consequence that a complete theory of intentionality requires an account of consciousness. And our consciousness is consciousness of something. Thus persons have the essential feature of consciousness. There is an interconnection between person, mind, and consciousness. Empirically, there are distinctions between them. But transcendentally, they point in the same direction. It is right to say that a person is a mental being, and the essence of mind is consciousness. Therefore, the concept of the mind, the person, and the consciousness go together. Thus consciousness is related to mind, which also belongs to a person.

P.F. Strawson has adopted the term 'person' for a philosophical use which comes rather closer to common usage than did Locke's usage of the form, while it raises philosophical problems of its own perhaps it is less disreputable to hold that the person is a primitive concept. This is because the Lockean view of the concept of person is forensic concept, but Strawsonian concept of person is a metaphysical concept like the concept of the self and therefore is not merely a social or a forensic concept.[16] Pradhan[17] points out that it is metaphysical, precisely because it shows how it can be used to describe the minded being as the unique substance which is not identical with the body, though it is necessarily linked with the body. That is to say that persons have material bodies and yet are not on the same levels as physical bodies or organisms. Persons, therefore, are not physical things at all and this is because persons transcend their physical existence. The person-substance as a minded being tended to be the 'I' or the self in the sense that, though it is a continuant being in the spatiotemporal world, yet it does not belong to the world in the way the human belongs. It is because human beings belong to this world, but the 'I' or the self is not an entity in the world at all. The fact is that the self is not a physical body to be counted in the world. Hence there is a sense of transcendence underlying the concept of self, though it does not mean that selves or persons for that matter are mysterious entities.

The transcendental qualities, however, shows that persons are explainable from the first-person perspective. The first-person perspective grants a unique individuality or an 'I' who experiences, as Wittgenstein[18] points out that even it is not 'name' which can substitute 'I.' Therefore, the first-person is not the descriptions of any human being, because it refers to the third-person perspective, but it refers to the person himself or herself. This does not mean that person is distinct from this world, but a person is part of this world. As a Strawsonian person, to begin with, is to be understood as distinct from a mere material body, which retains the contrast customarily observed between a person and things.

Let us now think of some ways in which we ordinarily talk of ourselves, certain things which we do, and which are ordinarily ascribed to ourselves. We ascribe to ourselves actions, intentions, sensations, thoughts, feelings, perceptions and memories. However, we ascribe to ourselves location and attitude. Of course, not only we ascribe ourselves temporary conditions, states, situations, but also enduring characteristics, including physical characteristics like height, shape, and weight. That is to say, those among the things that we ascribe to ourselves are

those that we also ascribe to material bodies. But there are things and attributes that we ascribe to ourselves, but cannot dream of ascribing to material bodies.

Moreover, the persons are human beings and have the history of an organism belonging to a natural kind. At the same time, they are not physical beings at all, and they transcend their physical existence, i.e., self-conscious moral being. This transcendental nature of a person cannot be merely only a social construction. Thus persons are metaphysical beings who could be all these and yet must be claimed in the sense that they could not be what they are without a metaphysical essence.

**Personal Identity as Moral Identity**

In this section, we aim to show that personal identity as moral identity. Many scholars like Kant, who claim that the fundamental idea of philosophy is human autonomy. According to Kant, 'autonomy' literally means giving the law to oneself.[19] Kant's moral philosophy is also based on the idea of autonomy. He holds that there is a single fundamental principle of morality, on which all specific moral duties are based. He calls this moral law the categorical imperative. The moral law is a product of reason, for Kant, while the basic laws of nature are products of our understanding. Moral rightness and wrongness apply only to free agents who control their actions and have it in their power, at the time of their actions, either to act rightly or not.[20] Thus freedom depends on our moral identity and humans have complete liberty to practice morality. Because a person claims that 'I should do justice' or 'All men are equal' or 'I am an honest person' or 'I am a good human[21] Thus it is a morality that leads to freedom, and it is a problem of human to claim that freedom is necessary being a moral agent. The freedom is not something given; rather it is something achieved. This is because a person is potentially free, but certain obstacles that s(he) has ignorantly put around herself or himself appear to limit her or his freedom. But Rajendra Prasad pointed out that for some philosophical systems, freedom is only instrumental good, that is, it is good because it is helpful to, is a necessary condition for, to achieve something higher; for some other systems, it is intrinsically good, that is, a state desirable for its own sake. There are also many systems which consider both instrumentally and intrinsically good. Lastly, many believe that freedom is a state of mind and not of the body.[22]

Let us look at the idea of moral identity. If persons are not moral, they are not persons. This is because persons are rational beings, and being rational is natural to the person. If a person does any irrational act, it becomes self-denial to her or him as rationality is natural. The moral identity modified into self-identity.[23] Therefore, rationality is one of the in born qualities of human being and thus his or her identity became a moral identity. If we accept persons are material beings, the question of morality and freedom do not arise. There is an idea of self determination is exist in the case of moral identity, but not in the case of persons are material identities. Therefore, personal identity as moral identity is because being a person is self-conscious and ontological being.

However, the claim is that to be moral is natural for a person as it is for her/him to be rational. Humans are *naturally moral beings* in the sense that every human person has the natural propensity to act morally unless otherwise constrained.[24] In this way, a person's moral nature follows jointly from her/his being rational as well as spiritual. It is spiritual because person urge rises above self-interest and ego-centricity and are motivated to act morally from the universal point of view. The view that rational discernment of the moral form the immortal and spiritual urge for self-excelling put us naturally on the path of moral progress. Hence, the moral inclination is innate to persons, who are constitutively endowed with the capacity of rationality and spirituality. This is because moral personhood is constitutive in nature.

The person as the ideal moral agent makes choices which are rationally guided but are not the result of a deliberative of a process undertaken by an abstract moral reason. These choices flow from the internal moral commitments of the person. The moral commitments are the commitments of the normative kind which are located in the moral nature of the person concerned.[25] We would like to reiterate that it is futile to ask for the justification for persons being rational. This shows not that rationality lacks a necessary justification but that it needs no justification because it cannot intelligibly be questioned unless it is already presupposed. It is not the case that men are rational because of the circumstances, but a person's rational identities are beyond any kind of conditions. The natural fact that persons are capable, by rationality, not only of theoretic discriminations of sorts but also of judging what we should do and should not do. This is because a person is a self-knowing agent in the sense that s(he) is aware of his or her moral commitments as a person having willingly made rational choices. Following Pradhan, we would like to point out that self-knowledge in this sense is not merely knowing what one does, but willing to undertake to do certain actions. The self-knowing person is a willing and performing person. Self-willing is as important as self-knowing. Therefore, self-knowledge is a part of the moral personhood.

Self-knowledge is very much related to the idea of the person being rational is evaluative consciousness, which is characteristic of personal consciousness. This is because of the moral reflection on a person's actions. Moral reflection, by which we engage in the moral evaluation of ourselves, our actions and those of others, is thus constitutive of personhood; for moral reflection is the activity of rational agents and only persons are rational agents. The relation of 'constitution' is crucial to explain the notion of person and personal identity. The Constitution is a widely occurring relation in the world around us. What needs to be clarified first is that constitution is neither a relation of identity nor a relation of separateness.

However, the constitutive nature of personhood is very much related to the reflective nature of moral action. The objects of one's moral action may be a reflection of one's own intention, actions as well as those of others and such reflections are usually expressed in the form of moral judgments.

However, the moral judgments lead to a person committing to those beliefs and desires in the sense that these in-

tentional states play a constitutive role in the unfolding of plans and projects of the person. The person is so-called because of the plans and projects which s(he) intends to execute. For this kind of execution, the person needs to have rational choices. This is one of the important aspects of being moral personhood. Again, this fact is very much related to the notion of 'freedom' and freedom is the basis of rational choice. If the person had not a rational nature, s(he) could not have exercised freedom because freedom is the very essence of rationality and morality.[26] On the other hand, the idea of freedom is sometimes formulated independently values, preferences, and reasons. Therefore, freedom cannot be fully appraised without some idea of what a person prefers and has reason to prefer. Here, there is a basic use of rational assessment in appraising freedom. In this sense, freedom must depend on a reasoned assessment of having different options. As Amartya Sen says, "Rationality as the use of reasoned scrutiny cannot but be central to the idea and assessment of freedom."[27] If we take freedom otherwise, rationality depends on freedom. To practice a rational choice in civil society, freedom is necessary. But if we see freedom as a person's natural right, freedom is necessary for her or his existence. If one views rationality or morality from a Kantian perspective, it is a person's primary duty to be moral as being only a rational animal in the universe. Our aim in this paper is not to discuss the philosophical problems exist about person's external world, rather analyze person's natural nature, as we have been arguing is an inherent feature of personal identity, which is morality or rationality. We do not deny that freedom can be valued for the substantive opportunity it gives to the pursuit of our objectives and goals.

Moreover, while responding to one of the important questions like "why should I be rational?" The first answer would be: it is up to the person. Here, the idea of freedom inherently present in a person's existence. As Rajendra Prasad said, that person should be moral because being moral has an intrinsic dignity of its own or that it is a necessary condition for having personal happiness, social harmony, etc.[28] Our answer would be rational is part of the definition of being a person. There may be several non-ethical points of view from which we may demand reasons for action, for example, self-interest. From impartial spectator, it may be the reason for transcending self-interest and act on universal judgments. This transcendental nature of a person leads to spirituality. This is because moral judgment is possible through spiritual endeavors. This leads to the fact that there is an innate potentiality which when actualized or realized, leads the person to rise above his self-interest. This spirituality has an innate capacity, which is the realization through intellectual maturity. This intellectual maturity is under the spirituality that the person is a self-reflective and self-evaluator. Therefore, there is a spiritual urge for self-transcendence underlies a person's life of self-appraisal that leads to self-transcendence. The self-transcendence requires that the person's motivational system be controlled by his valuation system with which he identifies his or her reason and has to control his or her passion.

In the above passage, the concept of freedom plays a vital role. It is because a person's commitments and choices are based on a person's freedom. Freedom as a social value is the power to choose between real alternatives and to realize the chosen alternative to further one's own, someone else's or some group's interests. By calling it power, we want to emphasize the fact that it is not a state or occurrence but ability. It is not a state of mind, nor something that occurs to it, but a power or ability to work in a certain manner. This power may or may not be exercised on a certain occasion, but this does not mean that on that occasion it has ceased to exist. But whatever it is exercised, in a conscious manner. It is a conscious power because it is a power to choose, and choice, by its very definition, involves conscious selection and rejection. It is contradictory to claim 'I chose x in the context c out of x, y, z, but was not aware of my choice[29] This is why many philosophers argue that it is a natural propensity of a person to act morally and in the same manner to become a moral person one depends on his or her freedom. Freedom of the individual, which is the driving force of self-excellence, is the prerogative of persons. It lies at the root of the urge to moral perfectibility and goes to the making of what we have been referring to as unfailing moral personhood. It is because persons always seek freedom and s(he) is engaged in reflective self-evaluation, and such is an embodiment of value. Freedom is important because, on Kant's view, moral appraisal presupposes that we are free in the sense that we can do otherwise. Freedom is not supposed as a character or quality; rather it is the very essence of the person. Any qualities or characters are something different from that to which it belongs, but freedom does not belong to the person, the person is freedom. Although there are various determining factors that exist in the universe, this does not mean freedom is not there in the universe. Therefore, freedom does not mean 'no-determination,' it means self determination.

We would like to point out that person's ability to be free, the ability to identify herself/ himself with the value judgment and in this way the person transcends the level of mere human existence, but the person never identifies with this material world. This does not mean that a person does not interact with this material world. Rather, the person attains rationality through the human body (material world).[30] A person alone attains perfection, but not any other beings in this world. This is how moral progress is possible. This progress is possible by moral striving, and moral striving that aims at moral perfection is not just rational but spiritual. Thus a person can exercise the spiritual self- transcendence, as well as the freedom that leads to perfections. Therefore, perfection is a matter of natural necessity as we have seen that morality is a rational affair and since rationality is constitutive of personhood, morality becomes intrinsic to personhood. That is why we do not find any difficulty to combine both spirituality and rationality to establish personal identity.

## Bibliography

Arthur C Danto, Encyclopedia *of Philosophy*, Vol. 6, Paul Edwards (Ed.), Macmillan and Free Press, New Delhi, 1967.

Bhat, P. R., Human Autonomy and Ramchandra Gandhi, in *Ramchandra Gandhi: The Man and His Philosophy*, A. Raghuramaraju (ed.), Routledge, New Delhi, 2013, Pp. 85-117.

Hardy, Sam A. et al., The Integration of Personal Identity, Religious Identity, and Moral Identity in Emerging, Identity: An International Journal of Theory and Research, 2017. http://dx.doi.org/ 10.1080/ 15283488.2017. 1305905.
Kant, Immanuel. *Grounding for the Metaphysics of Morals*. Translated by Ellington, James W. (3rd Eds.). Hackett, USA, 1993 [1785]
Korsgaard, Christine, Personal identity and the unity of agency: A Kantian response to Parfit. Philosophy and Public Affairs Volume:18, No. 2, 1989, pp.101-132.
Lock, John, *An Essay Concerning Human Understanding*, Clarendon Press, London, 1894.
Margolis, Joseph, *Persons and Minds*, D. Reidel Publishing Company, Holland, U.S.A., 1978.
Mohapatra, P. K., *Ethics and Society: An Essay in Applied Ethics*, Concept Publishing Company, New Delhi, 2008.
Moore, G. E., Philosophical Studies, London: Routledge & Kegan Paul Ltd., 1970.
Nath, Rajakishore, Machine intelligence, competence and creativity, *Artificial Intelligence and Society: Journal of Knowledge, Culture and Communication*, Volume 23, Number 3, May 2009, pp. 441-458. doi:10.1007/s00146-007-0080-y.
Parfitt, D., Reasons, and Persons, Oxford University Press, Oxford, 1986.
Pradhan, R. C. *Self-Knowledge, Moral Agency and Normativity*, a paper presented at International Conference on Self-Knowledge and Moral Identity at Department of Humanities and Social Sciences, IIT Bombay, India, January 13-16, 2015.
Pradhan, R. C., Persons as Minded Beings: Towards a Metaphysics of Persons, Journal of Indian Council of Philosophical Research, Vol. XV, No.3, 1998, pp. 17-25.
Prasad, Rajendra, *Karma Causation and Retributive Morality: Conceptual Essays in Ethics and Metaethics*, ICPR, New Delhi, 2004.
Searle, R. John, *The Rediscovery of Mind*, The MIT Press, London, England, 1994.
Sen, A., *Rationality and Freedom*, Oxford University Press, New Delhi, 2002.
Strawson, P. F. *Individuals*, University Paper Backs, London, 1963.
Vohra, Ashok, *Man, Morals, and Self*: *A Philosophical Perspective*, Viva Books, New Delhi, 2014.
Wilkerson, T. E., *Minds, Brains and People*, Clarendon Press, Oxford, 1974.
Williams, B., Problems of the Self, Cambridge University Press, Cambridge, 1973.
Wittgenstein, L., *Philosophical Investigations*, translated by G. E. M. Anscombe, Blackwell, Oxford, 1953.

## Notes

[1] Arthur C Danto, *Encyclopedia of Philosophy*, Vol. 6, Paul Edwards (Ed.), Macmillan and Free Press, New Delhi, 1967, p. 110.
[2] Strawson, P. F. *Individuals*, University Paper Backs, London, 1963, pp. 101-102.
[3] Williams, B., Problems of the Self, Cambridge University Press, Cambridge, 1973, pp. 1-81
[4] Parfit, D., Reasons, and Persons, Oxford University Press, Oxford, 1986, pp. 205.
[5] Hardy, Sam A. et al., The Integration of Personal Identity, Religious Identity, and Moral Identity in Emerging, Identity: An International Journal of Theory and Research, 2017. http://dx.doi.org/ 10.1080/ 15283488. 2017.1305905
[6] Pradhan, R. C., Persons as Minded Beings: Towards a Metaphysics of Persons, *Journal of Indian Council of Philosophical Research*, Vol. XV, No.3, 1998, p.20.
[7] Strawson, P. F. *Individuals*, University Paper Backs, London, 1963, pp. 81.
[8] Nath, Rajakishore, Machine intelligence, competence and creativity, *Artificial Intelligence and Society: Journal of Knowledge, Culture and Communication*, Volume 23, Number 3, May 2009, pp. 441-458. doi:10.1007/s00146-007-0080-y.
[9] Strawson, P. F. *Individuals*, University Paper Backs, London, 1963, pp. 100-101.
[10] *Ibid*.
[11] Margolis, Joseph, *Persons and Minds*, D. Reidel Publishing Company, Holland, U.S.A.1978, p.12.
[12] Moore, G. E., Philosophical Studies, London: Routledge & Kegan Paul Ltd., 1970, p. 25.
[13] Pradhan, R. C., Persons as Minded Beings: Towards a Metaphysics of Persons, *Journal of Indian Council of Philosophical Research*, Vol. XV, No.3, 1998, p.20.
[14] Wilkerson, T. E., *Minds, Brains and People*, Clarendon Press, Oxford, 1974, p.15
[15] See, Searle, R. John, *The Rediscovery of Mind*, The MIT Press, London, England, 1994, pp. 130-132.
[16] Lock, John, *An Essay Concerning Human Understanding*, Clarendon Press, London, 1894, p.230.
[17] Pradhan, R. C., Persons as Minded Beings: Towards a Metaphysics of Persons, *Journal of Indian Council of Philosophical Research*, Vol. XV, No.3, 1998, p.19.
[18] Wittgenstein, L., *Philosophical Investigations*, translated by G. E. M. Anscombe, Blackwell, Oxford, 1953, sec. 410.
[19] See, Kant, Immanuel. *Grounding for the Metaphysics of Morals*. Translated by Ellington, James W. (3rd Eds.). Hackett, USA, 1993 [1785]
[20] Korsgaard, Christine. 1989. Personal identity and the unity of agency: A Kantian response to Parfit. Philosophy and Public Affairs Volume:18, No. 2, 1989, pp.101-132.
[21] Hardy, Sam A. et al., The Integration of Personal Identity, Religious Identity, and Moral Identity in Emerging, *Identity: An International Journal of Theory and Research*, 2017. http://dx.doi.org/ 10.1080/ 15283488.2017. 1305905
[22] Prasad, Rajendra, *Karma Causation and Retributive Morality: Conceptual Essays in Ethics and Metaethics*, ICPR, New Delhi, 2004, P. 399.
[23] Bhat, P. R., Human Autonomy and Ramchandra Gandhi, in *Ramchandra Gandhi: The Man and His Philosophy*, A. Raghuramaraju (Ed.), Routledge, New Delhi, 2013, pp. 85-117
[24] Mohapatra, P. K., *Ethics and Society: An Essay in Applied Ethics*, Concept Publishing Company, New Delhi, 2008, p. 44.
[25] Pradhan, R. C. *Self-Knowledge, Moral Agency and Normativity*, a paper presented at International Conference on Self-Knowledge and Moral Identity at Department of Humanities and Social Sciences, IIT Bombay, India, January 13-16, 2015.
[26] Pradhan, R. C. *Self-Knowledge, Moral Agency and Normativity*, a paper presented at International Conference on Self-Knowledge and Moral Identity at the Department of Humanities and Social Sciences, IIT Bombay, India, January 13-16, 2015
[27] Sen, A., *Rationality and Freedom*, Oxford University Press, New Delhi, 2002, p.5.
[28] Prasad, Rajendra, *Karma Causation and Retributive Morality: Conceptual Essays in Ethics and Metaethics*, ICPR, New Delhi, 2004, P. 399.
[29] Prasad, Rajendra, *Karma Causation and Retributive Morality: Conceptual Essays in Ethics and Metaethics*, ICPR, New Delhi, 2004, p. 400.
[30] Vohra, Ashok, *Man, Morals, and Self*: *A Philosophical Perspective*, Viva Books, New Delhi, 2014, pp.59-62.

# The Concept of Freedom: Effects of Ockham's Revolution

*Kateryna Rassudina*

**Abstract:** In his famous book *Sources of Christian Ethics*, fr. Servais Theodore Pinckaers OP described the process of transformation that our understanding of the ethically right and wrong has undergone for the last two thousand years. Much of his attention the scientist gave to the problem of human freedom, on the decision of which largely depend the moral principles of a society. In particular, he clarified and analyzed the transition from the concept of freedom advocated by Thomas Aquinas to the William Ockham's concept of freedom. In this article, I will not only present this transformation, but also point out those of its consequences in European thinking, which we still observe in our societies.

**Keywords:** freedom, virtue, indifference, autonomy.

## 1. Freedom in the understanding of Thomas Aquinas

The problem of free will became one of the important topics in medieval Christian philosophy, primarily due to the comprehension of the interaction between the absolute will of God and the moral choice of man. In this article, I will mainly refer to the solution proposed by Thomas Aquinas, but before him about freedom not only Augustine Aurelius wrote, but also such apologists as Justin Martyr and Clement of Alexandria, such Church Fathers as Basil the Great and Gregory of Nyssa. In turn, medieval solution, early as well as high, is based on the ancient philosophy, for example, Aristotle and the Stoics, and, of course, on Holy Scripture.

Freedom, according to Aquinas, is an ability proceeding from the intellect and the will ("*liberum arbitrium dicitur esse facultas voluntatis et rationis*[1], which, in turn, are the inclinations to truth and good. The first dimension of free choice (the order of entity and specification) depends on the intellect: on the object of knowledge, as well as on the analysis of alternatives. Initially, the direction of judgment is determined by some intuitive comprehension, the ability to choose between good and evil in favor of the first (principle: *bonum est faciendum, malum est vitandum*), namely by synderesis. The second dimension of free choice comes from the subject (the choice between action and inaction). Judgments of the intellect, practical judgments, give to the choice a conscious motivation, however this choice is actualized by will, by effort of the person who strives for the good by means of the action, and thus it becomes free. It is worth noting that the original movement of the will, according to Thomas, like the intuitive knowledge of the first principles, is a spontaneous volition of such goals as good and happiness. The difficult task of freedom consists in choosing the means to achieve these goals, i.e. it is represented as a function of the will.[2]

As we see, according to Aquinas, will and intellect follow natural inclinations to good, truth, happiness, that's why he considers these inclinations as sources of freedom, and not as restrictions for its manifestation. Spontaneity, from which the freedom proceeds, is inherent in all people by nature; it ensures the achievement of happiness as a natural goal. In the realm of morality, this spontaneity is manifested in the sense of truth, good, right and love; i.e. they also do not limit, but reveal the freedom: "We are free, not in spite of them, but because of them. The more we develop them, the more we grow in freedom".[3]

Therefore, it is possible to evaluate the so-understood freedom only on the basis of the attractiveness of the object: it spontaneously manifests itself with respect to the things marked by good or truth. The subject can be mistaken and even deliberately strive for a false good – this is the weakness of freedom or rather of the being endowed with freedom. That's why a person should cleans his or her inclinations of external layers. The naturalness of striving for happiness means that counteraction to this desire cannot be called free.

On the other hand, the fact that we cannot choose the ultimate goal, but only the means to achieve it, also does not mean a limitation of freedom. According to the logic of Thomas Aquinas, the very our ability to unlimited truth and good is unlimited. Looking forward them (or loving God), we do not experience pressure, because this gravity, and consequently, the will motivated by it is natural for us. Thanks to our spiritual abilities, our unlimited freedom, our will becomes free also in relation to all final and individual goods, means on the way to the goal.

The virtues help us to advance in the way to the final goal, to adapt to the true good by developing of our perfect natural abilities. Virtues, according to Aquinas, are the perfections of strength, of dynamics in a person ("*virtus, ex ipsa ratione nominis, importat perfectionem potentiae*"[4]). We can say that virtue is a property of character, or a permanent predisposition to actions that are agreed with the good. Virtue is directly connected with the act of making a decision, with a choice: we manifest virtues when perpetrate acts of will, but also our acts of will, our choices are conditioned by character traits. This is the case with all habitus – predispositions that govern desires and behavior in general (virtues are the most perfect of them), abilities to act and develop his- or herself (with regard to Tomas, habitus should not be confused with habits as psychological mechanisms). At first glance, such

St. Thomas Aquinas Institute of Religious Sciences
Kyiv, Ukraine.
email: rassudina.k@gmail.com

Philosophical Readings XII.3 (2020), pp. 370-375.
DOI: 10.5281/zenodo.3903494

a system of habitus and virtues should limit freedom, prompting certain decisions to be made. However, in the views of Aquinas, habitus and, especially, virtues, in the case of their development, on the contrary, allow freedom to unfold, as they promote progress towards the ultimate goal, i.e. toward happiness.

As I have already noted, virtues can be and must be developed, and the path of such improvement coincides with the way of strengthening the will, the intellect and, as a result, the educating of freedom. Surprisingly, this education begins from discipline – not as a violent direction of the will, but as an appeal to natural inclinations, spontaneous sense of the good, to a conscience. The task is not to inculcate behavioral habits that restrict freedom, but, on the contrary, to discover the opportunity that it has: the desire of unlimited truth and unlimited good.

At the second stage of that formation, a disciplined person takes his or her life – first of all, moral one – in own hands and, so to say, develops the taste for values. Instead of the limitations that characterize discipline, the person is motivated by the desire for progress of his or her own virtue for its own sake.

At the third stage of the freedom's formation Thomas sees not only moral maturity, but also fruitfulness, as well as the responsibility of a person:

> Due to the gradual development of his faculties, the human person is now capable of viewing his life in its entirety. He performs his actions personally according to a plan, a higher goal which will profit himself and others. This leads him, through patient acceptance of all trials and obstacles, to the fulfillment of a lift project which gives meaning, value and fruitfulness to existence.[5]

As we will see later, in this definition of Pinckaers there is something from Sartre's attitude, however, he expresses the thought of Thomas Aquinas. The similarity is natural, since it is a question of one subject – freedom. On the other hand, the point of view of Aquinas cannot be estimated worthily without specifying that the formation of freedom should include not only development and harmonization of one's personality, but also an openness to another person. Agape love is the perfection of such an openness.

Thus, from the point of view of Thomas Aquinas, freedom is a possibility of development coordinated with human rationality. However, the medieval theologian does not exclude the ability to resist the prompts of intellect, also calling it freedom. Contrary to teleology, there is also the possibility not to strive for happiness. So, Aquinas overcomes the threat of determinism on the part of rationality; however, he believes such a decision is not the maximum manifestation of the freedom, but, on the contrary, a sign of its weakness (something like childish whim).[6]

In conclusion, I should note that, Aquinas adopts the Aristotelian definition, according to which a man is free, because he is the goal for himself, and does not exist for something else. However, unlike Aristotle, who considered some people to be slaves by nature, Thomas Aquinas asserted the radical and universal equality of all people. This remark allows us to assert that "freedom by nature" is considered by Thomas as "one of the refinements of existence in a particularly excellent way: existence not as a means to an end, but as an end in itself"[7], which inherent in a person.

## 2. William Ockham's revolution

British Franciscan William Ockham was almost a contemporary of Thomas Aquinas. However, his attitude to freedom and morality in whole is fundamentally different. He asserts that freedom is a power to cause, indifferently and accidently, different effects, in such a way that a person can bring about an effect or don't do it. Freedom is postulated by Ockham as the ability to choose between opposites; he reduces it to indifference in relation to those values that could determine this choice. The indifference to the objects of choice allows the will to proceed when it makes a decision solely from itself.

Such conclusions Ockham reached analyzing the nature of God. For Ockham, the will of God is so free that it opposes determination even by His goodness. Good, according to Occam, is what God desires, but not contrary. The same can be said about the moral order in the created world: he divine will at any time can change its own precepts, even the first commandment. Thus, the original freedom inherent in God is indifferent to the objects of choice. Freedom becomes the main property of God; He becomes the freedom himself.

Since there cannot be two definitions of freedom in one system, Ockham gives a person the same independence from his own nature as God. However, God also is omnipotent, He can impose His will on man: this is the only difference between divine and human freedom. Accordingly, the relationship between God and man can rest only on the "right of the strong" (first of all, of God); the strong enforces the low, and the weak has the obligation to fulfill it. Taking into account the fact that, according to Ockham, God is free to change any of His precepts, we see here the way to moral relativism: morality rests on obligation versus law regardless of its content.

Freedom as the "right of the strong" to impose his law on the basis of non-determinism by external objects and internal characteristics also means that it turns out to be the primary force in relation to reason and will. It is a matter of free choice – to learn or not to learn, to desire or not to desire. In relation to the reason, this means that the will can resist to its decisions. Thus, in spite of the fact that freedom is declared to be the primary fact of experience, practically it is identified with the will, it becomes the source and conductor of human actions. The will receives a new interpretation.

> It was no longer defined as an attraction toward the good, exercised in love and desire, as in St. Thomas and the Fathers. It became a radical indifference, whence proceeded a pure will, actually ab imposition of will on itself or others, 'a conscious pressure of self upon self', to use E. Mounier's definition. This was to become the modern understanding of will.[8]

But most acutely Ockham opposed freedom to the natural inclinations of human – to good, truth and happiness. Similarly, sensuality and, in particular, passions become

threats, obstacles to freedom, not only: the will strengthens itself in opposing them (the rigorism in morality). As a motivation, alien to the will, Ockham considers also a virtue.

Since the inclinations are in subjection to freedom, the desire for happiness ceases to be the determining factor for the human. A person is free to choose happiness, unhappiness as well as his or her nonexistence. Hence the negation of the finality, i.e. the existence of the ultimate goal, which is happiness.[9]

The negation of the finality is directly related to Ockham's nominalism. If, according to Thomas Aquinas, ultimate goal is a principle of the unity of all particular person's actions, then Ockham's absolute indeterminacy of the individual actions breaks the foundation of such unity. An independent action takes place after an instant unrelated decision. Therefore, Ockham considers the goal only as a single, isolated one of a concrete action.

A person whose dignity is the ability to do at any moment what he or she wants becomes elusive – result of collecting a "puzzle" from various, inconsistent acts.[10] This disappearance of the person behind a variety of actions did not signify, however, its destruction, but the concentration on him- or herself in the name of the most important willing – the assertion of the freedom. Indifference of the will in relation to inclinations and prompts of reason does not mean apathy, but, on the contrary, implies a kind of passion: "the human will to self-affirmation, to the assertion of a radical difference between itself and all also that existed"[11]. This requirement to provide a person, rests solely on his or her selfness, the ability to choose between opposites, we now, after Kant, call autonomy. We understand it as independence from the law, norms and everything that does not come from the self, as an opportunity to refuse, to contradict, to accuse and generally to be against something, and also to enjoy the pleasure in arbitrariness of our actions.

### 3. Development of the concepts of freedom and autonomy after Ockham's "revolution"

At the time when the philosophy of the Middle Ages was transformed into a Modern one, Ockham's understanding of freedom was firmly entrenched in European thinking. Descartes, for instance, argues that the will by nature is so free that it can never be constrained, limited. Contemporary of Descartes, Thomas Hobbes, being a materialist and emphasizing the egoistic character of human nature, pays much attention to the motives of person's activity. He already doesn't understand why Thomas Aquinas has placed inclinations and reason in the foundation of freedom. Therefore, although natural law, according to Hobbes, is a freedom (liberty) to do anything against anyone, referring to own interest (to own goal), this freedom is limited by the natural law in the form of reason's direction. At the same time, the inborn desire, the will to claim the benefit for oneself, is in such a disunity with this natural law that the latter can only be guaranteed by the dictates of the state. It is the basing of freedom as an unlimited choice between alternatives that prevents Hobbes from asserting the existence of real freedom – permissiveness – in civil society: according to its nature, such a society deprives the individual to carry freely certain actions in favor of preserving the harmony of interests.[12]

Jean-Jacques Rousseau also thinks in the same way: man is born free, i.e. unlimited in his choice of actions. Naturally, liberty is "bounded only by the strength of the individual"[13]. Liberation, therefore, consists in overcoming the restrictions (imposed by other people) and gaining the benefit of one's own strength without meeting resistance. However, this freedom from limitations is combined in Rousseau's theory with positive freedom: it is into it that the social contract transforms the permissiveness. When concluding the contract, people put all their strength under control of the common will; now they make decisions not individually, but together. Such civil liberty implies the right to preserve property and enjoy the rights of a citizen (first of all, the right to take part in collective decision-making), as well as moral freedom.

Pointing to this transformation of the liberty, Rousseau, however, realizes that in fact it is its limitation and even destruction. A natural human being who cares only about natural liberty does not have a sufficient basis to be tempted by the delights of civil liberty and moral freedom. The social contract, thus, annuls freedom, but only because Rousseau thinks of it in the terms of Ockham, as an unconditional arbitrariness.

In the current discussion, we cannot ignore David Hume's attitude. Only Ockham's conception makes him regard freedom of will as a fiction. According to the British empiricist, human actions are led by passions by analogy with the attraction and repulsion forces in the theory of Newton. According to Hume, a passion is "a violent and sensible emotion of mind, when any good or evil is presented, or any object, which, by the original formation of our faculties, is fitted to excite an appetite"[14]. In other words, the motive power of each action is desire; i.e. every action always has its a reason (on the principle of association in our minds). If person's actions have a rigid connection with motives or personal characteristics, then they are so caused as the movements of material objects. Identifying freedom with chance, Hume does not find it in human consciousness, that's why he declares it to be the same fiction as the independent existence of self.

Immanuel Kant also shares this point of view, but only when writes about the world of phenomena. This world, which includes, among other things, the actions and mental states of a person (e.g., empirical character), everything that constitutes experience, that is organized by the forms of space and time and by the categories. We can find a reason to every action of a person, it can be deduced from the phenomena of a natural order. Kant clearly states that there is no freedom with respect to this empirical character. However, it is in the world of noumena, lying beyond the forms of time and space and the category of causality. The best proof of this is the existence of morality which is impossible without freedom.

Such, noumenal, freedom is thought by Kant, first of all, as an independence from the laws of nature, but, in turn, as an obedience to those laws that a man "makes himself and in virtue of which his maxims can have their part in the making of universal law (to which he at the same time subjects himself)"[15]. Such freedom cannot be

either learned or understood from any experience, it holds only as "a necessary presupposition of reason in a being who believes himself to be conscious of a will – that is, of a power distinct from mere appetition"[16]. The will is aware of its ability to initiate a chain of causes by the power of reason, i.e. freely; and therefore it is a creator of those laws to which it obeys.

According to Kant, will is a kind of causality of intelligent beings, and freedom is its property, which allows to action independently of extraneous reasons. Freedom, then, turns out to be independence, first of all, from external world which affects the will, that is from sensual inclinations, pleasure, pain and the pursuit of happiness, of any empirical interest, as well as from all other inclinations that do not proceed from the will. Kant, of course, does not forbid a person to wish happiness, as well as take care of his or her health. However, it is not these aspirations, but obligation, that must be taken into account when someone choose an action. We see here that Kant accepts the nominalistic concept of free will.

According to Kant, freedom should be postulated, first of all, because otherwise it is impossible to imagine the action of practical reason, which establishes norms and rules of human behavior. Only being independent of any external factor in relation to it, the reason "exercises causality in regard to its objects"[17]. Otherwise the subject would attribute the determination of his power of judgement to an external to his will impulsion. This is unacceptable, since only one impulsion attaches moral value to an action, namely, respect for the law which the will itself lays down.

Thus, Kant postulates autonomy as a property of will, through which it is a law to itself, independently of the objects for which it wishes, i.e. self-sufficiency of the will. If the will seeks the law that is to determine it in the characters of objects, it is heteronomy. In that case the object gives the low to the will, and it acts according to the laws of nature.[18]

I should note, however, that autonomy, in the form in which Kant postulates it, does not mean the justification for any human actions. It is the first imperative that limits it: "Act only on that maxim through which you can at the same time will that it should become a universal law"[19]. Autonomy is predetermined by the universal nature of the reason, not by certain needs that are the consequences of freely chosen actions. Any action that does not proceed from my reason cannot be truly mine, it is heteronomous. In turn, any action that proceeds from my reason must meet its requirements: I must act only according to the principle that I would like to see as the basis for such actions of all people. This demand of reason, expressed as a categorical imperative, is a formal law (it cannot be formulated in terms of consequences, and it recognizes only one motive – action according to it). Freedom can be postulated only in relation to this demand:

> This amounts to freedom, because acting morally is acting according to what we truly are, moral/rational agents. The law of morality, in other words, is not imposed from outside. It is dictated by the very nature of reason itself. To be a rational agent is to act for reasons. By their very nature, reasons are of general application. […] So if the decision to act morally is the decision to act with the ultimate purpose of conforming my action to universal law, then this amounts to the determination to act according to my true nature as a rational being. And acting according to the demands of what I truly am, of my reason, is freedom.[20]

Morality, therefore, is identified by Kant with the universal demand of reason, to listen to which we are constrained by obligation. The latter is the need for an action dictated by respect for this demand as a law. It is obligation that frees a person from subordinating empirical necessity. Fulfilling the obligation, a person becomes both moral and free at the same time.

It is worth mentioning that Kant's ideas, like also Locke's, formed the basis of the phenomenon that we call Western liberalism. In terms of politics liberalism is the doctrine according to which the right and freedom of the individual have a priority in relation to the intervention of the state in the life of citizens. The main reason for this attitude of the state is the idea of a person as a rational being, capable to learn the truths on one's own, to identify one's will in accordance with that truths, and therefore to direct it to creation and not to destruction, as well as to respect the similar actions of another person. Such a person knows what his or her interest is, and intelligently realizes it in a healthy competition with others, finding a compromise that does not disparage the dignity of any of the parties. Finally, the whole society embodied in the state cannot consider the person as a means for even the highest goals, and therefore violate his or her rights, treat the person as heteronomous in relation to that society.

However, at the origins of liberalism one more ethical direction, the opposite of Kant's deontology, lies, namely, utilitarianism. Proceeding from the understanding of the good as an individual's benefit, and the common good as a collection of interests of members of society, Jeremy Bentham protects the ideas of liberalism in his own way. The state's non-interference in private life, especially the economic one, is necessary, in his opinion, in order that nothing prevents the individual from maximizing pleasure and minimizing suffering.

John Stuart Mill writes about freedom in his work *On Liberty*.[21] He argues there that self-determination and the possibility of choice are the components of the concept of happiness. Any restrictions imposed by the state, in this case, should be justified only by mitigating harm to other members of society. Restrictions of freedom, therefore, are permissible only to ensure safety of its manifestations.

Speaking about freedom, about its modern understanding, we cannot help mentioning the influential concept of Jean-Paul Sartre. Without going into the details of his paradoxical definition of freedom as such that has no essence, I will only reveal those moments that show us that Sartre's understanding inherits the tradition that Ockham laid. Thus I will consider the freedom with respect to the motives and end of the action or choice.

It is important to note that the topic of freedom in *Being and Nothingness* Sartre especially examines in connection with the study of human actions in the world or the realizing of personal project. Such an action always proceeds from a desire as recognizing a certain lack in oneself, and therefore it is intentional, i.e. it has an end. In

turn, the end refers to a cause or a motive. Thus, determinism seems to be justified, however, Sartre dodges from it, arguing that it is action that makes a decision about causes and ends, and action is an expression of freedom. After all, to be free means to determine oneself independently to wanting or choosing: to choosing ends and acting for their achievement.

First of all, Sartre points to the stupidity of disputes between determinists and proponents of freedom as indifferent, unmotivated choice. That proponents, being the followers of Ockham, "are concerned to find cases of decision for which there exists no prior cause, or deliberations concerning two opposed actions which are equally possible and possess causes (and motives) of exactly the same weight"[22]. At first glance, the determinists seem to be more sensible, because the action, in order to be a human one, and not a physical event, must have motivation and end. However, here Sartre puts the question of the very end and the very cause.

Sartre's analysis shows that the end is something nonexistent. A human actions to fill a certain lack, emptiness, in him- or herself. Similarly, the matter is with the cause. It appears only then when a person, being-for-itself, attaches to it the significance of the motive, and it happens not because of referring to some previous being, but because of intentional involvement in nothingness. Thus, not the independent from the being-for-itself cause creates the motive for the action, but the totality of his or her projects (e.g., not suffering, but a decision to change the situation in which a person suffers). Such a motive is an integral part of an action that turns out to be free. The action itself makes a decision about the ends and causes. Only through actions can we get an answer to the question: what we are; and even the feelings we are refer to justify the choice, in fact, are the product of an action. When we evaluate our behavior, only its consequences are important, because per se there are no causes. But there are not the ends also, because the action proceeds from nothingness. Thus, a person is free even when faces with obstacles in the pursuit of aspirations, for the main thing in freedom is not achievements, but the desire itself, proceeding from negativity.

Speaking of the cause, Sartre agrees that it becomes the basis for the action only as a totality of rational considerations that justify it, or as an objective evaluation of the situation. In other words, freedom requires reflection. Through reflection, a person discovers that he or she is a nothingness for him- or herself, and affirms the freedom from everything external, including the cause. Sartre writes: "To adopt Husserl's famous expression, simple voluntary reflection by its structure as reflectivity practices the *έποχή* with regard to the cause; it holds the cause in suspense, puts it within parentheses"[23]. Thus rationality, in spite of the fact that it is necessary for freedom, is nevertheless secondary to it, because proceeds from it as a negation of the reality which exists outside the person (freedom from), in contrast to which the person becomes a being-for-itself.

The same can be said about the will, which is a *négatité* and a force of nihilation. It does not create the end that freedom seeks, but is an instrument for achieving it: "Freedom is nothing but the existence of our will or of our passions in so far as this existence is the nihilation of facticity; that is, the existence of a being which is its being in the mode of having to be it"[24].

The fact that freedom precedes both will and consciousness does not, however, means the chronological precedence. Freedom exists in the very action, in the very choice. Hence it follows that the choice cannot be non-free, i.e. it has not conditions, and the motive of the action is in the action itself. Like Ockham, Sartre comes to the absolutism of freedom. In other words,

> by the sole fact that I am conscious of the causes which inspire my action, these causes are already transcendent objects for my consciousness; they are outside. In vain shall I seek to catch hold of them; I escape them by my very existence. I am condemned to exist forever beyond my essence, beyond the causes and motives of my act. I am condemned to be free.[25]

## 4. Conclusions

Traditional concept of freedom followed mainly Aristotle. It was developed by Christian philosophers of the Middle Ages, first of all, in the teachings of Thomas Aquinas. He didn't consider the teleology of person's action and the conscious motivation for it as barriers for personal freedom.

Ockham defined freedom as the ability to choose indifferently between opposites. Thus he began the tradition not to call free such actions that proceed from the pursuit of happiness, from natural inclinations, personal convictions. In other words, freedom, in his opinion, precedes the self and determines this self. Freedom becomes the defining property of a person.

The influence of Ockham's conception can be seen in such different thinkers as Descartes, Hobbes and Rousseau. The latter argues that freedom as arbitrariness of will is a natural state of human being. Even civil liberty and moral freedom, according to Rousseau, are limitations.

Following Ockham's definition, Hume declares freedom to be a fiction, and Kant agrees with him when it comes to the world of phenomena and the empirical character. Kant postulates the existence of the freedom in the world of noumena. Nevertheless, he still proclaims it incompatible not only with feelings, but also with any inclinations and aspirations that do not arise from the pure will.

On the basis of Kant's concept of autonomy and utilitarian views, so-called Western liberalism was developed. But especially clear the absolutization of freedom appears in the philosophy of the XX-th century, primarily, in existentialism. According to Sartre, not motives and ends determine the actions of the individual, but vice versa. He claims freedom to be the ability to choose these ends. A person assigns both the end and the motive to be what they are.

## Bibliography

Hobbes, Thomas. *Leviathan*. London: Printed for Andrew Crooke, 1651.
Hume, David. *A Treatise of Human Nature*. Garden City: Doubleday, 1961.<http://www.gutenberg.org/files/4705/4705-h/4705-h.htm#link2H_4_ 0080>

Kant, Immanuel. *Groundwork of the Metaphysic of Morals*. Tr. H.J. Paton, New York, Hagerstown, San Francisco, London: Harper & Row, Publishers, 1964.
Mill, John Stuart. *On Liberty*. London: Longman, Roberts & Green, 1869. < https://www.bartleby.com/130/>.
Piechowiak, Marek. „Tomasza z Akwinu egzystencjalna koncepcja osoby i jej godności. Komentarz do Summy teologii, część I, kwestia 29, artykuł 1", in M. Piechowiak, T. Turowski (ed.) *Szkice o godności człowieka*. Zielona Góra: Uiniwersytet Zielonogórski, 2012. 35-47.
Pinckaers, Servais Théodore OP. *Sources of Cristian Ethics.* Tr. M. Th. Noble, Washington, DC: The Catholic University of America Press, 1995.
Rousseau, Jean-Jacques. *The Social Contract.* Tr. G.D.H. Cole, London: J. M. Dent & Sons Ltd., 1923. <http:// www. constitution.org/ jjr/socon_01 .htm>.
Sartre, Jean-Paul. *Being and Nothingness*. Tr. H.E. Barnes, New York: Washington Square Press, 1956.
Taylor, Charles. *Sources of the Self*, Cambridge, Massachusetts: Harvard University Press, 2001.
Thomas Aquinas. *De Malo.* < http:// www.corpusthomisticum.org/ iopera. html>.
–. *Summa theologiae.* < http:// www.corpusthomisticum.org/ iopera. html>.

## Notes

[1] Thomas Aquinas, *Summa theologiae,* I-II, 83, 2.
[2] Cf. Pinckaers, 379-385.
[3] Pinckaers, 358.
[4] Thomas Aquinas, *Summa theologiae,* I-II, 55, 2 co.
[5] Pinckaers, 366.
[6] See Thomas Aquinas, *De Malo*, qu. 6 co; ad 7.
[7] Piechowiak, 46.
[8] Pinckaers, 332.
[9] Cf. Pinckaers, 259-260.
[10] Cf. Pinckaers, 338.
[11] Pinckaers, 338.
[12] See more: Hobbes, 75-82.
[13] Rousseau.
[14] Hume.
[15] Kant, 103.
[16] Kant, 127.
[17] Kant, 116.
[18] Cf. Kant, 108.
[19] Kant, 88.
[20] Taylor, 363.
[21] The 4th edition of the work which is available online: Mill, John Stuart. *On Liberty*. London: Longman, Roberts & Green, 1869.
[22] Sartre, 436.
[23] Sartre, 451.
[24] Sartre, 444.
[25] Sartre, 439.

# Was The Buddha a Reductionist About The Self?

*Peter Donegan*

**Abstract:** This paper examines whether a reductionist view of the self can be found in the *Suttas* of early Buddhism. I will argue that the views of the self exemplified in the *Path of Purification* (*Visuddhimagga*) and the *Abhidhamma* Commentaries are similar to reductionist views of the self put forward by Western philosophers such as David Hume and Derek Parfit. I shall argue that the views of the *Visuddhimagga* can be seen as a legitimate development of the ideas of the *Suttas*. I will also argue however that an opposing view, namely the rejection of a 'realist semantics', can be found in nascent form in the *Suttas*. I will demonstrate that, when legitimately developed, this view can be seen to contradict the reductionist view of the self.

**Keywords:** Buddhism, reductionism, Parfit, Buddhaghosa, Nāgārjuna, chariot simile, *Mūlamadhyamakakārikā,* conventional truth, ultimate truth, *Abhidhamma,* semantic realism, *papañca*.

Reductionism in regards to the self is the view that all facts about a person can be reduced to a set of more particular impersonal facts (Parfit, 1984 p. 210). In Western philosophy these impersonal facts may be facts about the operation of the brain and body. In Buddhism, in particular the Buddhism of the *Abhidhamma* Commentaries of Buddhaghosa, these facts are the operation of the five aggregates (*khandhas*). On this view a 'person' or 'self' is merely a convenient label for this set of deeper facts and a comprehensive account of ultimate reality could be given without reference to 'persons' or 'selves' at all. In this paper I will take the reductionism of Derek Parfit as a standard for what a reductionist position looks like. I will then attempt to address a number of sub-questions before returning to the overarching question of whether the Buddha held a similar reductionist view of the self. I will first outline the standard account of the Buddhist No-Self (*anattā)* doctrine as enumerated by Buddhaghosa and show how this is similar to Western reductionist views of the self such as that of Derek Parfit. In section two I will discuss the objection that the Buddha was not a philosopher at all but rather was a pragmatist attempting to make a soteriological point by teaching *anattā*. I shall argue that this view is mistaken and that the *Suttas* display clear evidence of systematic philosophy. I shall then explore, in section three what some scholars, notably Ronkin, (2005) have taken to be the Buddha's view of language, through a close analysis of the *Madhupiṇḍika Sutta*. I will show how in this *Sutta* seeds of the notion that language is a mere convention are present, I shall argue along with Ronkin (2005) that this rejection of a 'realist semantics' has analogues in the Madhyamaka of Nāgārjuna. I shall take the position of Parfit as a benchmark for what a reductionist position looks like and shall argue that the position of Nāgārjuna is ultimately different in content -- the differences hinging on divergent accounts of language, truth and reality. In the final section I shall fully expound the view of the self found in the *Abhidhamma* Commentaries. I shall show how the seeds of this view can be found in the *Suttas*, particularly the *Vajirā Sutta* and I shall show how the view of the self here is closely analogous to the reductionism of Parfit. I will conclude by arguing that both the 'reductionist realist' position of Buddhaghosa and the 'semantic anti-realism' of Nāgārjuna are supported in the early tradition and that, depending on the text analysed, it is possible to conclude that the Buddha both was and was not a reductionist about the self.

## 1. The Standard View

The standard understanding of the Buddha's teaching of No-Self (*anattā)* as enumerated by Buddhaghosa is as follows: in the *Anattalakkhaṇa Sutta (PTS: S iii 66)* the Buddha explains that there are five aggregates (*khandhas*) which make up a human being: matter or form (*rūpa)*, feeling (*vedanā*), apperception / conceptualisation (*saññā),* volition (*saṅkhāra*) and consciousness (*viññāṇa*). The Buddha asserts that these aggregates are not fit to be regarded as Self (*attā*), this is because, ultimately one cannot absolutely control them, one cannot say 'let my material body be this way!' or 'let me feel this way!' and because they are impermanent (*anicca*) and lead to suffering (*dukkha*).

In brahminical thought it was believed that one could attain universal power through the realisation that the microcosm of the true Self (skt. *ātman*, Pāli: *attā*) and the macrocosm of the universe (*brahman*) were identical. Control over the Self would lead to control over the universe with which it is identical (Collins, 1982 p.97). The Buddha, taking aim at this philosophy argues in the *Anattalakkhaṇa Sutta* that there is no 'inner controller' and that one ultimately does not have control over the constituents of the self. One cannot say 'let my body be thus', 'let me feel thus' rather there are aspects of the body and the self that are beyond conscious control. The two other prongs of the Buddha's argument in the *Anattalakkhaṇa Sutta* are directed against the Upaniṣadic notion of *brahman* as 'being', 'consciousness', and 'bliss' (*satcitānanda*). The Buddha argues that there is no element of the self that is permanently existent and blissful and that all five of the

Indipendent scholar
Oxford, UK.
email: pj_donegan@outlook.com

Philosophical Readings XII.3 (2020), pp. 376-389.
DOI: 10.5281/zenodo.3903513

*khandhas*, including consciousness (*viññāṇa*) are impermanent (*anicca*) and tend towards suffering (*dukkha*). In doing so the Buddha rejects Upaniṣadic concepts of being and the corresponding notions of *ātman* and brahman. Indeed the three hallmarks of existence (*ti-lakkhaṇa)*, impermanence (*anicca*), suffering (*dukkha*) and Not-Self (*anattā*) betray Upaniṣadic reasoning in that they are in direct opposition to the triad of *satcitānanda* (Gombrich, 2009 pp. 69-70).

On this standard view of the No-Self doctrine, exemplified in the teachings of Buddhaghosa the five *khandhas* are seen as constituting an exhaustive list of the series of processes which we conventionally designate by the term 'person': 'it is only these that can afford a basis for the figment of a Self or of anything related to a Self' ( *Visuddhimagg*a XIV. 218) (translation from Siderits 2007, p. 37). Accordingly the *Anattalakkhaṇa Sutta* gives a philosophical analysis of the five *khandhas* as an exhaustive list of the individual and asserts that within the individual no permanent, unchanging essence (*attā)* is to be found.

According to the *Abhidhamma* the only ultimately existing phenomena are those which possess 'intrinsic nature' (Sanskrit: *svabhāva*; Pāli: *sabhāva*). What is only conventionally real disappears under analysis but what is ultimately real cannot be broken down into further parts. To take a well known example a chariot is conventionally real, a conceptual fiction, because upon analysis one can see that a chariot is composed entirely of parts, such as the axle, the spokes, the ropes, the wheels etc. which in turn can be broken down into further parts. Parts which are elementary and cannot be further broken down are termed *dhammas*. This category includes a variety of non-material (mental) *dhammas* and matter (*rūpa*) which can be further analysed in terms of the four primary elements (*cattāri māha bhūtāni*) of solidity (earth), adhesion (water), heat (fire) and motion (wind) (Mendis, 1985 p. 23). Whilst this theory may resemble atomism it is important to note that a fire *dhamma* is merely a momentary occurrence of the property of heat not an atom. Likewise an earth *dhamma* is not a very small substance but rather an 'occurrence of solidity in a very small region of space' (Siderits 2007. p. 15). The *Abhidhamma* asserts that the person can be reduced to a bundle of *dhammas*: matter (*rūpa)*, feeling (*vedanā*), apperception (*saññā),*volition (*saṅkhāra*) and consciousness (*viññāṇa*) and that the 'self' as a conceptual fiction borrows its existence from these five more basic elements. Take away the five *khandhas* and nothing of the 'self' remains. The person is therefore a collection of mental and physical *dhammas* which themselves are impermanent, causally conditioned and subject to dependent origination (*paṭiccasamuppāda)*. According to the parlance of everyday language it is true that 'persons' exist. We can refer to persons, predicate things of them and describe their actions. Nevertheless at the ultimate level of reality, the level of the *dhammas*, it is the case that there are no persons and there is no *attā*. There are only the five *khandhas*. Persons are not 'in the language of "ultimate" truth, ultimately real Existents' (Collins, 1997, p.478). Hence the statement of Buddhaghosa:

'Just as when trunk, branches, foliage, etc., are placed in a certain way, there comes to be the mere term of common usage "tree," yet in the ultimate sense, when each component is examined, there is no tree. ..so too, when there are the five aggregates [as objects] of clinging, there comes to be the mere term of common usage "a being," "a person," yet in the ultimate sense, when each component is examined, there is no being as a basis for the assumption "I am" or "I"' (Visuddhimagga, XVIII, 28) [Trans. Ñāṇamoli (2010)].

This view is a type of mereological *reductionism* (Siderits 2007, p.54). This is the view that parts are real but wholes are not; wholes consist entirely in their parts and can be reduced to them. It is also a version of *realism*, as it assumes that there are ultimately real phenomena (*dhammas*) that can be described using language. Just as we use the word 'chariot' to designate an assemblage of parts (to use an example from the *Vajirā sutta)*, so too we use the term 'self' to designate the *khandhas*. A 'self' is ultimately unreal and all that exists is the psychophysical continuity of the *khandhas*. We use the word 'self' only as a conventional shorthand for this deeper truth.

This *reductionist realist* view of the self has been compared to the views of Western philosophers such as Derek Parfit and David Hume. Parfit argues that a 'person' just consists in the existence of a brain and body and the occurrence of certain mental and physical events and processes such as thoughts, actions and experiences. To use an analogy a 'person' is rather like a 'nation'. A nation consists entirely in a body of people inhabiting a certain territory united by a common history, culture, or language, and is nothing over and above this. Likewise, for Parfit, a person consists entirely in the mental processes of brain and body and is, ontologically, nothing extra (Parfit, 1984 pp. 199- 218). We conceive of persons (or 'selves') as permanent entities above our thoughts and experiences only because of language, which imposes a stable concept of 'person' upon multiple transient processes. Hence:

'Even Reductionists do not deny that people exist. And, on our concept of a person, people are not thoughts and acts. They are thinkers and agents. I am not a series of experiences, but the person who *has* these experiences. A Reductionist can admit that, in this sense, a person is *what has* experiences, or the *subject of experiences*. This is true because of the way in which we talk. What a Reductionist denies is that the subject of experiences is a *separately existing entity*, distinct from a brain and body, and a series of physical and mental events' (Parfit, 1984 p. 233).

Parfit asserts that the Buddha would have agreed with this reductionist realist view of the self (Parfit, 1984 p. 273). Likewise Hume saw persons as a 'bundle' of different perceptions with no enduring substance.

'For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe any thing but the perception.' (*Treatise*, 1.4.6.3)

As such both 'Hume and the Buddha were unable to find a self when they looked within' (Siderits, 2007, p.45).
In the following sections I will examine a number of challenges to this standard understanding and some responses.

In the next section I will first address the question of whether or not the Buddha was a philosopher. I will argue against the view of Thanissaro (2011) that the *anattā* doctrine is a purely practical strategy and I will seek to demonstrate that it is not misleading to characterise the Buddha as a philosopher.

## 2. Was the Buddha a Philosopher?

As reductionism is a philosophical position it will be first necessary to establish whether the Buddha was interested in presenting a philosophically coherent doctrine in regards to the self. There are a number of scholars who argue that he was not. Gombrich (2009, p. 60) argues that to compare the words of the Buddha to philosophers such as Hume is to remove them from their historical context and to potentially obscure them. For Gombrich the Buddha's approach to the problem of suffering (*dukkha*) and its cessation (*nirodha*) is pragmatic. As a consequence the Buddha was not interested in philosophical questions for their own sake (Gombrich, 2009 p. 67).

Gombrich asserts that the Buddha 'condemned all theorising which had no practical value' and 'tended to be quite hard on those who indulged in metaphysical speculations' (Gombrich, 2009 p. 166). Furthermore he 'was not interested in presenting a philosophically coherent doctrine' (Gombrich, 2009 p. 164) but was rather focussed on providing pragmatic advice which was nonetheless underpinned by systematic thought.

Support for the view that the Buddha was not interested in purely philosophical speculation is to be found by considering the unanswered questions (*abyākatā pañhā)*. As explained in the *Cūḷamāluṅkyovada Sutta* (PTS: M i 426) the Buddha refused to comment on metaphysical questions such as:

1.) Is the cosmos eternal?
2.) Is the cosmos not eternal?
3.) Is the cosmos finite?
4.) Is the cosmos infinite?
5.) Are the soul and the body the same?
6.) Is the soul one thing and the body another?
7.) Does the Tathāgata exist after death?
8.) Does the Tathāgata not exist after death?
9.) Does the Tathāgata both exist and does not exist after death?
10.) Does the Tathāgata neither exist nor does not exist after death?

These 'ten indeterminates' (*dasa avyākatavattūni*) 'formed a kind of questionnaire with which the ancient Indians used to confront any religious teacher of note' (Ñāṇananda, 1971 p. 17). The Buddha rejects the questionnaire *in toto* and advises that the questions are the outcome of wrong reflection and hence do not warrant a reply (Ñāṇananda, 1971 pp. 17- 18). The questions are unhelpful and are compared to a man shot with a poisoned arrow enquiring as to the name, height and town of origin of the man who shot him, instead of receiving the necessary treatment. Spending time in speculation, rather than attending to the immediate problem of *dukkha* is of no benefit and will not lead to *nibbāna*. A similar point is made in the *Aggivacchagotta Sutta* wherein the Buddha repudiates the same list of questions as 'a thicket of views' (PTS: M i 483. Trans: Bodhi [2009]).

This view that the *anattā* doctrine is not philosophy is more radically taken up by Thanissaro Bhikkhu. Thanissaro (2011) states that 'the Buddha in teaching not-self was not answering the question of whether there is or isn't a self. This question was one he explicitly put aside' (Thanissaro, 2011 p. 1). Thanissaro argues that the Buddha's teaching just covers two issues, how suffering is caused and the path to its cessation (ibid.) Thanissaro also draws upon the unanswered questions and the simile of the poisoned arrow to assert that the Buddha was only concerned with providing a pragmatic answer to the problem of *dukkha* and nothing further. For Thanissaro the *anattā* doctrine is not meant to answer questions such as 'do I have a self' or 'do I exist' (Thanissaro, 2011 p. 3). Thanissaro argues that this view has some textual support: in the *Ānanda Sutta* (PTS:S iv 400) the Buddha refuses to answer Vacchagotta's question of 'is there no Self?' Similarly in the *Alagaddūpama Sutta* it is stated that the brahmins falsely misrepresent the teaching of the Buddha stating 'the recluse Gotama is one who leads astray; he teaches the annihilation, the destruction, the extermination of an existing being' (PTS: M i 141: Trans.Bodhi [2009]). Thanissaro presents the Buddha not as enquiring philosophically into the nature of the self but rather attempting to answer the question of whether it is beneficial and skillful (*kusala*) to regard anything as 'mine' or 'my self' (Thanissaro, 2011 p. 4). To view the world as 'not-self' is therefore a strategy for helping one avoid unskilful action and put an end to suffering. Thanissaro argues that the Buddha was not attacking the Upaniṣadic notion of the self, nor any other philosophical conception of self rather he completely put aside the philosophical question of whether or not there is a self and in its place recommended that one practically view the world as 'not mine, not myself' as a therapeutic means to eliminate suffering (*dukkha*) (Thanissaro, 2011 p. 10) as a consequence 'the Buddha is not interested in defining who you are or what your self is' (Thanissaro, 2011 p. 13). To support this view Thanissaro refers to the *Siṃsapā Sutta* herein the Buddha states:

'Bhikkhus, the things I have directly known but have not taught you are numerous, while the things I have taught you are few. And why, bhikkhus, have I not taught those many things? Because they are unbeneficial...and do not lead to cessation... to Nibbāna.... And what, bhikkhus, have I taught? I have taught: 'This is suffering'; I have taught: 'This is the origin of suffering'; I have taught: 'This is the cessation of suffering'; I have taught: 'This is the way leading to the cessation of suffering.' (PTS: S v 437. Trans. Bodhi 2000)

Ruegg (1995) notes that the view that the Buddha was not a philosopher claims support from within the *Suttas* themselves, (as already noted) particularly the *Cūḷamāluṅkyovada Sutta* (PTS: M i 426) wherein the Buddha refuses to answer the questions put to him by *Māluṅkyaputta* (Ruegg, 1995 p. 149). The Buddha is the true doctor who pragmatically removes the poisoned arrow rather than dwelling fruitlessly on what sort of person shot the arrow. The Buddha's teaching is therapeutic and questions which do not serve this therapeutic need are

to be excluded as irrelevant. Ruegg notes that the question of whether the Buddha was a philosopher, unsurprisingly, hinges on the definition of philosophy. If philosophy is understood as a pure analysis of concepts, language and meaning, Buddhism as a practical soteriology would not be considered purely philosophical (Ruegg, 1995 p.151). Nevertheless in Buddhism epistemology and metaphysics are closely bound up with the soteriological schema of the Four Noble Truths *(cattāri ariya saccāni*). Ruegg contends that what the Buddha was rejecting was 'disputing for the sake of disputing rather than useful discussion and analysis' (Ruegg, 1995 p. 154). The Buddha rejects disputation and contentiousness for its own sake but not philosophical reasoning *per se*.

I would contend that the Buddha was a philosopher. Van Inwagen notes that a useful definition of metaphysics is as such: 'metaphysics is the study of ultimate reality' (Van Inwagen, 2015, p.1). Perhaps to avoid the ambiguities inherent in the word 'ultimate' we could define metaphysics as 'a general picture explaining the relation between human thought and the world' (Ronkin 2005, p. 247). Though the Buddha was concerned with providing pragmatic advice to realise the cessation (*nirodha*) of suffering (*dukkha)* this does not necessarily discount him from being categorised as a philosopher. This is particularly evident in light of the fact that a sceptic such as Pyrrho (c. 360 – c. 270 BC) who emphasised indifference, impassivity and tranquillity in light of the fact that we must not trust perceptions (Svavarsson, 2010 p. 47), and was perhaps 'an unsystematic moral sage' (Svavarsson, 2010 p. 37), is treated as a philosopher *par excellence*. Similarly Sextus Empiricus, who emphasised the suspension of assent to any belief (advising that tranquillity would ensue as a result), (Pellegrin, 2010 p. 125) is undoubtedly categorised as a philosopher within the Western tradition. If we are to take Ronkin's definition of metaphysics here of 'a general picture explaining the relation between human thought and the world' (Ronkin 2005, p. 247) we can see that some of what is written in the *Suttas* contains what could be termed metaphysics. This is particularly evident when examining the schemas of the aggregates (*khandhas*) and dependent origination (*paṭiccasamuppāda*). The khandhas are an attempt to provide a general picture of human experience by dividing it into matter (*rūpa)*, feeling (*vedanā*), apperception (*saññā),*volition (*saṅkhāra*) and consciousness (*viññāṇa*) and analysing these elements of experience to demonstrate that no permanent self can be found. This, according to Ronkin and Van Inwagen's definitions, would certainly constitute metaphysics. Similarly dependent origination which explains the arising of old age and death (*jarāmaraṇa*) from ignorance (*avijjā*) seems to be in its most basic sense a picture explaining the relation between human thought (and action) and the world.

In addition I would also assert that Thanissaro's (2011) assessment of the *anattā* doctrine is mistaken. In addition to being a practical strategy the *anattā* doctrine is a philosophical position which asserts that a substantial self cannot be found in reality. The Buddha's silence in regards to the unanswered questions can be understood by carefully analysing the texts. In regards to the 'unanswered questions' it is plausible that questions one to four were dismissed purely as pointless speculation but it seems that there must be a further reason for the Buddha's refusal to answer questions five to ten, particularly in the light of the Buddha's teaching of No-Self. Collins notes that in the Commentaries Tathāgata is glossed once as *attā* (UdA. 340) which suggests that the questions regarding the state of the Tathāgata after death are rejected because they assume the existence of an *attā*: 'they use personal referring terms, which according to Buddhist thinking have no real referent' (Collins 1982 p. 133). Philosophically this is a case of 'presupposition failure'. The presupposition here is that there 'ultimately' exists an entity denoted by the word *attā*. This presupposition in Buddhist thinking is false and hence the sentence 'does the Tathāgata exist after death' cannot be meaningfully decided because the terms employed 'relate to null subjects' (Ruegg, 1977 p. 2). At the ultimate level the Tathāgata does not exist, it is therefore not meaningful to ask whether he exists after death. In the same way as the 'soul' does not exist, it is not meaningful to ask whether it is identical with the body. 'The *attā*' is like the son of a barren woman in that there is nothing in reality to which it actually refers.

The No-Self doctrine is therefore a middle way (*majjhimāpaṭipadā*) between the extreme positions of annihilationism (*ucchedavāda*) and eternalism (*sassatavāda*). As explained in the *Brahmajāla Sutta* the core of the eternalist position is that the Self and the world are eternal (PTS: D i 1 ). Eternalism comprises, among others, the view that the true Self (Skrt: *ātman*) and the macrocosm (*brahman*) are identical and change is an appearance (a belief found in the Upaniṣads, notably the *Chāndogya* and *Bṛhadāraṇyaka* and in later Advaita Vedānta) (Bodhi 1978 pp.16-19). Annihilationism, in its common form, teaches that death is the destruction of a substantially existing being and that death brings an end to all individual experience. One form identifies the self with the body and its elements which are dissolved upon death and as such death involves the annihilation of an existent being (PTS: D i 1 ). The No-Self doctrine is not annihilationism because there is no 'Self' to destroy. Nor is it eternalism, as an eternal Self (*attā)* is repudiated. It is rather a middle way between the two, a conventional self is accepted which is nonetheless ultimately a conceptual construction.

The Buddha's refusal to answer Vacchagotta's question of 'is there no self?' in the *Ānanda Sutta* (PTS:S iv 400) can be explained as a refusal to side with the annihilationists.The Buddha refused to reveal the metaphysic of the *anattā* doctrine to Vacchagotta on this occasion because he was already confused. The same *Sutta* states: 'if...I had answered, “There is no self”, the wanderer Vacchagotta, already confused, would have fallen into even greater confusion, thinking, ‘It seems that the self I formerly had does not exist now.' (PTS:S iv 400. Trans. Bodhi [2000]). Bhikkhu Bodhi commentating on the same *Sutta* argues that the reason the Buddha does not state outright that 'there is no self' is because such terminology was used by the annihilationists and he wished to avoid aligning his teaching with theirs in front of the bewildered Vacchagotta (Bodhi, 2000 p. 1457). This point is made more plausible when it is taken into account that some variants of the annihilationist teaching were accompanied by moral nihilism. This is true of the teaching of Ajita

Kesakambalī who taught the extermination of the existing being upon death with the result that there is no karmic fruit of good or bad deeds (Collins, 1982, p. 36). Indeed later in the same *Sutta* after the confused Vacchagotta has left, the Buddha states to Ānanda 'all phenomena are non-self' (*sabbe dhammā anattā)* which is a restatement of the metaphysic of the No-Self doctrine. Far from being a purely practical strategy the anattā doctrine is a metaphysic which states that no enduring, substantial self can be found in reality.

Siderits (2015) notes that the Buddhist *Suttas* were written down several centuries after the Buddha's death and are in a language (*Pāli*) that he is unlikely to have spoken. It is therefore theoretically possible that the Buddha in person rejected metaphysics and that the *Suttas* of the Pāli Canon are not faithful to his original intentions. Nevertheless the *Suttas* are the only comprehensive account we have of the earliest teaching of the Buddha and the only way to identify the 'historical Buddha' is through the lens of the *Suttas*. It cannot be fruitful therefore to speculate on theoretical possibilities such as the *Suttas* being unfaithful to the Buddha's original intentions. I would argue along with Siderits (2015) that the best way to understand the 'original' teaching of the Buddha is to engage with the *Suttas* via the medium of the later Buddhist philosophical tradition. In the remainder of this paper I will discuss two philosophical interpretations of the *anattā* doctrine: that found in the *Abhidhamma*, and that espoused by Nāgārjuna. It is to the latter that I will turn first.

### 3. The Buddha's View of Language as Interpreted by Ronkin, Ñāṇananda and Nāgārjuna

In this section I shall explore the view that the Buddha rejected a 'realist semantics' and how this view conflicts with the reductionist realism of Parfit. I will discuss the works of Gombrich (2009) who has argued that the Buddha rejected the brahminical theory of language and Ñāṇananda (1971) who has argued that the *dhamma* can be conceived as a schema for the 'deconceptualisation of the mind' (Ñāṇananda, 1971 p. 27). I shall also analyse the work of Ronkin (2005) who has argued that the views of Nāgārjuna are closer to the Buddha's original message than the writings of Buddhaghosa. I will show how support for this thesis can be found in the *Suttas*. Finally I shall demonstrate how the view of Nāgārjuna is similar to that of Parfit at the level of conventional (Skrt: *saṁvṛti*) truth, but significantly different at the level of ultimate (Skrt: *paramārtha*) truth.

Gombrich (2009, p. 149) posits that informing the teaching of the Buddha was a fundamental rejection of the brahminical theory of language Mī ̄māṃsā. Mī ̄māṃsā as one of the six schools of orthodox (skt: āstika) philosophy upheld a 'robust form of linguistic realism' (Arnold, 2011 p. 138). For the Mīmāṃsaka the Vedas are authorless, uncreated and eternal, and consequently the Sanskrit language from which they are composed is of an eternal nature. In addition, according to this theory the relation between the Sanskrit word and the state of affairs denoted by the word is non-conventional. The word 'tree' (Skrt: *vṛkṣa)* for example describes an object that embodies a real and objectively existent universal: 'treeness'. The relation between the word and the universal is not arbitrary but is rather a timeless and objectively real relation (Arnold, 2011 p.143). Gombrich argues that, in light of the feature of the Buddhist worldview that all phenomena are impermanent (*anicca*), it is the case that there are no fixed and unchanging entities that can be captured by words. Indeed Gombrich states 'the Buddha concluded not merely that languages were conventional, but that it is inherently impossible for any language to fully capture reality' (Gombrich, 2009, p. 149). It is the case that conceptualisation (*saññā*) which necessarily operates via fixed concepts such as 'man', 'woman', child and 'Buddha' cannot fully capture a reality which is radically transient. Indeed in the *Anattalakkhaṇa Sutta* the Buddha explicitly states that the process of conceptualisation (*saññā*) is impermanent and unsatisfactory. It stands to reason that such an impermanent (*anicca*) and unsatisfactory (*dukkha*) process could never fully capture the reality of the world (*loka*) which is also impermanent and unsatisfactory.

This theme of the inadequacy of conceptualisation (*saññā*) is taken up by Ronkin who argues that the teaching of the early *Suttas* emphasises the inadequacy of the process of thought and by extension language. The point here is that the naming process itself is a product of apperception (*saññā*) and 'making manifold' (*papañceti*) and hence dependently originated. To rely on language to convey ultimate truth is therefore to grasp at conditioned reality (*saṃsāra*) which will result in suffering. The rejection of all speculative opinions by the Buddha therefore amounts, according to Ronkin, to a rejection of the human cognitive apparatus itself and an assertion that mental proliferation (*papañca)* and the process of making manifold (*papañceti*) are the root of the disease of suffering (*dukkha*). (Ronkin, 2005, pp. 246-247). Similarly Gombrich, states 'the very act of conceptualizing, the Buddha held, thus involves some inaccuracy. His term for this was *papañca*' (Gombrich 2009 p.150). Language can point to reality but can never fully capture its fluidity.

Of particular relevance to this point is the *Madhupiṇḍika Sutta*. Here it is asserted that:

> 'What one perceives, that one thinks about. What one thinks about, that one mentally proliferates (*papañca*). With what one has mentally proliferated as the source, perceptions and notions [born of] mental proliferation beset a man with respect to past, future, and present mind-objects cognizable through the mind...Bhikkhu, as to the source through which perceptions and notions [born of] mental proliferation beset a man: if nothing is found there to delight in, welcome, and hold to, this is the end of the underlying tendency to lust, of the underlying tendency to aversion,. . of the underlying tendency to ignorance...here these evil unwholesome states cease without remainder,' (PTS M i 108: Trans.Bodhi [2009])

In order to understand this *Sutta* we must first understand the notion of *papañca*. *Papañca* conveys such meanings as 'spreading out', 'expansion', and 'manifoldness' (Ñāṇananda, 1971 p. 4). It is a comprehensive term 'hinting at the tendency of the...imagination to break loose and run riot' (Ñāṇananda, 1971 p. 4). It can mean 'mental proliferation', 'prolific conceptualisation' or 'mental rambling'.

Papañca tends to obscure subjective experience by giving rise to obsessions. As in the legend of the resurrected tiger who ate the magician that brought it to life, concepts and linguistic conventions overwhelm the man who has evolved them (Ñāṇananda, 1971 p. 6). Language assumes an objective character due to its stability and assails the individual with a tangled net of views *(diṭṭhijāla)*. According to Bodhi *papañca* is 'the propensity of the worldling's imagination to erupt in an effusion of mental commentary that obscures the bare data of cognition' (Bodhi, 2009 p. 1204). Necessarily bound up with mental proliferation is the separation of the world into self and non-self and the accompanying views 'this is mine' (*etaṁ mama*) and 'this is my self' (*eso me atta*). From this matrix of superimposed ego arises views (*diṭṭhi*), thirst (*taṇhā*) and conceit (*māna*). (Ñāṇananda, 1971 pp. 6-12)

The *Madhupiṇḍika Sutta* asserts that in order to end the obsessions of views and the suffering that it entails one must dis-identify with the process of prolific conceptualisation *(papañca)*. The argumentative thrust of the *Sutta* therefore suggests that thinking itself is fundamentally a product of *saṃsāric* conditioned existence and intimately bound up with suffering (*dukkha*). According to Ñāṇananda the *dhamma* can be seen as a schema for the 'deconceptualisation of the mind' (Ñāṇananda, 1971 p. 27): wholesome thoughts are used to drive out unwholesome ones just as a carpenter would use a small peg to knock out a larger one (PTS: M i 118). Yet wholesome thoughts are ultimately to be replaced by wisdom (*paññā*) which is not characterised by discursive thought and is 'immediate and intuitive' (ibid.). Indeed in the *Poṭṭhapāda Sutta* the stages of jhāna meditation are described until the summit of consciousness is reached – the cessation of conceptualisation (*saññā*).

> As he remains at the peak of perception, the thought occurs to him, 'thinking is bad for me. Not thinking is better for me....so he neither thinks nor wills, and as he is neither thinking nor willing, that perception ceases and another, grosser perception does not appear. He touches cessation. This, Potthapada, is … the ultimate cessation of perception [saññā]. (PTS: D i 178. Trans. Thanissaro [2013])

Ñāṇananda argues that the nature of sense experience is such that as soon as one conceptualises the raw sense data of experience one is led astray. The Goal is being able to correctly perceive sense data without being misled by their implications in terms of mental proliferation. According to Ñāṇananda the sense data that enters the mind of the arahant does not reverberate as conceptual proliferation by way of thirst, conceit and views (*papañca- taṇhā- māna- diṭṭhi*). Rather the emancipated sage is characterised by inner stillness, having gone beyond all speculative views (*diṭṭhi*) (Ñāṇananda 1971 pp. 31-39) and particularly having gone beyond the notion of the self (*attā*).

In opposition to the Commentaries which take *diṭṭhi* to mean the sixty-two false views *(micchā- diṭṭhi)* as opposed to the right view *(sammā-diṭṭhi)* of the path (*magga*) Ñāṇananda defines *diṭṭhi* as a 'deep seated tendency.... to be beguiled by concepts' (Ñāṇananda 1971 p. 40). As per the simile of the raft in the *Alagaddūpama Sutta* even the *dhamma* as right view (*sammā-diṭṭhi*) is only a raft for the purpose of crossing over to the far shore of *nibbāna* not for grasping on to. As such one 'should let go even of right mental objects (*dhammas*), to say nothing of wrong ones' (PTS: *M i 130*. Trans: Ñāṇananda ). Right view (*sammā-diṭṭhi)* therefore contains 'the seed of its own transcendence as its purpose is to purge the mind of all views including itself' (Ñāṇananda 1971 p. 43). As the purpose of this practice is to purge the mind of all views (*diṭṭhi*) even those views which are of great help to the practitioner must eventually be expelled. (Ñāṇananda 1971 p. 44). The state of the one who has attained *nibbāna* (as described in the *Nibbāna Sutta*) is such that:

> There is neither earth, nor water, nor fire, nor wind; neither dimension of the infinitude of space, nor dimension of the infinitude of consciousness, nor dimension of nothingness, nor dimension of neither perception nor non-perception; neither this world, nor the next world, nor sun, nor moon. And there, I say, there is neither coming, nor going, nor staying; neither passing away nor arising: unestablished, unevolving, without support [mental object]. This, just this, is the end of dukkha. (PTS: Ud 80. Trans Thanissaro [2012])

The liberated sage who has attained unbinding sees through the concepts of the phenomenal world and experiences emptiness (*suññāta*) (Ñāṇananda 1971 p. 73). The arahant therefore attains 'the utter cessation of the world of concepts' (Ñāṇananda 1971 p.109). The early Buddhist attitude was to 'realise the imperfections of language and logic by observing the internal and external conflict it brought about' (Ñāṇananda 1971 p. 108). It was not necessary to counter every thesis with an antithesis rather the general principle of *dukkha* and its cessation was advanced 'which provides the true impetus for the spiritual endeavour to transcend all theories by eradicating the subjective bias' *(ibid.)*. Ideas of permanence (*nicca*) and Self (*attā*) are what the mind imposes upon objects. The *attā* therefore is something experienced within the mind, it is not a phenomenon 'out there'. The Buddha is primarily attacking the tendency of the mind to subjectively superimpose the notion of an *attā* upon experience. (Ñāṇananda 1971 pp. 103- 108).

From this view of language and reality a radical metaphysics can be discerned – that is if we take metaphysics in its broadest sense to mean 'a general picture explaining the relation between human thought and the world' (Ronkin 2005, pp. 247). The Buddha rejects the notion that the world is composed of existing substances and that language can refer to these substances (Ronkin, 2005, pp. 244-247). This entails a rejection of *semantic realism*. Semantic realism is the view that a statement has an objective truth value in virtue of its relation to an independently existing reality and statements are true or false independent of our means of knowing them (Hale, 1997). This view is to be contrasted with semantic anti-realism which denies that statements are true or false in an objective or ultimate sense. This distinction can be made clearer by examining the realism/anti-realism debate in the philosophy of mathematics: the realist about mathematics would accept the existence of numbers and sets which exist objectively, independent of the human mind. The anti-realist on the other hand would deny the mind-independent ex-

istence of numbers and may assert in its place a constructivist theory in which numbers, sets and the like are fundamentally the product of human thought and do not have a basis in 'mind-independent' reality (Hale, 1997). Likewise the Buddhist semantic anti-realist asserts that truth and language are fundamentally a product of the human mind. Sentences are true or false depending on how they accord with commonly accepted conventions. They have no 'ultimate' or objective truth value independent of the human conventions of language. This semantic anti-realism is the position of the Buddha according to Ronkin, for the Buddha 'reality' is subjective experience and it is not meaningful to speak of a mind-independent reality: 'It is in this fathom-long body endowed with perception and mind that I proclaim the world, the origin of the world, the cessation of the world, and the way leading to the cessation of the world' (PTS: A ii 49: Trans. Bodhi [2012]). According to Ronkin 'the awakened mind breaks up the apparently solid world that we construct for ourselves' (Ronkin, 2005 p. 246) and realises that words and concepts do not ultimately name anything. This is the silence of the eschewal of all views. Hence the statement of the Buddha '"speculative view" is something that the Tathāgata has put away'. (PTS: M i 487. Trans. Bodhi [2012]).

Ronkin posits that by the time of the Abhidhamma Commentaries this view of the nature of language had been lost and the No-Self doctrine had become one of reductionist realism. The *dhammas* becoming phenomena which exist in mind independent reality (in virtue of possessing *sabhāva*) and the 'self' being a convenient label for a bundle of mental and physical *dhammas*. This assumes a realist ontology (the *dhammas* being ultimately real) and a realist semantics in which words bear an objective relation to ultimate reality (Ronkin 2005 p. 249). For Ronkin the *Abhidhamma* dissects experience and brings it within our conceptual and linguistic framework, yet 'this conceptual delineating or giving of boundaries is exactly what the verb *papañceti*, "making manifold" means' (Ronkin 2005 p. 249). The practice of taking concepts to accurately reflect reality and clinging to them as views (*diṭṭhi*) is precisely what the Buddha had warned against in the *Madhupiṇḍika Sutta*.

According to Ronkin 'the earliest Buddhist teaching discloses tenor analogues to Nāgārjuna's Madhyamaka' (Ronkin, 2005 p. 248). Key to understanding the thought of Nāgārjuna is the notion of the two truths: conventional (Skrt: *saṁvṛti*) and ultimate (Skrt: *paramārtha*). In The Fundamental Wisdom of the Middle Way (*Mūlamadhyamakakārikā)* Nāgārjuna states 'the dharma teaching of the Buddha rests on two truths: conventional truth and ultimate truth' (MMK, 24.8) (Siderits and Katsura, 2013 p.272). At the conventional level Nāgārjuna accepts a non-substantial self which can be reduced to the functioning of the five aggregates (Pāli: *khandhas,* Skrt*: skandhas*): matter (Pāli: *rūpa,* Skrt: *rūpa* ), feeling (Pāli: *vedanā,* Skrt: *vedanā*), apperception (Pāli: *saññā,* Skrt: *saṃjñā)*, formations (Pāli: *saṅkhāra,* Skrt: *saṃskāra* ) and consciousness (Pāli: *viññāṇa,* Skrt: *vijñāna* ). This 'self' can persist over time, be reborn, perform skilful and unskilful actions and can reap the karmic fruits of past deeds. Yet unlike Buddaghosa, for Nāgārjuna this 'truth' of a self which can be reduced to the five *khandhas* is only assertable in virtue of the commonly accepted linguistic practices of human beings. It is a conventional truth not an ultimate truth. At the ultimate level there is only the truth of emptiness (*śūnyatā)*.

According to the doctrine of emptiness (*śūnyatā)* nothing in the world exists with inherent existence or 'substance' (*svabhāva)*. Objects possessing s*vabhāva* constitute the 'ontological rock bottom' (Westerhoff, 2017) of the world, and are those entities which exist in their own right. This notion is best understood by use of example: one may assert that 'tables' do not exist in reality but only the atoms which make up the table, the 'table' itself being a conceptual construct. In this example the atoms would possess *svabhāva*. Nāgārjuna asserts that ultimately nothing in the world exists with s*vabhāva* and hence asserts that nothing exists objectively or in its own right – everything is empty. For the Mādhyamika, ultimately, there is no 'way things objectively are', there is no inherent and objective structure to reality independent of the human mind. The Mādhyamika will not therefore accept semantic realism, this being the theory that statements are objectively true or false in virtue of their relation to independently existing reality. What is asserted in its place is a conventional theory of truth in which statements can be asserted only if they agree with commonly accepted practices and conventions (*ibid*.). For Nāgārjuna even the ultimate truth of emptiness is itself empty – it is not ultimately true in the sense that it corresponds to the way reality 'really is' independent of human conventions. There are not objects in the world which have the property of being empty. Rather the doctrine of emptiness is rather an attempt to stop the mistaken ascription of s*vabhāva* to phenomena by the human mind (Westerhoff, 2017).

This can be investigated by exploring the Dispeller of Disputes (*Vigrahavyāvartanī*) of Nāgārjuna. Herein Nāgārjuna states 'I do not have any thesis...while all things are empty... free from substance, from where could a thesis come' (Westerhoff, 2010 pp. 29-30). According to this passage Nāgārjuna states that ultimately he does not have any substantial thesis, that is any thesis which can be described according to a realist semantics – a semantics wherein statements can capture real and objectively existing entities in the world. All statements are conventional as the very notion of a realist semantics is ruled out by the Madhyamaka theory of emptiness (Skrt: *śūnyatā*) (Westerhoff, 2010 pp. 64-65). Propositions are only applicable to those entities which are a product of mental proliferation *(papañca,* Skrt *prapañca*) (Ruegg, 1977 pp. 12). In The Fundamental Wisdom of the Middle Way (*Mūlamadhyamakakārikā*) Nāgārjuna states 'independently realised, peaceful, unobsessed by obsessions, without discriminations and a variety of meanings: such is the characteristic of truth' (MMK 18.9) (Kalupahana, 1986 p. 270). Nāgārjuna here is making an epistemological point, according to the Buddha, man searching for permanence misses the impermanent nature of experience. In referring to 'peace' (Skrt: *śānta*) and 'obsession' (Skrt: *prapañca*) (MMK 18.9) in this context Nāgārjuna is referring to the peace of mind attained by one rid of the delusion of permanence, free from the obsession (Skrt: *prapañca*) of the search for ultimate truths in terms of existence or non-

existence (Kalupahana, 1986 p. 272). This notion mirrors Ronkin's interpretation of the *Madhupiṇḍika Sutta* which warns against the same 'obsessions' or mental proliferations (*papañca*) and the misleading nature of conceptualisation.

When we assess the view of the No-Self doctrine as presented by Nāgārjuna in this light we can see that there are key differences between this doctrine and reductionist notions of the self as espoused by philosophers such as Derek Parfit. Parfit claims that the facts of personal identity 'just consists in the holding of certain more particular[ that is more basic] facts.' (Parfit, 1984, p. 210). This view is firmly *realist*, it assumes that there are certain facts about the way the mind-independent world is and that these facts can be fully described using language. Parfit states: 'we could give a complete description of *reality* without claiming that persons exist' [emphasis added] (Parfit, 1984 p.212).

At the conventional level the views of Parfit and Nāgārjuna are similar, both accept a non-substantial self which can be reduced to either the processes of brain and body or the *khandhas*. Yet at the ultimate level Nāgārjuna does not accept reductionism as he rejects a realist semantics. For Nāgārjuna concepts can only conventionally, not factually describe the world. This is in contrast to the view of Parfit (and as I shall more fully demonstrate later, Buddhaghosa), who assumes that ultimately the external world can be described objectively and factually. Parfit describes the reductionist position as such: 'the *fact* of a person's identity over time just consists in the holding of certain more particular *facts*' (Parfit, 1984 p. 210). With this in mind we must conclude that at the ultimate level (as opposed to the conventional level) Nāgārjuna is not a reductionist about the self. At the ultimate level Nāgārjuna rejects the notion of an objective 'factual' reality, he therefore does not accept the key tenet of reductionism that 'the fact of a person's identity over time just consists in the holding of certain more particular facts'. For Nāgārjuna ultimately there are no such facts, there is only emptiness: 'for the Mādhyamika not only is there no substantial self, there is also no substantial basis on which a nonsubstantial self could be built' (Westerhoff, 2009, p. 163). This is in opposition to the reductionist view of Buddhaghosa –who holds the *khandhas* to be the substantial basis upon which a non-substantial self is built – and Parfit who holds the processes of brain and body to be such a basis. It is the case therefore that Nāgārjuna accepts the No-Self doctrine but rejects the reductionist No-Self doctrine of Buddhaghosa as an ultimate truth.

A strong case can be made that the views of Nāgārjuna can be traced back to the early *Suttas*. Kalupahana presents Nāgārjuna not as an innovator but rather as a grand commentator on the early Pāli *Suttas* who did not try to improve upon the original teachings of the Buddha but rather explicated them in their original form (Kalupahana, 1986 p. 5). For Kalupahana Nāgārjuna's philosophy is a restatement of the original philosophy of the Buddha (Kalupahana, 1986 p. 8). According to the Pāli *Kaccāyanagotta Sutta* '"everything exists"...is one extreme' and '"everything does not exist"...is the second extreme' 'without approaching either extreme the Tathāgata teaches you a doctrine by the middle' (PTS: S ii 16. Trans. Kalupahana, 1986, pp. 10-11). This middle way is the teaching of dependent arising (*paṭiccasamuppāda*) which is the view that phenomena (that is elements of experience) give rise to other phenomena under certain causal conditions and these phenomena cease when the relevant causal conditions are absent. This is the mechanism by which the Buddha was able to explain phenomena without resorting to the concept of a permanent entity (Kalupahana, 1986 p. 34). Nāgārjuna wrote the *Mūlamadhyamakakārikā* both as a commentary on the *Kaccāyanagotta Sutta* (Kalupahana, 1986 p. 5) and as an attempt to discredit 'heterodox' schools particularly the Sautrāntikas and Sarvāstivādins (Kalupahana, 1986 p. 26). In the *Mūlamadhyamakakārikā* the key themes of early Buddhism are taken up not, according to Kalupahana, to reject them but rather to rid them of metaphysical explanation in terms of absolute existence (eternalism) or non-existence (annihilationism) (Kalupahana, 1986 p. 29) these are the same views that the Buddha was rejecting in the *Kaccāyanagotta Sutta*.

For Nāgārjuna belief in substantial entities such as the Self (Skrt: *ātman*) gives rise to grasping, possession and in turn obsessive mental proliferations (Skrt: *prapañca*). This process can be stopped by the recognition of emptiness (*śūnyatā*) in regards to the *ātman* and all phenomena (Skrt: *dharmas*) (Kaluphanana, 1986 p. 56). Antecedents to this emptiness doctrine can be found in the *Suttas* of the Pāli canon, notably the *Pheṇa Sutta*. Herein it is said:

'Bhikkhus, suppose that this river Ganges was carrying along a great lump of foam. A man with good sight would inspect it...it would appear to him to be void, hollow, insubstantial. For what substance could there be in a lump of foam? So too, bhikkhus, whatever kind of form there is, whether past, future, or present, internal or external, gross or subtle, inferior or superior, far or near: a bhikkhu inspects it, ponders it, and carefully investigates it, and it would appear to him to be void, hollow, insubstantial. For what substance could there be in form?' (PTS: S iii 140 Trans. Bodhi [2000]).

'Form is like a lump of foam, Feeling like a water bubble; Perception is like a mirage, Volitions like a plantain trunk, And consciousness like an illusion, So explained the Kinsman of the Sun. "However one may ponder it and carefully investigate it, it appears but hollow and void when one views it carefully".' (PTS: S iii 140 Trans. Bodhi [2000]).

Commentating on this *Sutta* Bhikkhu Bodhi notes that the imagery relating to the empty nature of conditioned phenomena was taken up by the Mādhyamikas (Bodhi, 2000, pp. 1085-6). The emptiness doctrine of Nāgārjuna is therefore to be found in a nascent form in the early *Suttas*. The seeds of this doctrine in the *Suttas* giving rise to the fully grown sprout of the Madhyamaka philosophy of emptiness.

It is also interesting to note that the tetralemma (*catuṣkoṭi*) found in the *Cūḷamāluṅkyovada Sutta* (PTS: M i 426) was taken up by the Mādhyamikas 'to establish the inapplicabilty of any imaginable conceptual position...that might be taken as the subject of an existential proposition' (Ruegg, 1977 p. 9) and therefore become the basis for a set of doctrinal extremes. The tetralemma is as follows:

1.) Does the Tathāgata exist after death?
2.) Does the Tathāgata not exist after death?

3.) Does the Tathāgata both exist and does not exist after death?
4.) Does the Tathāgata neither exist nor does not exist after death?

For Nāgārjuna the Buddha rejected the tetralemma in regards to the state of the Tathāgata because his existence was asserted in a 'real and absolute sense' (Kalupahana, 1986 p. 309) it is a substantial view of the self which leads to speculations as to the state of the Tathāgata after his death and this in addition to any corresponding speculations should be rejected. Hence according to the *Mūlamadhyamakakārikā* 'when he is empty in terms of self-nature, the thought that the Buddha exists or does not exist after death is not appropriate' (MMK 22.14) (Kalupahana, 1986 p. 309). Just as the Tathāgata is devoid of self-nature so too 'the universe is also devoid of self-nature' (MMK 22.16) (Kalupahana, 1986 p. 310). Nāgārjuna asserts '"empty", "non-empty", "both" or "neither" -these should not be declared. It is expressed only for the purpose of communication' (MMK 22.11) (Kalupahana, 1986 p. 307). Here Nāgārjuna rejects any theorising regarding emptiness, using the four cornered logic of the *catuṣkoṭi*. The terms 'empty' or 'non-empty' are only used for the purpose of communication and should not be reified and taken as something possessing essential nature (Skrt: *svabhāva*) (Kalupahana, 1986 p. 308). This should be seen as the rejection of concepts as 'incorruptible reals' (*ibid*.) and as an assertion of a conventional and pragmatic theory of truth and language.

Based on the above it can be surmised that the views of Nāgārjuna can be traced back to the *Suttas* of the early Canon. At the very least it is possible to assert that the seeds of the Madhyamaka emptiness doctrine and corresponding theory of language can be found in the *Suttas*, particularly the *Pheṇa Sutta* and the *Madhupiṇḍika Sutta,* seeds which bore their fully developed fruit in the Madhyamaka philosophy. For Nāgārjuna as the world is empty of *svabhāva* there is no fixed way the world ultimately 'is'. This worldview results in a rejection of a realist semantics. As I have previously asserted this view is very different to the reductionism of Parfit, who accepts a realist semantics, or at the least assumes it in *Reasons and Persons*.

I would assert that it can be maintained that there is an element of reductionism in the philosophy of Nāgārjuna at the level of conventional truth. This is because Nāgārjuna accepts a conventional self which can be reduced to the functioning of the five *khandhas*. Nevertheless for Nāgārjuna at the ultimate level all *dhammas* are empty including the *khandhas*. This doctrine of emptiness has the implication of the rejection of a realist semantics rendering the position of Nāgārjuna very different to that of reductionism as espoused by Parfit. Parfit asserts that the self is unreal (or, more specifically, real only in virtue of the way we use language), whereas the parts, that is the psycho-physical processes of brain and body are ultimately constituents of reality. Nāgārjuna on the other hand asserts that all of reality is empty of *svabhāva* and hence there is no fixed structure to 'reality' at all. Whereas there is a conventional self which can be reduced to the *khandhas* ultimately even the *khandhas* do not possess 'reality' or inherent existence (*svabhāva)*.

In the next section I will examine the reductionism of the *Abhidhamma* and the *Abhidhamma* Commentaries of Buddhaghosa. I will show how this view of the self is reductionist in the truest sense. I will assert that the view found in the *Abhidhamma* Commentaries displays more similarities to the reductionism of Parfit than the views of Nāgārjuna.

## 4. The Position of the Abhidhamma and its Commentaries.

In this section I will outline the view of the Buddhist doctrine of No-Self (*anattā)* as exemplified in th*e Path of Purification* (*Visuddhimagga*) of Buddhaghosa and the *Abhidhamma* Commentaries. I will show how this view is similar to reductionist views of the self put forward by Western philosophers such as David Hume and Derek Parfit.

The *Abhidhamma* which can be translated as 'about (*abhi-*) the *dhamma*' (Anālayo, 2014 p. 78) can be seen as growing out of the early followers of the Buddha's psychological need to make the teachings of the Buddha as comprehensive as possible (Anālayo, 2014 p. 169) and to clarify any ambiguity or incomplete aspects of the original teaching. This drive to give a complete picture of all aspects connected with the Buddhist Path grew interdependently with the idea that the Buddha attained omniscience upon his awakening (*ibid*.). The tradition of the Commentaries holds that the *Abhidhamma* was expounded by the Buddha to an assembly of deities (*devas*) in the *Tāvatiṃsa* heaven, and hence is the authentic word of the Buddha himself (*buddhavacana*) (Bodhi, 2000A p.11). Yet the *Abhidhamma* may have had more mundane origins and Anālayo asserts that the *Abhidhamma* seems to have had its origin as a communally recited commentary on the *Suttas* which gradually evolved into a 'higher' teaching: 'abhi-' itself taking on the meaning of 'higher' rather than simply 'about' (Anālayo, 2014 p. 116).

The Theravādin *Abhidhamma* is comprised of seven books the *Dhammasaṅgaṇi* (Enumeration of Dhammas), the *Vibhaṅga* (Analysis), the *Dhātukathā* (Discourse on Elements), the *Puggalapaññatti* (Designation of Persons), the *Kathāvatthu* (Points of Discussion), the *Yamaka* (*Pairs*), and the *Paṭṭhāna* (Causal Conditions) (Ronkin, 2017). They are treatises in which the doctrine of the *Suttas* have been systematised, tabulated and meticulously organised (Bodhi, 2000A p. 2). The cornerstone of the *Abhidhamma* is the *dhamma* theory which maintains that the 'fundamental components of actuality' (Bodhi, 2000A p. 3), that is the building blocks of ultimate reality are *dhammas*, 'momentary mental and material phenomena which constitute the process of experience' (ibid.). Reality is constructed upon the foundation of the *dhammas* which possess, ultimate reality and determinate existence (*sabhāva*) 'from their own side' (*sarūpato*) (ibid.). Bodhi notes that in the *Abhidhamma* the *dhamma* theory is not 'expressed as an explicit philosophical tenet' (ibid.) but is rather implicit in the texts. This implicit doctrine being expressed as a fully fledged theory in the Commentaries of Buddhaghosa. These Commentaries are the *Atthasālinī* (The Expositor) the commentary on the *Dhammasaṅgaṇi*;

the *Sammohavinodanī*, (The Dispeller of Delusion) the commentary on the *Vibhanga*; and the *Pañcappakaranaṭṭhakathā*,the commentary on the remaining five treatises (Bodhi, 2000A p. 13). In the *Atthasālinī* it is written 'there is nothing else [but *dhammas*], whether a being, or an entity, or a man or a person' (*Atthasālinī* 155) (Trans. Ronkin 2017). The principle here being that the phenomenal world at the ultimate level is a world of *dhammas* and there is no other ultimate reality apart from the reality of the *dhammas* (Ronkin, 2017). Bodhi maintains that such a conception of reality is already present in nascent form in the *Suttas* (Bodhi, 2000A p. 3), particularly in the Buddha's analysis of the five aggregates (*khandhas*) (PTS: S iii 66), the analysis of the six sense bases *(saḷāyatana)* (PTS: M iii 215), and the enumeration of dependent origination (*paṭiccasamuppāda*) (PTS: S ii 2 ). Here there is a tendency to analyse experience into its fundamental building blocks and from this foundation explain the arising of suffering causally.

In a statement that resembles Parfit's reductionist view of the self Bodhi describes the *Abhidhamma* project as attempting to drive a wedge between ultimately real entities, that is the *dhammas*, and those entities which exist only as conceptual constructs, such as the self (Bodhi. 2000A p. 4). The *Abhidhamma* primarily seeks to comprehend the nature of experience and the reality that is the focus of the *Abhidhamma* is the conscious reality of human experience *(ibid.)*. Nevertheless by the time of the *Abhidhamma* Commentaries *dhammas* come to be taken to exist in mind-independent reality and are invested with ontological significance, the *dhammas* being those entities which comprise the ultimate constituents of reality in virtue of being invested with *sabhāva* (Ronkin 2005 pp. 248-249). Ronkin asserts that the worldview of the Commentaries resembles Leibniz's monadological metaphysics (*ibid.*) which is the view that everything in the world is composed of simple substances, which form the foundation of reality (Look, 2017). This latter view present in the Commentaries is therefore a fully fledged reductionism with the acceptance of a realist semantics. The self does not exist in reality but is rather a conceptual fiction. Just as a 'nation' can be reduced to those entities which exist in reality e.g. human beings located in a certain geographical area acting in certain ways, so too the 'self' can be reduced to those entities which actually exist in reality: for Parfit these are the psycho-physical processes of brain and body, for Buddhaghosa it is the *dhammas*.

As the *Abhidhamma* Commentaries accept the position that there are ultimately real entities, that is those entities possessing *sabhāva* a strong case can be made that the *Abhidhamma* Commentaries accept a realist semantics. The *Abhidhamma* Commentaries accept that there is a way ultimate reality is and that language can describe this ultimate reality. Statements about the self will be true or false depending on their relation to this ultimate reality of *dhammas*. Ronkin asserts that the Commentators 'endow the final products of their analysis with the status of ultimate facts', the *dhammas* are understood as the building blocks of reality and to understand that reality is composed of *dhammas* is to understand the way things really are (Ronkin, 2005 p. 119). The *Abhidhamma* metaphysics paves the way for the acceptance of a realist semantics (Ronkin 2005, p. 153): the theory that a statement has an objective truth value in virtue of its relation to an independently existing reality. The position of the *Abhidhamma* Commentaries is therefore a reductionist realism that is similar in structure to the view of Parfit. Just as for Parfit the 'self' or a 'person' can be reduced, ultimately to facts about the brain and body (Parfit, 1984, p. 210) and presumably further to a set of deeper facts about atoms, electrons, quarks etc. In the same manner in the *Abhidhamma* Commentaries all facts about the self can be reduced to impersonal facts about the *dhammas*. The *dhammas* can be understood conceptually and can be described using language with sentences which are ultimately true or false. This is reductionism in the truest sense and the views of the *Abhidhamma* Commentaries and of Parfit display striking similarities in this respect.

Bodhi asserts that the *Abhidhamma* does not proclaim a new doctrine not found in the *Suttas* (Bodhi, 2000A p. 5). The difference between the *Abhidhamma* and the *Suttas* is one of scope and method (*ibid.*). In the *Suttas* the Buddha makes use of conventional parlance for pragmatic reasons in order to guide his audience which had differing capacities for understanding his message. The *Abhidhamma* on the other hand 'rigorously restricts itself to terms that are valid from the standpoint of ultimate truth' (*paramatthasacca*) (Bodhi, 2000A p. 6) -- though this distinction is not absolute and the *Suttas* themselves contain discourses strictly relating to ultimate truth. Bodhi, in opposition to the views of Ronkin asserts that whilst it is tempting to trace some historical development of ideas between the *Abhidhamma* and the Commentaries this line of thought should 'not be pushed too far' (Bodhi, 2000A p. 14). This is due to the fact that the *Abhidhamma* requires the Commentaries to give context and provide a unified and systematic meaning to the material. Bodhi contends that without the Commentaries important elements of meaning would be lost and as such the Abhidhamma and the Commentaries should be taken as a whole (Bodhi, 2000A pp.13-14).

The reductionist realism which is to be found in the *Abhidhamma* Commentaries is not necessarily a novel development not found in the *Suttas*. Rather it is the case that the seeds of the reductionist realist position can be found in the *Suttas* themselves, as even Wynne, (an opponent of the reductionist realist position) notes (Wynne 2010 p. 157). These seeds developed into the fully fledged reductionist realism of the *Abhidhamma* Commentaries.. In the Vajirā *Sutta* the nun Vajirā exclaims to Māra, the Evil One:

> Why now do you assume ‘a being’? Māra, is that your speculative view? This is a heap of sheer formations: Here no being is found. “Just as, with an assemblage of parts, The word ‘chariot’ is used, So, when the aggregates exist [*khandhesu santesu*], There is the convention ‘a being.’ “It’s only suffering that comes to be, Suffering that stands and falls away. Nothing but suffering comes to be, Nothing but suffering ceases (PTS: S i 134 Trans. Bodhi 2000)

Here the reductionist realist view is stated explicitly: it is conventionally true that there is a self as a collection of *khandhas*. Nevertheless it is an ultimate metaphysical truth that 'no being is to be found'. The statement of Vajirā here implies that there is a way reality is ultimately

structured and that language is capable of describing this reality: the '*Vajirā Sutta* is both reductionistic as well as realistic, for it speaks of the aggregates "existing" (*khandhesu santesu*) and of the failure to "find" an essential being in them (*na yidha sattūpalabbhati*)' (Wynne 2010 p. 157). The presence of such a clear statement of reductionist realism in the early *Suttas* calls into question the view that a radical shift in metaphysical outlook can be traced between the philosophy of the *Suttas* and the philosophy of the *Abhidhamma* Commentaries of Buddhaghosa. Indeed Buddhaghosa echoes the sentiment of the *Vajirā Sutta* precisely in the *Visuddhimagga:*

'Therefore, just as when the component parts such as axles, wheels, frame poles, etc., are arranged in a certain way, there comes to be the mere term of common usage "chariot," yet in the ultimate sense when each part is examined there is no chariot... —so too, when there are the five aggregates [as objects] of clinging, there comes to be the mere term of common usage "a being," "a person," yet in the ultimate sense, when each component is examined, there is no being' Visuddhimagga (XVIII, 28) [Trans. Ñāṇamoli (2010)]

In addition antecedents of the *dhamma* theory can be traced back to the *Suttas*. The *Mahāhatthipadopama Sutta* (PTS: M i 184) provides an exhaustive list of the human being in terms of the earth, wind, fire, water and space elements. Whilst asserting that these basic elements exist it concludes that no independent *attā* can be found upon analysis of the constituents of a human being. The *Sutta* therefore espouses a realist ontology, the elements ultimately existing whereas the *attā* does not (Wynne 2010 p. 159). This echoes the later *dhamma* theory in which form (*rūpa*) is further analysed in terms of the four primary elements (*cattāri māha bhūtāni*): the *dhammas* of solidity (earth), adhesion (water), heat (fire) and motion (wind) (Mendis, 1985 p. 23). These elements constituting the fundamental building blocks of material reality. The *Sutta* also accepts a variant of a realist semantics: 'the ultimate truth of things is here captured in words, and is not something beyond logic and the conceptual construction of consciousness' (Wynne, 2010 p. 159-160). Similarly the statement all phenomena are not Self *(sabbe dhammā anattā )* found many places in the canon, such as in the *Dhammaniyāma Sutta (*PTS:A i 286), and the *Ānanda Sutta* (PTS:S iv 400) is most readily interpreted in line with the later *dhamma* theory of the *Abhidhamma*. Rahula notes that by the phrase 'all *dhammas* are not Self' the Buddha is explicitly claiming that the *attā* is nowhere to be found in reality (Rahula 1959, p.58) – a statement which seems to assume that there is a fixed external reality which is ontologically composed of the *dhammas* but in which ultimately no *attā* is to be found. On the evidence of the *Vajirā Sutta* we must accept that the reductionist realism of Buddhaghosa can be traced back to the early *Suttas* themselves – although this position may be explicitly stated rarely and may not be as fully developed as in the *Abhidhamma* Commentaries. This view of the self found in the *Abhidhamma* Commentaries shares a great deal of similarities with Western reductionist views of the self such as Parfit's.

Like the Buddha Parfit is arguing against substantialist views of the self. In Western philosophy perhaps the most famous proponent of the substantialist view of the self is René Descartes. Descartes in his Book 6 of the *Meditations* asserts that the mind and body are of different substances, because it is possible to imagine the existence of the mind without the body (Descartes, 2008 pp. 73-83). The essence of the human being is intellect and the mind is a purely thinking substance separate from the body. Hence the famous Cartesian slogan '*cogito ergo sum*' (I think therefore I am) (Descartes, 2008 p. 13). Parfit claims that the notion of a Cartesian Ego is intelligible only if this Ego were to manifest itself empirically, but that no such Cartesian Ego is to be found upon empirical investigation (Parfit, 1984 p. 227). In addition Parfit argues that it is not intelligible to argue that Cartesian Egos exist but are not empirically observable (Parfit, 1984 p. 228).

This line of reasoning is similar in style to the reasoning of the Buddha in the Second Sermon (PTS: S iii 66). The Buddha asserts that if the Self (*attā*) existed it would be observable, yet upon examining the entirety of human experience through the schema of the five *khandhas* the Buddha demonstrates that no permanent blissful *attā* is to be found, rather all phenomena are impermanent (*anicca*) and unsatisfactory (*dukkha*). Furthermore it is unintelligible to assert that the *attā* exists but is not connected with human experience in any way. This notion is expressed in the *Mahānidāna Sutta* here it is stated 'where nothing whatsoever is sensed (experienced) at all, would there be the thought, "I am"?' (PTS: D ii 55. Trans Thanissaro, [2013]). This passage is interpreted by Harvey to be rejecting the possibility of a transcendent *attā* above and beyond the empirical experience of the five *khandhas*. (Harvey, 1995, pp.31-33)

According to the *Abhidhamma* Commentaries what the Buddha asserts does exist is a conventional self which is dependent upon and can be entirely reduced to the *dhammas* of the five *khandhas*. It is the dhammas which are, ontologically, ultimately real. What guarantees personal identity across time according to this view is that the *khandhas* of one conventional self at time x-1 stand in the appropriate causal relation to the *khandhas* of the same conventional self at time x. This continuity of karma is compared in the *Milindapañha* (3.2.6) to the fruit (in this case a mango) arising from the seed planted in the ground. Though the mango and the seed are not physically identical it is the case that, conventionally, they are the same mango because the fruit stands in the appropriate causal relation to the seed. In the same manner the karmic seeds sown by one conventional self at time x-1 produce causal conditions which produce the karmic fruit for the same conventional self at time x.

Parfit also accepts the existence of a conventional self which ultimately can be reduced to a set of impersonal facts about the brain and body. He states 'though persons exist, we could give a complete description of reality without claiming that persons exist.' (Parfit 1984 p.212). This apparent paradox is resolved by appealing to the way we use language: 'facts about people cannot be barely true. Their truth must consist in the truth of facts about bodies, and about various interrelated mental and physical events ' (Parfit, 2011 p. 424). He illustrates this by means of example. Suppose I have detailed scientific informa-

tion about a group of trees growing on a hill. I then learn the further piece of information that this group of trees can be referred to as a 'copse'. In this instance I have not learnt any new factual information about external reality but a fact about the way we use language. The case is similar for persons, all the non-linguistic facts about a person can be reduced to psychological and physical facts about the brain and body. The fact that we call this bundle of psychological and physical processes a 'person' is only a fact about language. (Parfit, 2011 p. 424). This resembles the already quoted passage of Buddhaghosa: just as when the axles, wheels and frame are arranged in a certain way there is the linguistic convention to call this aggregation a chariot, so too when there are the *khandhas* there is a person *Visuddhimagga* (XVIII, 28). Indeed Parfit makes expicit comparison of his view to that of the Buddha stating:

'I claim that, when we ask what persons are... the fundamental question is a choice between two views. On one view, we are separately existing entities, distinct from our brain and bodies and our experiences.... The other view is the Reductionist View. And I claim that, of these, the second view is true. As Appendix J shows, ***Buddha would have agreed*** [emphasis in original]. The Reductionist View is not merely part of one cultural tradition. It may be, as I have claimed, the true view about all people at all times.' (Parfit, 1984 p. 273)

Parfit goes on in Appendix J to quote among other texts the *Visuddhimagga (XVIII 3 31):* 'The mental and material are really here, but here there is no human being to be found' (Parfit, 1984 pp. 502-3) and the chariot simile in the already quoted *Vajirā Sutta*. Thus stressing himself the similarities between Buddhist reductionism and his own position.

Nevertheless the notion of karmic continuity is alien to much of Western philosophy and it is here that key differences arise between the reductionism of Buddhaghosa and the reductionism of Parfit. For Parfit personal identity over time is established by the degree of psychological continuity between entities over time. Parfit terms the relation of psychological continuity 'relation R'. Collins notes that in Buddhism by contrast relation R need not hold. Sentient beings may be reborn in a manner that is completely psychologically disconnected from their previous life. One who is reborn may (and frequently does according to the tradition) have no memories of their past life. In addition to this a being who was a human may be reborn as an animal with strikingly dissimilar cognitive capacities (for example a human and an ant) and hence any psychological continuity from one being to the next must be ruled out (Collins, 1997 p. 482). From this we may conclude that what matters in Buddhism is not Relation R, psychological continuity, but rather *karmic* continuity, what Collins terms Relation K (*ibid*.).

It is the case that there are large similarities between the views of Parfit and Buddhaghosa and it can be said that both are reductionist realists about the self. Parfit claims that all facts about persons can be reduced to impersonal psychological and physical facts about the world: 'the *fact* of a person's identity over time just consists in the holding of certain more particular *facts*' (Parfit, 1984 p. 210). This view accepts semantic realism – it accepts that there is a way 'ultimate reality' is structured and that language can refer to this reality with statements that are true or false. Similarly for Buddhaghosa ultimately reality is structured by the *dhammas*, those entities possessing *svabhāva*. As the *dhammas* can be described using language and sentences containing propositions about the *dhammas* will be true or false depending on how they relate to ultimate reality, the view of Buddhaghosa also accepts semantic realism. Similarly both Parfit and Buddhaghosa accept a conventional self. In both views a self is a conceptual fiction that is superimposed upon more basic elements of ultimate reality: for Parfit this reality is the processes of brain and body, for Buddhaghosa this ultimate reality is the *khandhas*.

Nevertheless, as is to be expected, there are some key differences between the two views, for Buddhaghosa the criterion for determining personal identity over time is karmic continuity, whilst for Parfit it is relation R, psychological continuity. As I have shown however the two views are close enough in character to warrant informative comparison and the reductionism of Parfit can be meaningfully used to elucidate the reductionism of Buddhaghosa. Finally, I have argued that the reductionist position of the *Abidhamma* Commentaries can be traced back to the *Suttas* themselves. This is particularly true of the *Vajirā Sutta* which is explicitly reductionist in content, using the well-known chariot metaphor which was later taken up by Buddhaghosa. On the evidence of this *Sutta* therefore it is possible to conclude that seeds of the fully developed reductionist position are already present in the *Suttas* and therefore that the Buddha was a reductionist about the self.

## Conclusions

I have argued that it is correct to categorise the Buddha as a philosopher and that the *Suttas* contain elements of moral philosophy and what could be termed metaphysics. Explicitly metaphysical doctrines being those of dependent origination ( *paṭiccasamuppāda)* and the five aggregates (*khandhas*). I have asserted that the No-Self doctrine itself is a middle way (*Majjhimāpaṭipadā*) between the philosophical positions of eternalism (as found in the Upaniṣads) and annihilationism. I have raised the common objection that the Buddha shuns metaphysical speculation as counterproductive as in the *Cūḷamāluṅkyovada Sutta* and have responded that the Buddha rejects disputation and contentiousness for its own sake but not philosophical reasoning *per se*. I have explicitly argued against the views of Thanissaro (2011) who has put forward the position that the *anattā* doctrine is a purely practical strategy with no metaphysical implications. In response to this position I have argued that the Buddha's silence in response to certain metaphysical questions must be understood in the context of the situation. In the *Ānanda Sutta* the Buddha refuses to answer Vacchagotta's question of 'is there no self?' because a denial of the self here would be mistaken for annihilationism by the already bewildered Vacchagotta. Similarly the Buddha refuses to answer questions relating to the state of the Tathāgata after death because such questions are a case of presupposition fail-

ure. They assume the existence of a substantial entity denoted by the term Tathāgata and 'use personal referring terms, which according to Buddhist thinking have no real referent' (Collins 1982 p. 133).

I then went on to discuss the views of Gombrich (2009), Ronkin (2005) and Ñāṇananda (1971). I noted that Gombrich (2009) holds that the Buddhist philosophy of language constitutes a rejection of Mīmāṃsā. I discussed the notion of mental proliferation (*papañca* ) found in the *Madhupiṇḍika Sutta* and argued that this idea could be interpreted as pointing towards the inadequacy of conceptualisation ( *saññā*). *Papañca* for Ronkin points towards a rejection of semantic realism. For Ronkin 'The awakened mind breaks up the apparently solid world that we construct for ourselves' (Ronkin, 2005 p. 246) and realises that words and concepts do not name anything. I then explored the thesis that the rejection of *papañca* found in the *Madhupiṇḍika Sutta* anticipates the philosophy of language of Nāgārjuna. Nāgārjuna asserts that the 'characteristic of truth' (MMK 18.9) is free from mental proliferation (Skrt: *prapañca* Pāli: *papañca*) and Westerhoff (2017) interprets Nāgārjuna as rejecting a realist semantics. As no entities possess s*vabhāva* there is no way 'things objectively are' hence language and truth are only conventions. I then argued that the views of Nāgārjuna can be traced back to the *Suttas* of the early Canon and that the seeds of the Madhyamaka emptiness doctrine and corresponding theory of language can be found in the *Suttas*, particularly the *Pheṇa Sutta* and the *Madhupiṇḍika Sutta*. I concluded by demonstrating that whilst conventionally Nāgārjuna is a reductionist he is ultimately not a reductionist about the self as he rejects semantic realism. Parfit states the definition of reductionism as such: all facts about a person can be reduced to a set of more particular impersonal facts (Parfit, 1984 p. 210). Whereas Nāgārjuna accepts a conventional self which is constituted by the five *khandhas* Nāgārjuna does not accept that there is ultimately a fixed structure to reality which can be described using language. Nāgārjuna therefore rejects the view that there is a set of 'more particular impersonal facts' that the conventional self can ultimately be reduced to. He does not assent to the existence of such facts and asserts in its place the view of emptiness. It is the case therefore that Nāgārjuna accepts the No-Self doctrine but rejects the reductionist No-Self doctrine of Buddhaghosa and Parfit at the level of ultimate truth. I concluded by arguing that if we are to follow Nāgārjuna's interpretation of the Buddha's teaching it would be legitimate to assert that the Buddha was not a reductionist about the self at the level of ultimate truth.

I then examined the view of the self found in th*e Path of Purification* (*Visuddhimagga*) of Buddhaghosa and the *Abhidhamma* Commentaries. I demonstrated that in the Commentaries there is an acceptance of a realist semantics: The *dhammas* can be understood conceptually and can be described using language with sentences which are ultimately true or false. All facts about the self can be reduced to impersonal facts about the *dhammas* which are ultimately real and possess s*vabhāva*. I then argued that this view is strikingly similar to that of Western reductionists such as Parfit. Just as for Parfit all the facts about a person can be reduced to psychological and physical facts about the brain and body so too for Buddhaghosa all conventional facts about the self can be reduced to deeper facts about the *dhammas*. In addition both Parfit and Buddhaghosa accept a conventional self based on how we use language. Whilst there are some differences between the two positions such as between karmic continuity and relation R the positions are fundamentally similar. I also argued that this reductionist realism can also be traced back to the *Suttas* particularly the *Vajirā Sutta*.

Was then the Buddha a reductionist about the self? I have argued that seeds of both the reductionist realist position of Buddhaghosa and the semantic anti-realism of Nāgārjuna can be found in the early *Suttas,* specifically the *Vajirā Sutta* and the *Madhupiṇḍika Sutta* respectively. When taken in their fully developed form these ideas are in contradiction to one another, or as Nāgārjuna would assert, express different levels of truth. For Nāgārjuna reductionism is true at the conventional level but ultimately there is only emptiness. For Buddhaghosa reductionism represents the way reality 'really is' and hence is an ultimate truth. There is therefore a tension within the early tradition, there are seeds of certain ideas which when developed in certain ways yield mutually incompatible fruits. The Buddha of Buddhaghosa is a reductionist and a realist. The Buddha of Nāgārjuna ultimately rejects semantic realism and refuses to speak of an 'objective reality'. This former position would be recognisable to a Western reductionist about the self such as Parfit whereas the latter position would not. Both interpretations have textual support within the tradition and therefore it is possible to conclude that the Buddha of the *Suttas* was both a reductionist and not a reductionist about the self depending on which *Sutta* is focused on.

## Bibliography


Bhikkhu Anālayo. (2014). *The Dawn of The Abhidharma*. Hamburg: Hamburg University Press.

Arnold, D. (2011). Mīmāṃsā in *The Oxford Handbook of World Philosophy*. Oxford: Oxford University Press.

Bhikkhu Bodhi. (1978). *The All Embracing Net of Views*. The Brahmajāla Sutta and its Commentaries. Kandy: Buddhist Publication Society.

– (2000A). *A Comprehensive Manual of Abhidhamma*. Seattle: BPS Pariyatti Edition.

– (2000). *The Connected Discourses of The Buddha: a New Translation of The Saṃyutta Nikāya*. Boston: Wisdom Publications.

– (2009). *The Middle Length Discourses of The Buddha: A Translation of The Majjhima Nikāya*.
Boston: Wisdom Publications.

– (2012). *The Numerical Discourses of The Buddha: A Translation of The Aṅguttara Nikāya*.
Bristol: Pāli Text Society.

Buddhaghosa (2010). *The Path of Purification (Visuddhimagga)* Translated from the Pāli by Bhikkhu Ñāṇamoli. Columbo: The Buddhist Publication Society.

Collins, S. (1982). *Selfless Persons: Imagery and Thought in Theravāda Buddhism*. Cambridge: Cambridge University Press.

– (1985). Buddhism in Recent British Philosophy and Theology. *Religious Studies* 21, pp. 475- 493.

– (1997). A Buddhist Debate About the Self; and Remarks on Buddhism in The Work of Derek Parfit and Galen Strawson. *Journal of Indian Philosophy*, 25, 5, pp. 467-493.

Descartes, R. (2008). *Meditations on First Philosophy* translated by Mike Moriarty. Oxford: OUP.

Garfield, J. (1995). *The Fundamental Wisdom of The Middle Way:Nāgārjuna's Mūlamadhyamakakārikā*. New York: OUP

Glanzberg, M. (2016), Truth, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.).

https://plato.stanford.edu/archives/win2016/entries/truth/>.
Gombrich, R. (2009). *What the Buddha Thought*. London: Equinox Publishing Ltd.
Hale, B. (1997). Realism and its Oppositions, in B. Hale and C. Wright (eds.) *A Companion to the Philosophy of Language*. Oxford: Blackwell.
Harvey, P. (1995). *The Selfless Mind: Personality, Consciousness and Nirvana in Early Buddhism*. Richmond: Curzon Press.
Hume D. (2014) [1738]. *A Treatise of Human Nature* in Sir Lewis Amherst Selby-Bigge, and P. H. Nidditch (eds) David Hume: A Treatise of Human Nature. Oxford: OUP.
Kalupahana, D. (1986). *Na¯ga¯rjuna The Philosophy of The Middle Way*. Albany, N.Y.: State University of New York Press.
Look, B. (2017). Gottfried Wilhelm Leibniz,*The Stanford Encyclopedia of Philosophy,* Edward N. Zalta (ed.).
https://plato.stanford.edu/archives/sum2017/entries/leibniz/.
Mendis, N.G.K (1985). *The Abhidhamma in Practice*. Kandy: Buddhist Publication Society.
Bhikkhu Ñāṇananda. (1971). *Concept and Reality in Early Buddhist Thought*. Kandy: Buddhist Publication Society.
Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
– (2002). *Reductionism and Personal Identity* in Philosophy of Mind: Classical and Contemporary Readings, D. Chalmers (ed.). Oxford: OUP.
– (2011). The Unimportance of Identity in *The Oxford Handbook of the Self*, S. Gallagher (ed.). Oxford: OUP.
Pellegrin, P. (2010). Sextus Empiricus. In: R. Brett, ed.,*The Cambridge Companion to Ancient Scepticism*. Cambridge: Cambridge University Press.
Rahula, W. (1959).*What the Buddha Taught*. Bedford: Gordon Fraser.
Rhys Davis, T. W. & Stede W. (1952). *The Pāli Text Society's Pāli-English Dictionary*. London: The Pāli Text Society.
Rhys-Davids, T. (2012). *The Milindapañha*. Altenmu¨nster: Jazzybee Verlag.
Ronkin, N. (2005). *Early Buddhist Metaphysics*. London: Routledge-Curzon.
– (2017). Abhidharma. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.).
https://plato.stanford.edu/archives/sum2017/entries/abhidharma
Ruegg, D. (1977). The Uses of the Four Positions of the "Catus-koti" and the Problem of The Description of Reality in Mahāyāna Buddhism. *Journal of Indian Philosophy*, 5, 1-2, pp. 1-71.
– (1995). Some Reflections on The Place of Philosophy in The Study of Buddhism. *Journal of the International Association of Buddhist Studies* 18, 2, pp. 145-181.
Siderits, M. (2007). *Buddhism as Philosophy: An Introduction*. Aldershot: Ashgate.
– (2015). Buddha. *The Stanford Encyclopedia of Philosophy,* Edward N. Zalta (ed.). https://plato.stanford.edu/archives/spr2015/entries/buddha/>.
Siderits, M. and Katsura, S. (2013). *Nāgārjuna's Middle Way: The Mūlamadhyamakākarikā*. Boston: Wisdom Publications.
Svavarsson, S. (2010). Pyrrho and early Pyrrhonism. In: R. Brett, ed., *The Cambridge Companion to Ancient Scepticism*. Cambridge: Cambridge University Press.
Bhikkhu Thanissaro. (2011). *Selves & Not-Self*. CA: Metta Forest Monastery.
– (2012). *Nibbāna Sutta* (Ud 8.1), translated from the Pali by Thanissaro Bhikkhu. *Access to Insight*.
http://www.accesstoinsight.org/tipitaka/kn/ud/ud.8.01.than.html
– (2013). *Potthapada Sutta* (DN 9), translated from the Pali by Thanissaro Bhikkhu. *Access to Insight*.
http://www.accesstoinsight.org/tipitaka/dn/dn.09.0.than.html.
– (2013). *Maha-nidana Sutta* (DN 15), translated from the Pali by Thanissaro Bhikkhu. *Access to Insight*.
http://www.accesstoinsight.org/tipitaka/dn/dn.15.0.than.html.
Van Inwagen, P. (2015). *Metaphysics*. Boulder: Westview Press.
Westerhoff, J. (2009). *Na¯ga¯rjuna's Madhyamaka: A Philosophical Introduction*. Oxford: OUP
– (2010). *Na¯ga¯rjuna's Vigrahavyāvartanī The Dispeller of Disputes*. Oxford: OUP.
– (2017). Nāgārjuna. *The Stanford Encyclopedia of Philosophy* Edward N. Zalta (ed.).
https://plato.stanford.edu/archives/spr2017/entries/nagarjuna/>.
Wynne, A. (2010). The Atman and its Negation. *Journal of the International Association of Buddhist Studies* 33, 1-2, pp. 103–171.

# Semantic Externalism and Justified Belief about the External World

*Hamid Alaeinejad and Seyyed Mohammad Ali Hodjati*

**Abstract:** Philosophical skepticism about the external world seeks to call into question our knowledge of the external world. Some kinds of philosophical skepticism employ skeptical hypotheses to prove that we cannot know anything about the external world. Putnam tried to refute this kind of skepticism by adopting semantic externalism; but, as is now generally accepted, Putnam's argument is epistemically circular. Brueckner proposes some new, "simple" arguments that in his view are not circular. In this paper we evaluate Brueckner's simple arguments for refuting skepticism about the external world, and seek to demonstrate that they fail to prove that we can have knowledge about the external world. However, by appeal to the principle of privileged access, one of the Brueckner's arguments does indeed succeed in showing that we can have justified beliefs about the external world.

**Keywords:** Skepticism, Semantic Externalism, Privileged Access, Brueckner, Epistemic Circularity.

## 1. Putnam's Approach to Skepticism

According to the philosophical skepticism, propositions about the external world are not the possible objects of knowledge. In order to prove its claim, the general method of skepticism is first to establish a distance between, on the one hand, the evidence that could justify belief in a proposition, and on the other hand the truth of that proposition; and then to show that the evidence does not entail the truth. Accordingly, skeptics set out their claims in the form of a skeptical argument (Brueckner 1994: 827). In most cases, these arguments are based on hypotheses known as skeptical hypotheses. The content of such skeptical hypotheses includes a description of the world which is (i) different from the one we normally consider, (ii) consistent with our usual experiences, yet (iii) not distinguishable from the real world. One of the most famous of these hypotheses is the "Brain in a Vat" (BIV) scenario, which posits that humans are simply brains immersed in nutrients, whose (apparent) perceptions are being transmitted to their neural terminals by an advanced computer. Then the skeptic observes that assuming that we know some ordinary proposition about the external world is inconsistent with the fact that we do not know whether a skeptical scenario obtains or not; and from this, the skeptic concludes that we cannot know any proposition about the external world (Pritchard 2002: 217–18). Putnam argued that by accepting the thesis of semantic externalism, we can show that we know that we are not brains in a vat, and therefore block the skeptical result (Putnam 1981: 15). According to semantic externalism, the meaning and truth conditions of our propositions and the content of our intentional mental states are in some manner determined on the basis of the external, causal environment; in other words, the environment is effective in determining the content of intentional states (Kallestrup 2012: 62). In that sense, two persons may have identical intrinsic properties, but because of the difference in the environments in which they are located, differ as regards the content of their mental states. Putnam argues that by accepting externalism, the statement "I am a brain in a vat" comes out as false, independently of being expressed in the real world or in the world inside the vat, and therefore I do know that I am not a brain in a vat.

Many criticisms have been developed against Putnam's externalist argument, of which the most important —and the one with which this paper is concerned— turns on the claim that Putnam's argument is epistemically circular. By providing a detailed explanation of Putnam's argument, Brueckner (1986) sought to show what is needed to answer this criticism. He introduced some new versions of the externalist argument that are known colloquially as his "simple arguments." These arguments use other philosophical principles to avoid the problem of circularity. In this paper we evaluate these simple arguments, first carefully explaining the problem of epistemic circularity for Putnam's argument, and then examining Brueckner's proposals in its defense.

## 2. Epistemic circularity and Putnam's externalism

It is now generally accepted that Putnam's argument against skepticism is indeed epistemically circular (Brueckner 1986, Wright 1992, Davies 1995, Noonan 1998, Johnsen 2003). In response, philosophers have proposed alternative arguments that avoid this deficiency. In order to show precisely how Putnam's argument is epistemically circular, we focus on Brueckner's account of it. Brueckner's account, which he calls a "disjunctive argument," is based on the logically true proposition "Either I am a BIV or I am not" (Brueckner 1986: 154). According to Putnam's argument, if I am a BIV, the sentence "I'm a brain in a vat" is false because in that case, the words "brain" and "vat" refer not to the brain and vat, but to the brain* and vat*.[1] Also, if I'm not a BIV, the sentence

Department of Philosophy, Faculty of Humanities, Tarbiat Modares University
Tehran, Iran.
emails: hamid.alaeinejad@gmail.com ; hojatima@modares.ac.ir

Philosophical Readings XII.3 (2020), pp. 390-393.
DOI: 10.5281/zenodo.3903542

"I'm a brain in a vat" will be false. But Putnam wants to conclude that I'm *not* a brain in a vat.[2] So the argument is not complete: it is necessary to add to the above argument a premise such as the following, which will lead to the desired result:

(TC) The sentence "I am not a brain in vat" is true if and only if I am not a brain in a vat

But this leads the argument into epistemically circularity, because unless we can know that we are in the real world, we cannot know that the truth condition of the sentence (TC) is that I am a brain in a vat. If I *am* a brain in a vat, the truth condition of (TC) will be that I am a brain* in a vat*. Indeed, more generally we may observe that if it has been proven that a sentence *p* is true, the knowledge of the content of *p* does not necessarily result; from a proof that *p* is true we cannot conclude that we know the proposition that *p*.[3] Therefore, Putnam's argument against skepticism cannot properly show that I know that I am not a brain in a vat, and therefore does not succeed in rejecting skepticism (Brueckner 1986: 164–65).

### 3. Brueckner's simple arguments

Brueckner presents several externalist arguments that seek to avoid the problem of epistemic circularity, and at least at first glance do not have the complexity of Putnam's arguments; hence they are known as simple arguments. Consider the following argument (Brueckner 2012: 6, 2016: 21):[4]

*Brueckner's Simple Argument against Skepticism 1 (SA1)*

(A) If I am a BIV, then my use of the word "tree" does not refer to trees
(B) My use of the word "tree" refers to trees – So,
(C) I am not a BIV [(A),(B)]

Premise (A) comes from Putnam's semantic externalism. In the ordinary world, someone who uses the word "tree" refers to real trees because they have been causally associated with real trees in the external world. But a BIV has never been associated with real trees, and hence the truth conditions of the sentences (the meaning of the words) that a BIV expresses are different from the truth conditions of those sentences (those words) when asserted in the ordinary world. So if I am a BIV and say the word "tree," the word does not refer to trees, but to tree*s. Premise (B) suggests that when I use the term "tree," I am considering the real tree in the real world and I refer to it. Therefore, it is concluded that I am not a BIV. Brueckner (2010: 161) has also given another simple argument as follows:

*Brueckner's Simple Argument against Skepticism 2 (SA2)*

(A) If I am a BIV, then I am not thinking that trees are green
(B) I am thinking that trees are green
(C) So I am not a BIV

Again, Premise (A) comes from Putnam's semantic externalism. Since the BIV does not refer to real trees when it uses the term "tree," if it honestly states that it believes that trees are green, the content of its belief is not that real trees are green. A BIV cannot think that trees are green, but can only think that tree*s are green*. At first it seems that premise (B) causes the argument to be circular because if I am a BIV I cannot think that trees are green. But, based upon the principle of privileged access, this premise is justified. According to the principle of privileged access, when our faculty of introspection is functioning properly, we can know what we are thinking by introspection (McLaughlin and Tye 1998: 350). In other words, if we use our common abilities in the formation of second-order beliefs, then if we think that *p*, we can know that we are thinking that *p*. This knowledge is *a priori* and we do not depend on empirical examination of the outside world to achieve it. Such knowledge is not justified experimentally. So from these two premises we can conclude that I am not a BIV. We will return to this issue in the next section.

But these simple arguments still encounter problems. In the case of SA1, two criticisms can be made: one critique is that the use of the premise (A) causes the argument to be epistemically circular. The skeptics can claim that this premise is based on the assumption that the word "tree," in the language used in the vat or in the ordinary world, refers to *something*. But the point is that we can only know that the word "tree" refers *a posteriori*; but since the skeptical arguments are not based on experience, the use of such a premise is not permissible: hence the argument is epistemically circular. The second critique is raised against premise (B). According to the premise, I know that my word "tree" refers to trees; but the skeptic's claim is that we do not know whether we live in the real world or not, and therefore we do not know that we are faced with real trees. In fact, the claim that the word "tree" in our language refers to trees is based on accepting the anti-skeptical position, and hence the argument is epistemically circular. Brueckner accepts this critique, but states that it is possible to rewrite the premise and avoid this critique (Brueckner 2010: 159) [5]. He claims it is true that I do not know whether I am a BIV or a human in the ordinary world, but it is not the case that I do not know anything about the language I use. In fact, I know that if the word "tree" does refer in my language, it refers to trees. This is *a priori* knowledge of the semantics of my language, and hence to appeal to it is not begging the question (Brueckner 2012: 8–9). Brueckner then tries to solve this problem by rewriting these premises in a conditional way. He says that, drawing on externalism, I can claim that I know *a priori* that if the word "tree" refers, it refers to something with which I casually interact; however, I do not know *a priori* that the word refers to something. So we can rewrite SA1 as follows:

*Brueckner's Simple Argument against Skepticism 3 (SA3)*

(A) If I am a BIV, then it is not the case that if my word "tree" refers, then it refers to trees
(B) If my word "tree" refers, then it refers to trees –So,
(C) I am not a BIV

But in our view, although rewriting the premises of SA1 in conditional form responds to the criticism, we are still not justified in accepting premise (B) in SA3. We are justified in accepting this premise only if we know that we are not a BIV. The premise says that if the word "tree" refers to something, it will refer to the real trees. But how could we know this? True, I know that if the word "tree" refers, it refers to the *trees in my world*, but the problem is that I do not know in which world I live. So it is true that if the word "tree" refers in my language, it will refer to *trees in my world*, but this does not mean that it refers to real trees in a real world. So this argument is still circular. In fact, if we want to make an non-circular argument, then instead of premise (B) in SA3 we should use the following two premises:

**($B_1$)** If I am not a BIV, then if the word "tree" refers in my language, it refers to trees

**($B_2$)** If I am a BIV, then if the word "tree" refers in my language, it refers to tree*s

It is clear that from these two premises we may not conclude that I am not a BIV. Therefore, SA3 is also epistemically circular and does not work.

Argument SA2 faces a similar problem. Premise (B) in SA2 claims that I think that trees are green; but if I am a BIV I can only think that tree*s are green*. The claim that I know that I think trees are green again causes the argument to be epistemically circular, because if I do not know that I am not a BIV, I cannot know that I am thinking about not tree*s but trees. In fact, although premise (A) in SA2 seems to be acceptable, premise (B) causes the argument to be circular.

One can answer this problem, however, by using the principle of privileged access. In this case, our justification for accepting the premise (B) in SA2 is based on privileged access, and premise (A) in this argument is justified according to semantic externalism; so SA2 is not circular. But there remains another problem, which the following section addresses.

## 4. Privileged access and semantic externalist arguments against skepticism

At least at first glance, it does not seem possible to accept both the privileged access thesis (henceforth PA) and semantic externalism (henceforth SE). According to SE, the meaning of our words is partly determined by the reference of our words in the external world; so external experiences are needed in order to find out what the meaning of our words is, and thus to know what we are thinking. But according to PA, this kind of self-knowledge does not come about through empirical investigation, and we know the meaning of our words without any external experiences.[6] For example, consider someone who is thinking that drinking water quenches thirst. According to PA he knows what he is thinking about; but semantic externalists say that he does not know what the content of his belief is until he experimentally finds out what the word "water" refers to in the world he lives in. Consequently, the question arises of whether PA is intrinsically incompatible with SE, or whether this incompatibility is achieved, for example, by unjustified use of an epistemological principle; in which this case these two theses are not themselves incompatible.[7] Apart from the answer given to this question, it can be argued that the acceptance of the externalist arguments presented here is based on accepting the compatibility of these two theses. It is clear that if, in some way, it can be shown that PA and SE are incompatible, then the use of them in a single argument is not acceptable. There are compelling reasons in favor of SE, and PA is also intuitively acceptable, but there is no argument for the compatibility of these two theses. In fact, philosophers have gone to great effort to refute the arguments which have been presented to prove the incompatibility of PA and SE. But now for the sake of the argument, we assume that none of the arguments seeking to prove the incompatibility between PA and SE are successful, and that externalists have indeed shown that incompatibilist arguments are not sound.

Now the question is whether externalists can reject skepticism using PA to answer the circularity problem in SA2. We think that the answer to this question is negative: because from the fact that incompatibilist arguments are rejected, it does not follow that SE and PA are compatible; it only shows that they are *not in*compatible. In order to employ an argument that uses both SE and PA, we need to *know* that these two are compatible. To explain this, suppose that in a valid argument both SE and PA have used as premises and a conclusion $p$ has been drawn. If these premises are incompatible, the proof that $p$ is not epistemically significant, because from contradictory premises any result, including $p$, can be deduced. On the other hand, in order to be able to conclude that I know that $p$, we must also *know* the premises in order to, using the principle of epistemic closure, conclude that we know that $p$. If we know PA and SE, we do not need to have a separate proof for their compatibility, because knowing them would require their truth; but the problem is that we do not know PA and SE, but we are only *justified* in accepting them. Therefore, it is possible that, while accepting the two principles is reasonable, their combined application would lead to inconsistencies.

We would like to address a potential objection to the claim that since we do not know that SE and PA are compatible, we could not employ an argument that uses both SE and PA. Somebody may argue: you say any semantic externalist relying on privileged access must prove their compatibility. Call the idea that one cannot rely on the idea that SE and PA are compatible unless it has been proven that they are The Compatibility Principle. And call the idea that epistemic circularity is not accepted The Not-Accepted Principle. You rely on both The Compatibility Principle and The Not-Accepted Principle in your critique of SA2. The problem is that no one has proven these two principles are incompatible but no one has proven they are compatible either. Therefore your critique fails. But the answer is simple: we do not *know* that using SA2 we could not refute skepticism but we are justified to accept it. Since skepticism seeks to call into question our *knowledge* of the external world, our critique shows that using PA to answer the circularity problem in SA2 cannot refute skepticism about the external world. In other words, by using PA in SA2 we cannot conclude that we

can *know* propositions about the external world. But still we can justifiably accept them.

So, it can be said that the mere rejection of the arguments presented to prove the incompatibility between PA and SE does not establish their compatibility. Therefore, their compatibility is not proven. But since PA and SE are based on arguments that rely on strongly accepted intuitions, as long as their incompatibility has not been proven, we are justified in using them in a single argument. But it should be noted that we can no longer claim that we *know* the result of such an argument: in this case we are only *justified* in accepting the result.

We claimed above that SA2 is not successful in blocking skepticism about the external world. In this argument, premise (A) is based on the acceptance of SE and premise (B) is based on the acceptance of PA. But, as stated, this argument can only show that I am *justified* in accepting that I am not a BIV, but I still do not *know* that I am not a BIV. Therefore, according to what has been said, the use of the combination of SE and PA in an argument cannot lead to a rejection of skepticism about the external world. However, although such arguments do not rule out skepticism about the possibility of acquiring knowledge about the external world, by using both PA and SE we can show that we are justified in believing propositions about the external world. So SA2 is successful, at least, in blocking skepticism about the *justification* of our beliefs about the external world.

## 5. Conclusion

The aim of the present paper was to examine whether Brueckner's "simple arguments" were successful in blocking skepticism about the external world. Our evaluation showed that the first and third arguments (i.e., SA1 and SA3 respectively) are epistemically circular, and therefore unacceptable. Also, the second argument (i.e., SA2) uses both PA and SE, and we do not *know* that these two are compatible; so SA2 is not successful in blocking skepticism about the external world. However, SA2 succeeds in blocking the skepticism the justification of our beliefs about the external world.

## Bibliography


Brueckner, A., (1986), "Brains in a Vat," *Journal of Philosophy*, Vol. 83, No. 3, pp. 148–167.

Brueckner, A., (1994), "The Structure of the Skeptical Argument," *Philosophy and Phenomenological Research*, Vol. 54, No. 4, pp. 827–835, International Phenomenological Society.

Brueckner, A., (1999), "Semantic Answer to Skepticism," *Skepticism, a Contemporary Reader*, edited by Keith Derose and Ted A. Warfield, pp. 43–60, New York: Oxford University Press.

Brueckner, A., (2010), *Essays on Skepticism*, Oxford: Oxford University Press.

Brueckner, A., (2012), "Skepticism and Content Externalism," *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL= http://plato.stanford.edu/ archives/spr2012/ entries/ skepticism-content externalism/

Brueckner, A., (2016), "Putnam on Brains in a Vat," *Brain in a Vat*, edited by Sanford C. Goldberg, pp. 19-26, Cambridge: Cambridge University Press

Burge, T., (1988), "Individualism and Self-Knowledge", *Journal of Philosophy*, Vol 85, No. 11, pp. 649–663.

Davies, D., (1995), "Putnam's Brain-Teaser," *Canadian Journal of Philosophy*, Vol 25, Num. 2, pp. 203–228.

Falvey, K., Owens, J., (1994), "Externalism, Self-Knowledge, and Skepticism," *Philosophical Review*, Vol. 103, No. 1, pp. 107–137.

Johnsen, B., C., (2003), "Of Brains in Vats, Whatever Brains in Vats May Be," *Philosophical Studies*, Vol 112, pp. 225–249, Netherlands: Kluwer Academic Publishers.

Kallestrup, J., (2012), *Semantic Externalism*, New York: Routledge.

McLaughlin, B., P., Tye, M., (1998), "Is Content-Externalism Compatible with Privileged Access?" *Philosophical Review*, Vol. 107, No. 3, pp. 349–380.

Noonan, H., W., (1998), "Reflections on Putnam, Wright and Brains in Vats," *Analysis*, Vol. 58, No. 1, pp. 59–62.

Pritchard, D., (2002), "Recent Work on Radical Skepticism," *American Philosophical Quarterly*, Vol. 39, No. 3, pp. 215–257.

Putnam, H., (1981), "Brains in a Vat," *Reason, Truth and History*, pp. 1-21, Cambridge: Cambridge University Press.Tymoczko, T., (1989), "In Defense Of Putnam's Brains," *Philosophical Studies*, Vol. 57, pp. 281–297.

Wright, C., (1992), "On Putnam's Proof That We Are Not Brains-in-a-Vat," *Proceeding of the Aristotelian Society*, New Series, Vol. 92, pp. 67–94.


## Notes

[1] If I am a BIV, there is no real vat in my world but rather something else, perhaps electrons in a certain pattern, that causes my perceptions. So in a vat world I casually interact with something that is not a real vat. Following Brueckner, in order to distinguish between the real vat and the vat in the vat world, we call the latter the "vat*".

[2] Note that if the sentence "I am a brain in a vat" is false, then the sentence "I am not a brain in a vat" is true.

[3] Note that if *p* is proved to be false, it follows that I do not know *p*, because knowing *p* implies that it is true; but from the fact that I know that *p* is true, I cannot conclude that I know that *p*. Brueckner himself explains the epistemic circularity of Putnam's argument using the principle of disquotational truth (Brueckner 1999: 47). We do not go into this here.

[4] Prior to Brueckner, Tymoczko had also pointed out some simple arguments (Tymoczko 1989). But since Brueckner also tried to overcome the defects of these arguments, and finally presented an account that evaded at least some of the critiques of the simple arguments, we are content to call this category of arguments "Brueckner's simple arguments."

[5] By saying that my justification for believing (B) derives from my knowledge of the semantics of my own language, Brueckner (2016: 22-3) tries to show that SA1 is not circular. He thinks the circularity objection can be made to any argument with the *modus tollens* form. But the circularity objection raises here is not about the form of SA1. It simply says that I cannot know that my word "tree" refers to trees until I know that I am not a brain in a vat.

[6] Externalists have accepted that if one can show that acceptance of semantical externalism leads to the conclusion that an individual has to conduct an empirical investigation of the environment in order to know the content of his mental state, externalism has encountered a serious problem. See, for example, Burge 1988, Falvey and Owens 1994, McLaughlin and Tye 1998.

[7] In response to this question, philosophers are divided into two categories: compatibilists and incompatibilists, depending on whether they consider semantic externalism and the privileged access thesis to be compatible. For a detailed account see Kallestrup 2012.

# Reviews

Larry Krasnoff, Nuria Sánchez Madrid, Paula Satne (eds), *Kant's* Doctrine of Right *in the Twenty-first Century*, (Cardiff: University of Wales Press, 2018).

Incorporation of Kantian ideas into contemporary Anglophone political philosophy over the past fifty years has taken a peculiar path. Hanna Arendt's innovative rereading of the third *Critique* aside, it was "pure moral philosophy" of the *Groundwork* that attracted the most attention, followed by "a philosophical sketch" of *Towards Perpetual Peace* historically important for the scholars of international relations. John Rawls made use of the *Groundwork* when developing his arguments for the priority of the just over the good and modeling the subject in the original position. In the ensuing discussions Kant's *Doctrine of Right* tended to play a secondary role. One of the side effects of this – undesirable from the Kantian perspective – was the moralization of politics and naturalization of morals. Yet it was the *Doctrine of Right* that contained Kant's most systematic and detailed account of what he took to be political philosophy proper, preoccupied with establishing and maintaining juridical condition aiming to provide an equal share of liberty to each member. The volume under review capitalizes on some earlier efforts to bring Kant's *Doctrine of Right* to the wider philosophical attention it deserves.

One of the things that make the book stand out is that it is a truly cosmopolitan affair, bringing together voices from different parts of the globe. The result is by no means homogenous, as standpoints and opinions vary and sometimes conflict, making for engaging reading and allowing to appreciate the plurality of contemporary Kantianism(s). Yet the subject matter is one, and the architectonics of the volume are well designed in following Kant's own structure of the *Doctrine*.

Several themes run through the book. One is social contract. Macarena Marey opens the volume with 'The Originality of Kant's Social Contract Theory' providing a general theoretical account of Kant's social contract theory. This account brings together "pure" and prudential or instrumental arguments Kant has for exiting the juridical state of nature and examines them against several contemporary theories of social contract. The examination reveals important differences between Kant's views and both contemporary contractualism and contractarianism. The most fundamental difference is Kant's maintaining, "that a solid theory of the state could only be based upon a purely juridical or political starting point", and not on any prudential considerations (p. 11). The chapter provides a classification of contemporary theories, attempting to ground "naturalness" of social contract in certain human features, and closely examines one by David Gauthier. The author's conclusion is that Kant's social contract theory offers a fruitful alternative to contemporary theories: "Kant's social contract argument is tailored to give the political autonomy of the united will a strong theoretical basis without assuming a robust prepolitical starting point" (p. 24).

The topic of social contract remains prominent in 'Private Property and the Possibility of Consent: Kant and Social Contract Theory' by Alice Pinheiro Walla. Arguing from Kant's *Doctrine of Right*, the author examines a specific yet important aspect of social contract theory relevant to property right. Central to the discussion is the relationship between Kant's authorization of ownership of external objects as the means to avoid "a contradiction of freedom with itself" and his subsequent and seemingly redundant notion of "common possession of the earth" as the source for "legal title" (p. 29). The latter, a rather peculiar idea, has not been sufficiently attended to current Kantian discussions, moreover, it has sometimes been dismissed – a shortcoming the author is aiming to address. Pinheiro Walla proceeds by explaining the peculiarity awarded by Kant to property rights among other private rights established by mere consent of the parties, and by explicating the role of the notions of 'united will of all' and common possession in Kant's constractualism in relation to property.

Continuing the examination of the foundations of Kant's political theory, Eric Boot in 'Judging Rights by Their Duties: A Kantian Perspective on Human Rights' questions the ordinary treatment of Kant as just another liberal theorist of rights. The aim is to show the intangible connection Kantian rights have to duties to help prevent the explosion of rights claims happening in some contemporary discussions. The clarity is sought in Kant's distinction between duties of right and duties of virtue. The author claims that, unlike many current theories, including some claiming Kantian pedigree, in Kant's framework duties have to be established prior to establishing rights, so Kant's theory can properly be called 'duty-centered'. This is evident in the fact that connecting a right to a duty is a proper way to authenticate it (p. 47). The scope of rights is further and severely narrowed by the fact that only duties of right and not duties of virtue can produce corresponding rights. The result of careful (re-)reading of Kant's *Doctrine of Rights* can be sobering of many overblown claims of rights and, on the other hand, prevention of what the author indicates as the problem of 'rights inflation' (p. 63).

The important topic of rights to material well-being or welfare becomes the second cross-cutting theme of the volume. In 'The Proper Task of Kantian Politics: The Relationship between Politics and Happiness' Masataka Oki approaches the two seemingly distant topics to show the positive connection that is attributable to Kant: "we may regard ourselves happy as long as we live in a political system where every juridical claim of individuals for

what is their own would be justly settled through a rationally constructed system of laws" (p. 68). The author argues that Kant's well-known claim that promoting happiness by political means creates the risk of paternalism is compatible with his seeing happiness as the end of politics. An important by-topic of the chapter is perpetual peace, which is understood as the state of completion of the system of laws and so is the final step towards the possibility of happiness. Oki refers to Gerhardt's idea of 'vernünftiges Selbst' and its happiness in active freedom (p. 69), and Rawls's 'moral psychology' of happiness in following one's rational life plan seems another possible ally to this view. However, some ambiguity remains about whether we are to take a 'well-ordered' juridical condition as the sufficient condition of happiness, or, more plausibly, as a merely necessary one.

Welfare becomes the central topic in Nuria Sanchez Madrid's 'Kant on Poverty and Welfare: Social Demands and Juridical Goals in the *Doctrine of Right*'. The aim here is to warn against taking overly direct and unsystematic imports of Kant's concepts to support contemporary claims for redistribution of wealth. With attention to historical context and to text other than *Rechtslehre*, Sanchez Madrid reconstructs Kant's account of the problem of poverty and of the threats that it poses to society and shows the change these notions have undergone since eighteenth century to enter present discussions. The change has been profound: "there is no correspondence between our current notion of social welfare as a basic right and Kant's response to poverty, which he understands as a sort of institutional charity..." (p. 85). Sanchez Madrid examines Kant's reasons for poverty relief: maintenance of society and, perhaps more importantly, of its resistance to foreign threats. Here again the emphasis is placed not on the rights, as Kant doesn't provide for rights to be relieved from poverty, but on the duty of the state or its ruler to prevent the exclusion of the least well-off (p. 89).

The discussion of welfare continues in Larry Krasnoff's 'On the (Supposed) Distinction Between Classical and Welfare Liberalism: Lessons from the *Doctrine of Right*'. This chapter contributes to the important discussion of whether welfare liberalism *is* liberalism. The fundamental difference between the two liberalisms is framed as that "between views that regard freedom and individual rights as prior to political order, and views that regard freedom and individual rights as only possible through a political order" (p. 106). One of the achievements of the *Doctrine of Right* is that it shows the superficiality of this contradiction by offering another approach. The approach consists in differentiating between freedom as inner moral autonomy that belongs to all beings endowed with reason and as outer liberty to act or pursue ends. In Locke this latter liberty is granted by nature in form of rights, whereas in Kant it can be rightful only by convention. A set of recent arguments is used to illustrate, how Kantian approach produces plausible explanations to cases that were not around at Kant's time, like maintenance and regulation of public roads and mandatory health insurance schemes, avoiding the need to invoke the supposed difference between classical and welfare liberalism. It remains to be examined if the author's solution based on differentiating between inner and outer freedom requires a thoroughgoing shift in metaphysical paradigms from naturalist monism implied by contemporary Lockean liberal theories to a sort of Kantian transcendentalism – a shift that not all parties would accept.

Wendy Brockie's 'Resistance and Reform in Kant's *Doctrine of Right*' also contrasts Locke to Kant, this time regarding Kant's notorious prohibition of sedition and rebellion. This prohibition has caused numerous attempts to reinterpret or correct Kant based on other seemingly more fundamental provisions of his moral and (sometimes also) political philosophy. The author assesses these attempts and tests their applicability to some recent events like the 'Arab Spring'. Analyzing Kant's position, Brockie emphasizes the deeply skeptical foundations regarding the "impure" side of human nature, like dissemblance and "the corrosive effects of humans living together in communities" (p. 131), that underlie Kant's prohibition of active resistance together with more familiar arguments concerning the perfect duty to obey any law. The conclusion is that Kant, having discussed the complexity of factors relevant to the problem of active resistance, seems to live us without a justification for confronting unjust rule. At the same time his provisions for freedom of speech that he sees necessary for evolutionary development are vulnerable and can be ignored by unjust authority. This careful but limited conclusion warrants the continuation of the discussion.

The discussion continues in the next chapter by Alyssa R. Bernstein, titled 'Civil Disobedience: Towards a New Kantian Conception'. Here the reading that attributes to Kant an unconditional prohibition of disobedience is contrasted with what the author defends as a more plausible (and broadly Rawlsian) view that does provide for active resistance to tyrannical power. A brief but thorough restatement of Kant's arguments aims to show that they are relevant provided the government is legitimate and does not destroy law (p. 141). Subsequent survey of a host of recent interpretations leads to the conclusion that none of them is sufficient to justify the obedience to a tyrannical authority. Therefore, it is "permissible for individuals in a condition of barbarianism or a state of nature to use force" (p. 142), and "individuals are authorized to judge for themselves whether they are living under the rule of law or instead in a condition of barbarianism" (p. 141). The participation by Dr. Martin Luther King Jr. in demonstrations in 1963 is then taken as an example to illustrate how Kantian arguments provide for civil disobedience. The thesis of this chapter might benefit from a discussion of whether we, when judging ourselves to be in a state of nature, are to take into account the opinions of those people who happen to be under the same authority and might find it sufficiently legitimate.

Chapter nine has the title 'Kantian Insights on the Moral Personality of the State' and is contributed by Milla Emilia Vaha. The author summarizes and criticizes what she calls "liberal exclusionism", which purports to use Kant's idea of a moral personality of a state to deny a statehood to certain political sovereign formations that fail to meet certain moral criteria. Notable examples of liberal exclusionism are provided by Michael Doyle's "liberal peace theory", by Fernando Tesòn's claim that human rights protection is the criterion of state legitimacy, and by John Rawls' idea that only liberal and "de-

cent" peoples can be ascribed international agency and personhood. "All three authors seem to claim, relying on Kant, that the liberal states are *morally superior* and should thus enjoy rights that non-liberal states do not enjoy." (p. 167) While the "liberal exclusionary" position might have other theoretical pillars, it seems problematic from a Kantian perspective. To show this, the author undertakes the analysis of interpretations of Kant's moral personality of the state and finds support for a more generous understanding of moral states that expands the status to all but failed states, thus including the "despotic" states. The basis for this wider reading is chiefly provided by Kant's explicit principle of non-intervention and his idea of the importance of development towards republican constitution as a sign of a state's moral personality.

Sorin Baiasu continues the discussion of Kantian international relations with 'Kant's Guarantee for Perpetual Peace: A Reinterpretation and Defence'. Kant uses the notion of "guarantee" in *RL* and *ZeF* to denote the nature's will to override human reluctance to pursue the highest political good of perpetual peace. Baiasu points out four requirements that this guarantee has to meet in order to be coherent: it cannot transcend the cognitive limits set in the first *Critique*, it cannot be equated to a postulate of practical reason, it cannot depend on individual's decisions, and "it should have an epistemic status that provides some motivating force in addition to the normative force of the associated moral duty." (p. 193) Baiasu analyzes some of the recent attempts to explain Kant's guarantee and their not meeting all of the four requirements. The requirements are met, he believes, if we treat the problem of guarantee along the lines of Kant's argument in 'On Having Opinions, Knowing and Believing' of the first *Critique* (A820/B848-A831/B859). This allows the guarantee to act as an object of doctrinal belief. This solution rests on a strong notion of purposive nature that has an independent "will" in addition to "force" – a notion that introduces yet another set of theoretical complications.

In Kant's writings the topic of punishment is far more prominent than that of forgiveness. Paula Satne's 'Forgiveness and Punishment in Kant's Moral System' aims at examining the Kantian balance between the two concepts. Contrary to the view that Kant's system of morals provides no place for forgiveness, Satne offers a more charitable reading that finds forgiveness to be a "wide duty of virtue which is conditional on repentance." (p. 202) It is possible to speak of a *maxim of forgiveness* as a duty to have a forgiving character, although not of a perfect duty to forgive specific offenders. Satne's account relies on Kant's theory of rational agency from the *Groundwork*, his theory of evil from the *Religion*, and his moral metaphysics, thus providing several arguments to support the claim. The author offers to expand Kant's account of the states that have to be overcome by forgiveness to include not only hatred and vindictiveness, but also more general emotions like anger and resentment. This enhancement, while seemingly not contradicting anything Kant wrote, offers a more detailed treatment of forgiveness.

Kant is usually not the first authority when it comes to issues of marriage. In 'A Universal Estate: On Kant and Marriage Equality' Jordan Pascoe assesses Kant position in historic debates related to marriage and its treatment in Prussian Civil Code of 1794. Kant's opinion found expression in his lectures on moral philosophy and then found its way into the *Doctrine of Right*. Kant is seen as occupying middle ground in the debate between those who, like Fichte, awarded marriage a natural status and those who, like Hippel, treated it as a primarily juridical institution. Central to Kant's discussion is the idea of possession of another individual as a natural object for purposes of happiness and finding a proper juridical form to cultivate this relation. Pascoe relates the 18-century debates to current discussions of same-sex marriage to show that Kant's position predates arguments by those seeking to extend the right to marry. Correspondingly, those current contenders who stand for reforming the very notion of marriage would find little support in Kant's thought.

*Kant's* Doctrine of Right *in the Twenty-first Century* seems most inviting to two kinds of audiences. To those immersed in current social and political issues it offers a path to one of the most technical systems of thought, in which support for – or criticism of – one's position can reach metaphysical depths beyond ordinary arguments. And to those devoting most of their attention to the history of philosophy, and particularly to Kant's works, it might serve as a link or a clue to hot problems of the day. The book is successful in joining the two intellectual enterprises and setting them in motion towards each other.

Vadim Chaly
*Immanuel Kant Baltic Federal University*
DOI: 10.5281/zenodo.3903554

Donato Verardi, *La scienza e i segreti della natura a Napoli nel Rinascimento. La magia naturale di Giovan Battista Della Porta*, (Firenze: Firenze University Press, 2018).

A partire dagli anni ottanta del secolo scorso, grazie agli innovativi contributi di autori come Charles B. Schmidt e Charles Lohr[1], lo studio dell'aristotelismo rinascimentale ha goduto di grande fortuna e rinnovate attenzioni. Nel loro complesso, le nuove ricerche hanno reso il pensiero aristotelico del XVI secolo un campo di studio indispensabile e fecondo, sia al fine di approfondire il contesto in cui i tradizionali domini della conoscenza stavano mutando i loro confini, sia per valutare con il dovuto rigore i legami con saperi occulti come la magia naturale e l'astrologia, ben saldi nel periodo nei quali maturò la "rivoluzione scientifica". Il presente volume, dedicato alla proposta del celebre mago naturale Giovan Battista Della Porta (1535-1615), si inscrive, per taluni aspetti, in questo filone di indagine, costituendo un apporto significativo alla comprensione della complessa fisionomia dell'aristotelismo all'alba dell'età moderna. Il volume si divide in tre parti. Nella prima, l'A. discute la più accreditata storiografia relativa al contributo di Della Porta ai dibattiti scientifici del XVI secolo, nella seconda prende in esame il dibattito scientifico sui "segreti della natura" così come sviluppatosi a Napoli nel Rinascimento, mentre, nella terza, analizza la magia naturale di Della Porta con particolare riguardo al problema dell' "occulto" e del "segreto". Gli ultimi capitoli sono dedicati allo spinoso problema del

rapporto magia-stregoneria nei difficili decenni della Controriforma. Una delle acquisizioni più interessanti del libro è senz'altro la scoperta di una versione dell'aristotelismo rinascimentale diffusa in area partenopea e dotata di una propria specificità, dove l'indagine della natura si coniuga a studi magici ed astrologici ricchi, in alcuni casi, di apporti neoplatonici ed ermetici. Peraltro, dalle minuziose analisi dell'A., emerge come l'aristo-telismo si sviluppasse a Napoli in una significativa varietà di prospettive, talvolta divergenti proprio in relazione al problema inerente alla magia naturale. È ormai assodato che, pur nella mutata cornice del pensiero della prima età moderna, il paradigma aristotelico potesse giustificare teorie e pratiche attinenti alla magia, all'aritmologia ed alla cabbala. Ricerche condotte pochi anni orsono hanno potuto verificare questa tesi; è il caso del lavoro di Bernd Roling intorno alla figura ed all'opera di Paolo Ricci[2]. L'analisi di Verardi sulla magia naturale di Della Porta e sul *milieu* partenopeo in cui essa si formò rappresenta un'ulteriore dimostrazione di questa compatibilità. Mediante la descrizione dell'attività e delle opere dei principali filosofi naturali del Rinascimento partenopeo, nel resoconto acquistano la dovuta dignità profili di umanisti e studiosi finora poco o per nulla noti, ma che ebbero, all'epoca, un ruolo assai significativo. Tra di essi si impone Francesco Storella (1529-1575), professore di logica a Napoli negli anni in cui Della Porta scrive e pubblica il suo capolavoro, la *Magia naturalis* del 1558. Mettendo a frutto le acquisizioni dell'occamismo e della riforma rinascimentale della dialettica promossa da Lorenzo Valla (1405 o 1407-1457), l'aristotelico Storella (di origini salentine, ma formatosi a Padova) fornisce le basi concettuali per quella riforma della scienza che avrebbe permesso a Della Porta di fondare la propria indagine razionale sulle proprietà 'occulte', e pur naturali, dei segreti della natura (Cfr. pp. 31-45). L'ampia ricognizione del rinascimento culturale della Napoli cinquecentesca conduce l'A. a esaminare anche il dibattito inerente all'astrologia. In tale contesto, dopo la decisiva critica delle *Disputationes adversus astrologiam divinatricem* di Giovanni Pico della Mirandola (1463 –1494), la scienza degli astri appare contraddistinta da un significativo ritorno alle teorie del *princeps astrologorum* Claudio Tolomeo. Negli studi astrologici del rinascimento partenopeo, si impongono le opere di Giovanni Pontano (1429-1503) e di Giovanni Abioso da Bagnolo (fine XV sec.-inizi del XVI sec.), grandi interpreti, insieme a Girolamo Cardano (1501-1576), del lascito del 'vero' Tolomeo. Grazie a una ricerca di prima mano sui testi e le loro fonti, l'A. mostra come le proposte di tali filosofi-astrologi partecipino alla complessa delineazione del concetto di *influxus*, centrale nella successiva razionalizzazione dei segreti della natura proposta da Della Porta. Capacità ed effetti delle influenze celesti vengono esaminati scrupolosamente, secondo un approccio che risente dei conseguimenti del sapere ottico e matematico, sovente contrapposto da questi autori alle spiegazioni 'superstiziose' di negromanti e demonologi. Sempre per quanto concerne il versante astrologico, una figura di grande interesse indagata nel volume è quella dell'astrologo Matteo Tafuri (1492-1584), grande frequentatore del circolo intellettuale dei fratelli Della Porta, nonché profondo conoscitore dell'astrologia tolemaica. Tafuri, pur escludendo una visione fatalista dell'astro-logia, descrive un cosmo in cui gli astri si identificano con i *daemones*, finendo addirittura per chiamare in causa nei suoi pronostici l'allora proibitissima *Clavicula minor* di Salomone. Si tratta di un elemento che complica di molto lo scenario in cui viene a maturare la proposta peripatetica di Della Porta e grazie al quale l'A. può mostrare le differenti sensibilità che pure poterono influire, a più livelli, sul filosofo campano (cfr. pp. 56-59). L'immagine che l'A. fornisce di Della Porta è certamente quella di un pensatore influenzato dell'aristotelismo partenopeo, ma non per questo appiattita su di esso. L'A. mostra come Della Porta si apra alle istanze dell'albertismo rinascimentale, alla luce del quale egli rilegge la lezione di quello che sembra essere il maestro occulto di Della Porta, Cornelio Agrippa di Nettesheim (1486-1535). Da Agrippa, infatti, Della Porta sembra mutuare una serie di elementi filosoficamente pregnante, ma che egli rilegge alla luce di una sensibilità "albertina" essenzialmente estranea agli interessi demonici e cerimoniali che avevano contraddistinto, invece, la proposta del mago di Nettesheim (cfr. pp. 22-27). Al di là di facili schematismi che, da un lato, relegano Della Porta al ruolo di scialbo ripetitore di Agrippa e, dall'altro, lo confinano in quello di mero scopritore dei principi astrologici adatti a regolare le virtù occulte, emerge il ritratto di un profondo investigatore dei *secreta* naturali, attento per lo più al carattere "empirico" delle sue ricerche, tanto da privilegiare l'indagine degli effetti. Ciò che interessa Della Porta è *in primis* «il piano degli effetti della sostanza naturale, constatati nel corso dell'indagine scientifica e giustificati, in seconda battuta, tramite la ricerca razionale delle loro "cause prossime" e "sufficienti"» (p. 150). In equilibrio tra la giustificazione logica dei *secreta* e loro comprensione empirica, il mago naturale conduce le sue indagini procedendo dall'espe-rienza. La magia naturale rappresenta, perciò, un dominio di indagine nel quale ogni elemento *extra naturam* ed ogni *secretum* è sottoposto ad un radicale processo di naturalizzazione. Ciò vale anche per i segreti più "scabrosi", notoriamente ricondotti nell'alveo della stregoneria (cfr. pp. 123-138). A delineare ancora più chiaramente metodi ed oggetti della filosofia naturale di Della Porta interviene l'originale classificazione proposta dall'A., il quale dimostra come, per il filosofo campano, i *secreta* siano essenzialmente di tre tipi: 1) i «segreti» le cui virtù occulte sono regolate dalle leggi astrologiche che governano la «natura»: come, per esempio, l'attrazione del magnete sul ferro; 2) i «segreti» le cui virtù sono da ricondurre all'abilità «tecnica» dell'uomo nel manipolare gli elementi naturali. È questo il caso della testa parlante di Alberto Magno, le cui virtù sono solo impropriamente definibili «occulte», in quanto perfettamente spiegabili tramite un meccanismo che si serve dei quattro elementi (nel caso specifico, dell'aria); 3) i «segreti» nei quali la produzione delle virtù occulte necessita della contributo della «natura», con le sue leggi astrologiche, e della «tecnica» umana. (cfr. p. 149). Tale tripartizione è indubbiamente un contributo assai innovativo al dibattito relativo ai segreti della natura nella prima età moderna, in quanto evidenzia come l'attenzione di Della Porta non si limiti ai *secreta* spiegabili tramite gli influssi celesti. Il secondo genere di *secreta* nulla, infatti, ha a che fare con le virtù occulte di stelle e pianeti, e costituisce invece un significativo esito dello "sperimentalismo" di cui Della Porta si fa promotore nella sua Accademia dei

Segreti (cfr. p. 14). Il fenomeno della testa parlante di Alberto Magno, caso esemplare dei «segreti della tecnica», viene ricondotto, per l'appunto, a meccanismi del tutto naturali, messi in atto dall'abilità del mago naturale. Così il mago dellaportiano, infaticabile scrutatore della natura per mezzo dell'espe-rienza, si rivela per molti aspetti simile alla nuova figura del filosofo empirico che si affermerà durante il XVII secolo, differenziandosene, tuttavia, per un atteggiamento ancipite: se, infatti, da un lato le sue ricerche si volgono con sicura convinzione al piano degli effetti, dall'altro, contemplano ancora il sapere qualitativo, comune all'aristotelismo e all'astrologia.

[1] Cfr. Charles B. Schmitt, *Aristotle and the Renaissance*, Cambridge (Mass.) & London, Harvard University Press, 1983; Charles H. Lohr, *The sixteenth-century transformation of the aristotelian natural philosophy*, in *Aristotelismus und Renaissance. In memoriam Charles B. Schmitt*, hrsg. von Eckhard Kessler, Charles H. Lohr und Walter Sparn, Wiesbaden, O. Harrassowitz, 1988, pp. 89-99.
[2] Cfr. Bernd Roling, *Aristotelische Naturphilosophie und christliche Kabbalah im Werk des Paulus Ritius*, Tübingen, Niemeyer, 2007.

Marco Ghione
*École Pratique des Hautes Études, Paris-Università degli Studi di Genova*
DOI: 10.5281/zenodo.3903559

Marcus Willaschek, *Kant on the Sources of Metaphysics. The Dialectic of Pure Reason* (Cambridge: Cambridge University Press, 2018).

This book provides an extensive and insightful analysis of the Dialectic of Kant's first *Critique*. Willaschek's aim is to focus on the constructive side of the Transcendental Dialectic and, besides Kant's critique of speculative metaphysics, highlight the relevance of the Rational Sources Account (RSA), that is Kant's discovery that the sources of metaphysical thinking lie in reason itself. According to Willaschek, reason's metaphysical features follow from three main issues: 1) the discursive character of human thinking – for reason proceeds from elements to synthetic cognitive claims; 2) the iterative character of rational explanation – where every inquiry concerning reason-giving can be, in principle, always reiterated (if A is because of B, the question *why* B is always legitimate); 3) the rational need for completeness – for reason's satisfaction relies on answers which do not raise further questions.

Concerning its structure, the book is divided into two main parts. While in the first (chapters 1-5) Willaschek follows the path which brings reason towards metaphysical speculation, the second part mainly deals with the constitutive steps of the Transcendental Dialectic: the (inferential) derivation of the transcendental ideas, the paralogisms and the antinomies, the ideal of pure reason. Finally, in chapter 9 Willaschek argues for a possible rejection of Transcendental Realism (TR) – which is responsible for transcendental illusion – which does not compel us to accept Kant's Transcendental Idealism (TI). Concerning the RSA, the reader is strongly recommended to carefully look at section 0.3 in the Introduction, for there Willaschek points out very clearly that the RSA consists in a single complex argument which is articulated into four levels. The first is the transition from the Logical Maxim (LM) to the Supreme Principle (SP). While the LM "requires us to find a condition for each conditioned cognition" (p. 6), the SP states that if something conditioned is given, then the unconditioned complete series of conditions is given as well. The concept of the unconditioned is, at the second level, described in terms of the system of the transcendental ideas – where these concepts follow from natural and necessary inferences of reason. The focus of the third level is on reason's dialectical inferences, as originating from human reason itself and, therefore, being compelling to every reader of Kant's first *Critique*. The fourth and final level deals with two alternative uses of transcendental principles and ideas. While their regulative use allows for the search of unity among all empirical laws, their being taken for true representations of objects, that is constitutively, unavoidably brings reason towards transcendental illusion. This latter level seems to have a more specific methodological status, for the transition from the logical to the real use of reason and, in a complementary way, from the regulative to the constitutive use of reason's ideas and principles, represents the distinctive mark of reason's natural tendency to metaphysical speculation. One last remark: Willaschek stresses the relevance of the constructive side of the Transcendental Dialectic – namely the RSA – in order to counterbalance and mitigate the reading according to which the Dialectic would merely consists in Kant's "demolition of traditional metaphysics" (p. 9). Such an approach is very welcomed; its outcomes are undoubtedly stimulating for every Kant scholar and fruitful for contemporary epistemologists.

One of the first most relevant argumentative steps of the book is introduced at p. 46. There Willaschek – after having shown that even the ordinary employment of reason leads to metaphysical speculation – distinguishes between the LM and the SP. Willaschek refers to *KrV*, A307-8/B364 and assigns the LM to the logical use of reason (as it only deals with cognitions), while the SP belongs to the real use of reason (as dealing with objects). Some questions arise. As it is clear from Kant's statements, it is more appropriate to consider the LM as becoming the SP, and not to firmly distinguish the former from the latter. It is true that Willaschek, several pages later, clarifies that he does not mean to argue for two properly different principles but, rather, that "the transition passage must be understood on the model of sentences such as 'A bill can become a law only by an act of legislation' , or, more generally, '*X* can become *Y* only by way of *Z*,' where *X*'s becoming *Y* just consists in *Z*' taking place" (p. 122). The statement is clear: the LM becomes a principle of pure reason (A308/B365) by assuming the SP. This given, one could still wonder that another interpretative option remains available. While Willaschek correctly claims that there would be two, not three, principles at stake here, one could argue for one principle which takes different shapes. First, why may it not be the case that the LM becomes the SP in virtue of the fact that, otherwise, the unconditioned reason is in search of would not be really unconditioned? As Willaschek puts it, the LM deals with cognitions, while the SP deals with objects. In moving from LM to SP, Willaschek thus states

that LM presupposes SP; this is correct. However, it may still be possible that what the LM prescribes – namely the search for the unconditioned totality of the series of conditioned cognitions – cannot be fulfilled unless we move from cognitions to objects. There would be no truly meaningful logical employment of the logical maxim if we would not assume that the unconditioned is given, that is if we do not assume that the cognitions at stake are cognitions of objects. For sure, Willaschek's analysis is far from being inconsistent. What may be suggested is that the LM and the SP are so closely linked that they are one and the same thing. Instead of arguing for a maxim to become a principle of pure reason by means of another (supreme) principle, it may be the case that Kant means that the LM can play the role of a principle of pure reason insofar as the logical regression of conditions is also taken for a real one. In these terms, the SP would just represent a different use of the LM and there would be no need to say that the LM becomes a principle of reason via the SP – for the latter implies a transition from mere (logical) thinking to reality. In short, when dealing with the unconditioned, reason cannot be satisfied with a merely logical unconditioned – intended as the totality of conditioned cognitions. Reason's demand is higher: when taking into account the unconditioned, reason has to overcome the boundaries of logic – otherwise reason would be dealing with an unconditioned which would be limited, thus conditioned, by its not-being-real. Whatever may be the most appropriate reading, the main outcome remains unchanged: the dialectical inferences come from the transition from the logical to the transcendental (real) use of reason.

This transition is the key issue of chapters 4-5, which represent the core of Willaschek's interpretation. A closer look at section 4.2.2 leads the reader to important clarifications. In particular, Willaschek distinguishes between 1) the regulative and the constitutive use of reason and its principles, 2) the logical and the real (transcendental) use of reason, 3) the subjective and objective validity of the principles at stake. Moreover, Willaschek warns against assuming a symmetrical correspondence between the regulative/constitutive and the logical/real (transcendental) distinction. His aim is to propose a different reading, according to which the hypothetical use of reason – though non-constitutive – allows for the search for the unity of cognitions. Additionally, this use does not imply that reason's principles are false, but only that there is no warrant about them to be true. It follows that regulatively-used principles are assumed only problematically, for their prescriptive force is limited to reason's demand to "investigate the hypotheses that follow from the principle" (p. 115). Willaschek draws a further conclusion. By distinguishing between logical prescriptive principles and regulatively used transcendental (descriptive) principles, we may employ the latter as devices for assuming hypotheses about objects without using them constitutively – namely without applying them in order to determine objects. This reading coherently clarifies how reason's principles may refer to objects in two different ways: either regulatively, as heuristic devices for approximating systematic unity, or constitutively, thereby taking them to be objectively valid of nature itself. This picture is undoubtedly coherent. The only eventually missing argument here would be dealing with the following question: what are these principles meant to be constitutive of? It seems Willaschek takes for granted that in these pages of the Dialectic *constitutive* means *constitutive of experience*. In this case there would be obviously nothing to say, for reason's principles and ideas cannot play this role. However, what if these ideas and principles would be meant to be constitutive of the systematic unity of reason itself? Their prescriptive – normative – role would be stronger and, most importantly, there would be a less sharp distinction between constitutive and regulative. These two terms may not be alternative as it seems, for even the regulative use of reason's principles would be somehow constitutive of reason's systematic unity.

Moving to part II, two main points need to be discussed adequately. The first concerns the assumption of Transcendental Realism (TR) as responsible for transcendental illusion. The second consists in the rejection of TR without any complementary assumption of Transcendental Idealism (TI). To be precise, the topic of TR is first introduced in section 5.2.2, where Willaschek states that TR "is the key to understanding transcendental illusion in general, including the transcendental illusion involved in the transition from the regulative to the constitutive Supreme Principle" (p. 139). As already stated, the constitutive use of SP demands it to be valid of objects – far beyond the assumption of hypotheses. This real use of reason is responsible for all reason's dialectical inferences, also including Kant's arguments for the ideal of reason and God's existence. Chapters 7-8 – especially sections 7.2, 7.3.2, 7.4.2 – aim to show clearly how TR is at work in every step of the Transcendental Dialectic: from the transcendental ideas to the paralogisms, the antinomies and reason's ideal. If transcendental illusion follows from reason's constitutive use, TR – quite complementarily – takes 1) "rational principles to be constitutive of nature" (p. 165) and 2) this constitutive role to be a metaphysical insight and not a mere subjective projection. In a few words, according to TR that rational principles are more than merely rational, for nature's structure is assumed as completely corresponding to the principles at stake. The non-legitimacy of this use of reason is first evident in the paralogisms (section 7.2.2). At once, Willaschek reconstructs Kant's arguments and shows that the paralogisms entail a transition from the logical to the real use of reason. More specifically, the problem is that we "take the necessary conditions under which we represent things to be conditions of those represented objects" (p. 198). This makes evident that Willaschek does not limit his work to exegesis, for the consistency of his arguments follows from a critical re-definition of Kant's claims. In particular, concerning TR, Willaschek proposes the variation $TR_{rep}$:

> If, to be represented at all (by finite beings like us), some object *o* must be represented as being *F*, then *o* is *F*.

Going ahead with the antinomies and the ideal of reason, the reader can easily recognize that $TR_{rep}$ is always at work in reason's dialectical inferences and how it leads to transcendental illusion. This said, it becomes thus relevant to reject transcendental realism in order to both avoid

transcendental illusion and – at the same time – allow for the hypothetical use of reason. Quite surprisingly, Willaschek adds to this rejection a complementary argument about the non-necessity to assume Transcendental Idealism (TI). As it is very well known, Kant's solution to the paralogisms and the antinomies precisely relies on TI. However, Willaschek thinks that the rejection of TR does not imply any assumption of TI. This argument is developed in chapter 9. The key *demonstrandum* is that TI and TR are not contradictory, so that it is possible to reject the latter without assuming the former. According to Willaschek, TI implies that the properties of cognizable objects (appearances) depend on the possibility of being represented by finite rational beings. Though this is true, it would have been of help to have some reference to Kant's texts, since Willaschek always takes care of this in the whole book. Willaschek also says that this claim is paradoxical, for our cognitions of these properties are not representation-dependent. In addition to this, since many contemporary philosophers outside Kant scholarship do no assume TI, it would be better not to rely on TI in order to leave Kant's critique of metaphysics more philosophically attractive for contemporary approaches (pp. 248-249). This raises some perplexity. How could one think to give up arguing for one of the main achievement of Kant's critical philosophy – according to Kant himself – in order to, allegedly, leave another relevant achievement attractive for contemporary inquirers? One could wonder whether this attractiveness is too good a reason. Willaschek's argument relies on the following steps: 1) TI implies the falsity of TR for TI denies that TR holds for things-in-themselves (TR is thus restricted to appearances, that is empirical objects); 2) TI implies the falsity of TR for TI states that empirical objects do not coincide with the objects of a merely rational order (noumena in the positive sense); 3) neither the first nor the second argument for TR's falsity proves the truth of TI. Thus TI and TR are non-contradictory, for TI is merely the contrary of TR. It follows that "we can deny that the necessary conditions of representing some object in all cases are necessarily conditions of that object [TR] without accepting that empirical objects are mere appearances [TI]" (pp. 250-251). In these terms, Willaschek argument seems convincing. However, it may be the case that this perspective on TI is too simplified. The role of the pure forms of sensibility in allowing the receptivity of phenomena may require some additional discussion. Though it is true that TI holds for representation-dependent appearances, it is also true that the possibility of experience precisely relies on the fact that our representations can be valid of outer objects. Besides the perplexities of contemporary epistemologists and metaphysicians, Kant may still be right in claiming that it is impossible for us to represent anything outside space and time. Accordingly, by abstracting from the conditions which allow for an appearances to be represented *by us*, there would remain nothing but a mere (some)-thing *in itself*.

Willaschek's reconstruction of the RSA is not only complete and coherent, but also open to further developments into the practical sphere. Kant's practical metaphysics (Postscript, pp. 270-275) – with its postulates, the moral law, the highest good – represents Kant's path towards a non-dialectical and non-illusionary employment of purely rational principles. This perspective, though only sketched, is a significant legacy of *Kant on the Sources of Metaphysics* – together with the systematical reconstruction of Kant's RSA. Kant scholars may hardly avoid dealing with a book which has the merit of discussing a too often overlooked part of Kant's first *Critique*, presenting clear and insightful arguments for original views and accounting for reason's metaphysical drive.

Luigi Filieri
*Johannes Gutenberg-Universität Mainz*
DOI: 10.5281/zenodo.3903561

Diego S. Garrocho, *Sobre la nostalgia* (Madrid: Alianza Editorial, Madrid, 2019).

*Sobre la nostalgia* è l'ultimo saggio pubblicato da Diego S. Garrocho, professore di Etica e Filosofia Politica presso l'Universitá Autonoma di Madrid. Pubblicato in Alianza Editorial nel 2019, la scrittura di Garrocho è in grado di coniugare la filosofia con le più diverse espressioni della cultura umanistica, dalla storia alla politica, dalla mitologia al cinema, dall'architettura alla sociologia. Il testo ruota intorno all'analisi della nostalgia, descritta in relazione alla dialettica tra memoria e oblio, dolore e piacere, assoluto e storia, fino a tracciarne i connotati contemporanei in chiave politica. La riflessione sulla nostalgia prende piede da un'interpretazione peculiare della *damnatio memoriae* - il sottotitolo al saggio -, che Garrocho non intende solo come provvedimento punitivo nei confronti della memoria di qualcosa (di una persona, di un evento, di un luogo), ma come dolore causato dalla facoltà della memoria stessa. La memoria genera dolore proprio in virtú del ricordo, a prescindere dal contenuto: non è solo il ricordo di un evento spiacevole che causa dolore (il trauma), ma, anzi e soprattutto, il ricordo di un evento piacevole, proprio perché si colloca in una dimensione che risulta inevitabilmente inaccessibile, quella del passato. Se la memoria si oppone all'oblio, e varie sono state le tecniche proposte dalla tradizione occidentale per ricordare, non sembra esistere una tecnica per dimenticare realmente efficace. Non esiste l'imposizione o l'educazione all'oblio ed è proprio questa impossibilità di dimenticare che rende in qualche modo la memoria una forma di condanna ineludibile. La configurazione della memoria come generatrice di dolore scaturisce da una particolare concezione del tempo che appartiene all'uomo, che Garrocho mutua dalla distinzione aristotelica tra memoria e reminiscenza: mentre la prima facoltà consente di trattenere nel presente un evento del passato (capacità di cui altri esseri viventi sono dotati in diversi gradi), la reminiscenza è quella capacità di sentire come proprio, da sempre, la conoscenza di un determinato evento, quasi come se fosse una forma, per utilizzare un lessico piú contemporaneo e forse azzardato, di far emergere dall'inconscio qualcosa di cui non si aveva consapevolezza. In questo senso, si parla dell'associazione platonica tra conoscenza e ricordo, e proprio in quell'oscura mancanza, in quello stato di oblio inconsapevole in cui verte la condizione umana, si sviluppa l'esperienza nostalgica: la

concezione umana di tempo, che riconosce una distinzione tra passato, presente e futuro, ma che in qualche modo avverte una continuità portatrice di un'assenza, genera timore e speranza, curiosità e coscienza della perdita. La propensione dell'uomo sembra essere quella di avvertire una mancanza che vuole colmare: da una parte questa mancanza spinge alla conoscenza, dall'altra gli causa un dolore primordiale, atavico, nostalgico: in questo senso rimpiange, perché viene al mondo già perso, come la grande tradizione occidentale ha avuto modo di raccontare nelle sue sfumature religiose e filosofiche, dalla Genesi all'esistenzialismo. Nella riflessione aristotelica sulla reminiscenza, risulta fondamentale l'intervento non solo sull'anima, ma anche sul corpo. La questione del corpo è centrale, perché rappresenta la stessa base materiale del castigo inflitto dalla *damnatio memoriae*, che prevedeva gesti concreti come il ritiro del nome della persona da dimenticare o la decapitazione dei busti delle statue. In effetti, sembrerebbe che l'unico modo per imporre l'oblio sia quello di avere una superficie di intervento su cui far gravare il peso della condanna, una traccia materiale che però impossibilita un oblio totale, in quanto rimane una testimonianza. Nella stessa natura umana si conserva una traccia indimenticabile, anche se non si sa chiaramente di cosa, anche se non si è probabilmente mai conosciuta: in questo senso, il dolore del ricordo è un dolore decisamente umano, che quindi, per definizione di Garrocho, è un «animale che rimpiange». In spagnolo, il verbo che Garrocho utilizza è «añorar» la cui traduzione italiana «rimpiangere» va spogliata della connotazione di pentimento: il rimpianto a cui si riferisce Garrocho fa leva sulla particolare condizione di chi sa che qualcosa appartiene al passato e che non puó piú tornare. Il tentativo di cura della condizione esistenziale di rimpianto può essere trovata proprio nella funzione terapeutica dell'oblio, di cui la poetica si fa portavoce: attraverso la rielaborazione di un ricordo, di un gesto, di un esempio, si ricorda ciò che va ricordato, dimenticandone il dolore: «la poetica, nel senso piú classico, non è altro che una *tanatopraxia*: un esercizio di creatività al servizio di ciò che è già morto. Un intervento che altera e modifica un corpo, un oggetto o un'esperienza, per ingannarla e farne la sua migliore memoria».

In qualche modo, sembra che la natura dell'essere umano lo porti a provare continuamente un senso di assenza. La particolare sensazione di mancanza di qualcosa di perso assume i connotati specifici della nostalgia nel momento in cui viene inventato il termine alla fine del XVII secolo, precisamente nel 1688, quando Johannes Hofer 22 giugno 1688 coniò la parola per nominare una stato patologico nella sua *Dissertatio medica de nostalgia odes Heimweh*: in questo periodo nasce il dolore causato dal ricordo della terra natale nei soldati costretti a stare lontani dalla propria patria, dalla propria casa. La nostalgia nasce quindi come mal di patria, come sofferenza riferita allo spazio. Garrocho fa notare che, a partire dal XIX secolo, la nostalgia assume connotati sempre più metaforici e inizia a discostarsi da un'interpretazione puramente patologica. L'oggetto che causa la nostalgia non è più un luogo fisico, ma si trasforma nella mancanza di ciò che davvero non puó tornare indietro, il tempo. In effetti, anche riprendendo l'originale interpretazione, non è semplicemente il luogo fisico che provoca dolore, ma è il complesso di condizioni, è quel presente ormai passato, il tempo e come ci si sentiva a quel tempo: l'ambizione di tornare a una vita ormai passata è il motore del lungo pellegrinaggio di Ulisse, che al suo ritorno deve affrontare l'inevitabile scorrere del tempo, sebbene il luogo fosse rimasto lo stesso. La nostalgia si nobilita, riprende in qualche modo e rafforza la sua particolare connessione con la malinconia, esperienza emotiva anch'essa ispirata dal sentimento di mancanza.

Nella sua dimensione politica, la nostalgia sembra essere qualcosa di inevitabile e connaturato allo spirito stesso della Modernità, che nasce con il concetto di progresso: «Non c'è nulla di piú moderno della nostalgia perché non c'è niente di piú antico del futuro». La prospettiva di un futuro migliore si accosta al lamento del presente e all'ispirazione del passato, in una concezione del tempo tipicamente umana: l'utopia diventa, nel caso della nostalgia, un'utopia retrospettiva. La Modernità, nel clima di emancipazione, quindi di rottura, fa riecheggiare quel sentimento primigenio di mancanza, di perdita, di anelo a ritornare a qualcosa di perso. Sebbene originariamente nata come esperienza connessa al luogo, la nostalgia è attualmente viva nella sua dimensione metaforica. Oggi si assiste, secondo la visione di Garrocho, a una ancor piú moderna lettura nostalgica del mondo, in seguito al fallimento della promessa del progresso del XX secolo, in particolare con la caduta del muro di Berlino. Cosa succede quando la promessa viene disattesa? Si perde la speranza e lo sguardo sembra proiettarsi ossessivamente al passato a causa di una paura del futuro: ne sono prova, tra le tante, la recuperazione del passato all'interno della città, che si mobilita per dare spazio al *vintage* e la moda del *selfie*, che si orienta verso una doppia prospettiva: la necessità di immortalare un momento che non ha futuro e di rendere memorabile qualsiasi istante, anche il piú insignificante. Si assiste quindi all'ipertrofia della *damnatio memoriae*: la nostalgia sembra essere diventata una caratteristica generazionale, della cosiddetta generazione dei *millenials*: non è piú, in senso fisico o metaforico, la volontá di far tornare un passato perso Non basta piú recuperare il corpo fisico del passato, ora si vuole ricordare come si immaginava il futuro. Nell'ossessione del tempo e per il tempo, in un momento in cui il futuro ha smesso di offrire speranze, la caratterizzazione nostalgica degli ultimi anni sembra voler recuperare la stessa concezione del tempo nel passato, cioè come si viveva allora il futuro. Una riflessione particolarmente toccante di Garrocho è la proposta secondo la quale il contrario della speranza, di quella speranza disattesa nel nostro presente, non è la disperazione, quanto piuttosto l'inversione del tempo, cioè l'esplosione del sentimento nostalgico fino al punto di provare a tornare a una configurazione del passato-presente, vivere il presente con le stesse modalità e atteggiamento del passato.

Se, d'accordo con Garrocho, l'attenzione del presente sembra essere inconsciamente rivolta al passato e se davvero si assiste a una perdita di speranza per il futuro, varrebbe forse la pena capire quale futuro ci si prospetta nell'immediatezza. L'analisi del professore spagnolo compie salti tra il concetto di progresso e speranza, si muove tra la promessa e la disillusione, e sembra in qual-

che modo orientare la storia verso un punto di non ritorno. In questo senso, sorge naturale, come sempre e ancora piú di sempre, l'interrogazione sulla visione del tempo: come saremo visti nel futuro noi, che in questo momento vogliamo tornare a vivere il futuro come nel passato? Se il fallimento della speranza è dovuto alla promessa disattesa del progresso, è vero, come scrive Garrocho, che la risposta può essere l'inversione del tempo, la volontà di tornare a credere in quella promessa. C'è però forse anche un'altra via d'uscita: rinunciare a quel concetto di progresso o a quel concetto di promessa. Nel momento in cui avviene il tradimento, si può desiderare di tornare a vivere quel sodalizio immacolato con la promessa o, in alternativa, riconsiderare il contenuto di quella promessa. La domanda che genera *Sobre la nostalgia* è: si può ambire a costruire un futuro liberato dall'ossessione del progresso? Si può applicare, in chiave retrospettiva, una tradizione filtrata dall'insegnamento del tempo? Il fallimento della promessa può essere l'occasione per una nuova promessa, supportata dal concetto di perdono, quel perdono che Derrida suggeriva potesse essere possibile solo senza dimenticare il male. C'è qualcosa che va perdonato se si vuole ambire alla costruzione del futuro: se, come scrive Garrocho, il «futuro è l'unico passato che si può cambiare», allora forse l'esperienza del tradimento della promessa dovrebbe averci insegnato a perdonare: perdonare, nel presente, il fallimento della promessa che inevitabilmente ci riserva il futuro.

Fabio Scalese
*Centro de Ciencias Humanas y Sociales, Madrid*
DOI: 10.5281/zenodo.3903563

Johann Friedrich Blumenbach, *Contributi alla storia naturale*, a cura di Mario Marino, prefazione di Giulio Barsanti (Milano-Udine: Mimesis, 2018).

A pochi anni di distanza dalla ristampa moderna dei *Beyträge zur Naturgeschichte* di Johann Friedrich Blumenbach, avvenuta nel 2014 per la serie Historia Scientiarum dell'editore Olms, Mario Marino ne propone, ora per il pubblico italiano, un'accurata traduzione (*Contributi alla storia naturale*) che esce, nel 2018, per la collana di Mimesis Filosofia/Scienza a cura di Vallori Rasini. Certo la traduzione in italiano dei *Beyträge* colma un vuoto nella ricezione di Blumenbach in Italia, poichè oltre alla traduzione del saggio sul *Bildungstrieb* (*Impulso formativo e generazione*, Salerno: 10/17, 1992) a cura di Antonella De Cieri e alle traduzioni ottocentesche dell'*Handbuch der Naturgeschichte* (*Manuale di storia naturale*, Lugano: Vanelli, 1825, Milano: Fontana, 1826-1830), non si davano al momento significativi sviluppi nel panorama italiano degli studi blumenbachiani. Ma non si tratta solo della traduzione di uno dei più emblematici scritti di uno dei «Maestri della Germania» (L. Marino, citato a p. 11). Marino coglie l'occasione per ampliare le sue ricerche su Blumenbach e sugli stessi *Beyträge*, in parte avviate appunto quattro anni prima con l'edizione tedesca, in parte affidate ad articoli più recenti, come quello pubblicato insieme a R. Bonito Oliva e G. D'Alessandro sulla *Storia naturale e antropologia nei blumenbachiani "Beyträge zur Naturgeschichte",* «Studi filosofici», XXXIX, 2016, pp. 309-324, e proporre, con ciò nuovi spunti e direttrici di indagine, a partire dall'ampiezza e ricchezza tematica del testo blumenbachiano.

Insieme alla prima (1790) e alla seconda (1811) parte dei *Beyträge*, il curatore acclude al volume un essenziale apparato filologico con le aggiunte e le modifiche più significative apportate da Blumenbach nella revisione della prima parte (pubblicata nel 1806) e, a seguire, un appendice iconografico con le 'vignette' di Daniel Chodowieki (1726-1801), raffiguranti le cinque 'varietà' in cui Blumenbach aveva suddiviso la 'specie' umana. Sono tali appendici, non presenti nell'edizione tedesca del 2014 dei *Beyträge*, unitamente alla breve Prefazione di Giulio Barsanti e alla «lunga e sapiente introduzione» (così Barsanti a p. 7) del curatore sulla genesi del testo e sulla storia della sua ricezione (Introduzione, pp. 11-65), a fare del volume non la semplice trasposizione dei *Beyträge* in lingua italiana, ma un nuovo originale contributo nella storia editoriale del testo e non solo. Si tratta di un valido strumento che consente di ripercorrere lo sviluppo del pensiero e della metodologia della storia naturale di Blumenbach, finalizzato alla promozione e alla continuazione, anche in Italia, delle ricerche incentrate su questo autore.

Rispetto al contributo fornito nel saggio introduttivo all'edizione tedesca, Marino approfondisce qui in maniera decisiva le tappe che hanno condotto alla pubblicazione dei *Beyträge*, senza dimenticare la premura con cui Blumenbach si era assicurato di poter affidare le illustrazioni al «pittore dell'anima» Chodowiecki (Intr., p. 18), il «più ricercato, acclamato, costoso e rimunerativo incisore in Germania» (Intr., p. 21), noto anche per le sue illustrazioni nei *Göttingen Taschencalender* di Lichtenberg, nei *Physiognomische Fragmente* di Lavater e nell'*Elementarwerk* di Basedow (Intr., p. 20). Se è pur vero che l'attenzione per le immagini è indice degli intenti divulgativi dell'autore, preme a Marino sottolineare come essa racchiuda una valenza anche scientifica, attestando l'impegno decennale di Blumenbach per la realizzazione di quest'opera nonché il suo interesse per la componente concreta ed empirica, figurativa e comparatistica delle indagini naturalistiche e antropologiche. Così, a dispetto della ricezione non troppo fortunata dei *Beyträge*, di cui Marino individua alcune possibili cause (dal «loro carattere non strettamente accademico» al dichiarato «intento anche divulgativo» dell'autore), i *Contributi alla storia naturale* non sono affatto uno «scritto poco originale e di scarso impatto» ma, al contrario, – e questa è una delle tesi principali del curatore – si inseriscono in pieno nella produzione scientifica e originale del loro autore e ne riflettono, in ogni caso, le idee fondamentali rintracciabili nelle altre opere (Intr., p. 14). Testimonianza di ciò sono i rimandi ad altri scritti, come le *Institutiones Physiologiae* del 1787 (p. 86), il *De generis humani varietate nativa* (p. 91) e le *Abbildungen naturhistorischer Gegenstände* (1796-1810), viceversa, i riferimenti ai *Beyträge* presenti in altri lavori di Blumenbach e segnalati da Marino nel saggio introduttivo. Rispetto a scritti più incentrati tematicamente (come il *Bildungstrieb*) o decisamente più ampi e completi (come l'*Handbuch*), i *Beyträge* sono l'opera che maggiormente si presenta quale «sintesi così elementare, di dimensioni contenute e, al contempo, filosofica-

mente e intellettualmente aperta e significativa della storia naturale» (Intr., p. 16).

Il confronto puntuale tra le diverse edizioni dell'opera, reso possibile dall'accurato lavoro di Marino, consente di coglierne non soltanto l'ampiezza disciplinare e la successione cronologica degli interventi al testo, ma anche il tentativo blumenbachiano di elaborare le sue idee in una unità coerente, nonostante l'apparente eterogeneità tematica. Marino coglie, inoltre, la connessione tra prima e seconda parte come rapporto tra «assunti teorici, metodologici e disciplinari» – la prima – e le loro «verifiche empiriche» – la seconda (Intr., p. 17). Rivolgendosi a un pubblico ampio di lettori non specialisti – lo stesso autore lo dichiara all'inizio del Discorso preliminare della parte prima (p. 72) – Blumenbach intende, nello stesso tempo, divulgare le sue più recenti indagini nell'ambito della storia naturale ma anche emendare, ampliare, arricchire le sue ricerche. Ciò è evidente anche dal lavoro di revisione che egli compie tra il 1790 e il 1806 prima di pubblicarne una seconda edizione (1811).

Nell'appendice filologica (pp. 153-177) l'accento è posto sulle modifiche terminologiche e concettuali che consentono di cogliere i progressi compiuti da Blumenbach nel quindicennio tra la pubblicazione della prima e della seconda parte dei *Beyträge*. Attraverso una ricostruzione cronologica e comparativa tra le diverse edizioni, si mostra appunto il percorso compiuto da Blumenbach, l'evolvere delle sue posizioni rispetto ai dibattiti dell'epoca su questioni di tipo naturalistico e antropologico, la sua esigenza di correggere, rivedere e sostituire determinati passaggi, termini o concetti che va di pari passo con gli sviluppi compiuti nell'ambito delle sue indagini empiriche. A tal proposito, Marino individua alcune tendenze che segnano il passaggio da un'edizione all'altra riguardanti sia questioni concettuali che di metodo, mostrando come, in particolare, Blumenbach si fosse sforzato di rendere più sistematica e rigorosa la sua disciplina.

Tra gli esempi più significativi di tali tendenze individuate da Marino nel saggio introduttivo e rintracciabili dal confronto tra le edizioni del 1790 e del 1806 è l'«uso sempre più sistematico del termine 'razza'» che Blumenbach sostituisce, nella seconda edizione, a quello di «varietà» (*Varietät*, *Spielart*), inizialmente privilegiato per indicare i differenti tipi umani. In connessione a ciò è fondamentale notare come, sempre dal confronto tra le due edizioni, emerga anche l'impegno umanistico, le preoccupazioni di tipo anti-razzista e anti-schiavista (Marino sottolinea, a tal proposito, il contatto di Blumenbach con «le centrali internazionali dell'abolizionismo», Intr., p. 44) e la difesa della tesi dell'esistenza di un unico genere umano, sostenuta anche da autorevoli studiosi come Linneo, Bonnet, Haller (p. 88). Marino riporta anche le tre note conclusive aggiunte da Blumenbach nel 1806 alla prima parte dei *Beyträge*, che sono anch'esse decisive per inquadrare meglio dal punto di vista 'teorico' quel mutamento terminologico da 'varietà' a 'razza'. Affermare che nella natura vi è una «successione graduale» di configurazioni (*Bildungen*), tali per cui «l'una si riversa nell'altra» (un concetto ribadito più volte da Blumenbach), significa, a ben vedere, avvalersi di un'immagine ben consolidata nelle concezioni della natura tra Sei e Settecento, quella cioè della 'scala' o della 'catena'. Ma tale affermazione, ammonisce Blumenbach, non va intesa in senso 'metafisico', bensì – e qui a ragione Marino individua la matrice kantiana del discorso – nel suo «uso regolativo» (pp. 170-171; cf. Intr., pp. 28-29, 40 e 42, ma anche 64). Marino ricorda a tal proposito lo scambio epistolare tra Kant e Blumenbach, avvenuto proprio nel 1790 all'indomani dell'uscita della terza critica e dei *Beyträge* (che Blumenbach avrebbe inviato a Kant, il quale già li possedeva) e ascrive anche a tale scambio l'origine della revisione, da parte di Blumenbach, di alcuni punti della prima sezione dei *Beyträge* (Intr., p. 42 in nota, ma più nel dettaglio Oliva, D'Alessandro, Marino, *Storia*, pp. 319-320).

Le intenzioni di Marino sono anche quelle di far chiarezza sulle posizioni teoriche del cosiddetto «geometra della razza» (Marino ricorda la «formula a effetto» usata da S.J. Gould, Intr., p. 12), il cui pensiero fu facilmente e volentieri frainteso quando Blumenbach era ancora in vita. Inquadrare Blumenbach come momento chiave della «storia dell'antropologia razziale» significa cogliere il suo contributo come «ricerca empirica teoricamente fondata e metodologicamente determinata» (Intr. P. 64) che trova spazio e applicazione, per la sua pregnanza, anche al di là della storia naturale, mostrando la capacità di quest'ultima – ancora con Marino – «di servire da scienza ausiliaria ad altre discipline», non ad ultimo, alla politica e «al diritto sulle delicate questioni della schiavitù e della discriminazione razziale» (Intr. P. 17). Tematiche anche molto attuali come i concetti di 'catastrofe', 'mutabilità', 'estinzione', 'razza' fanno di questo volumetto un «documento prezioso lungamente dimenticato» (Intr., p. 65) e tuttavia ad oggi ricco di nuovi spunti di ricerca. Vi è da pensare che questo sia soltanto un punto di partenza per la prosecuzione, anche in Italia, degli studi su Blumenbach e sul suo ruolo nella storia del pensiero scientifico e filosofico di età moderna.

Laura Follesa
*Friedrich-Schiller Universität Jena*
DOI: 10.5281/zenodo.3903565

Sami Pihlström, Friedrich Stadler, Niels Weidtmann (eds.), *Logical Empiricism and Pragmatism* (Dordrecht: Springer, 2017).

The 19th volume of the Vienna Circle Institute Yearbook contains most of the papers presented at the international conference on "Logical Empiricism and Pragmatism," held at the University of Vienna on 7-9 November 2013. The volume explores several aspects of these two research programs, from both a historical and a theoretical perspective, in order to show to what extent they can be seen as comparable views of scientific knowledge. As one of the editors explains in his paper (p. 139 fn.) "logical empiricism" is preferred to "logical positivism," the former expression meaning "the somewhat broader set of ideas and the slightly more inclusive philosophical approach that survived the collapse of the Vienna Circle (and thus the collapse of logical positivism in a strict sense)." The aim is therefore not to focus exclusively on the Viennese movement, but rather to go beyond the limits of the schematization that traditionally can be encount-

ered in the history of the philosophy of science, and look at the connections existing between pragmatism and logical empiricism from a broader point of view.

Because of the importance of logical positivism for the history of philosophy, some of the contributions collected in this volume of course deal with the personal and scholarly exchanges that after the First World War involved members of the Vienna Circle (even in its early form) and pragmatist philosophers. But their investigation is not limited to a reconstruction of the relationship between Charles S. Peirce, William James and John Dewey, and leading figures of the Logical Empiricism movement. On the contrary, the aim of these papers is primarily to focus on the theoretical issues approached by members of the Vienna group and stress their compliance with some perspectives developed by pragmatist thinkers.

In his paper on *William James and the Vienna Circle*, for example, Massimo Ferrari deals with the Viennese approach to the problem of truth and the issue of scientific knowledge, and, broadly, with the anti-metaphysical attitude which is defended by Hans Hahn, Philipp Frank and Otto Neurath. As Ferrari shows, the interests of these authors overlapped with those of the American pragmatists (especially James), and it can be argued that "in the Viennese milieu it was possible ... to grasp the veritable core of pragmatist account of truth as opposed to the correspondence theory" (p. 29). Furthermore, Ferrari argues that "for Neurath, Frank and the other young supporters of the scientific world-conception the pragmatist method had a great significance in the struggle against metaphysics they were involved in" (*ibid*.). In fact, as Ferrari aptly maintains, early on 1933 Hans Hahn was aware of the fact that the "great problem" of truth required a new solution, and claimed to "side with the pragmatist conception of truth" (p. 15). Moreover, Frank later agreed with Hahn that "logic need[ed] a drop of pragmatic oil" (p. 16), and Neurath found "in James a new kind of thinker, close to the renewed epistemological perspective" (p. 32). As Neurath wrote in 1937, it was clear to the founders of the Vienna Circle that they would have found "a friendly welcome ... in a country in which Peirce, James, Dewey and others have created a general atmosphere that [was] empiricist in many respect. The very fertile American manner of thinking successfully combines with the European in this field" (p. 17).

Hahn's, Frank's and Neurath's interest in pragmatism is also addressed by Thomas Uebel (*American Pragmatism, Central-European Pragmatism and the First Vienna Circle*). In his contribution, Uebel argues that "the affinity between pragmatism's conception of meaning and the views developed in the Vienna Circle became plane [after 1928], for only then did Frank and some of his colleagues appreciate the relevance of pragmatism for the philosophy of science the Circle was developing" (p. 88). In his interesting reconstruction of the relationship (effective and/or elective, depending on the case) between the many authors involved, Uebel defends the thesis that "the early sympathies of some Vienna Circle members were based to a large extent on their appreciation of the work of [some] Central-European philosopher-scientist rather than merely the then prominent key text of pragmatism" (p. 83). That is to say that the reason why James's philosophy has been positively received in Austria does not reside primarily in "the close relationship between the truth concept of the modern logical movement and that of pragmatism" stressed by Frank during the 1929 *Congress of German Mathematicians and Physicists* held in Prague (p. 89), but rather in the predisposing role played, for example, by Ernst Mach and Ludwig Boltzmann (p. 96).

For what concerns Mach, his role is particularly emphasized by Friedrich Stadler, who contributed to the volume with a paper that focuses on Mach's 1895 book (*Ernst Mach and Pragmatism – The case of Mach's* Popular Scientific Lectures). In exploring the several connections that can be encountered between Mach and James, and in stressing the elements that Mach's epistemology has in common with fundamental pragmatist ideas, Stadler conceives the *Popular Scientific Lectures* as reflecting "the context of the time, while also revealing an approach to the sciences that places emphasis on the historical-genetic perspective and enriches the social role of research between the everyday world and the professional world even from today's perspective" (p. 13). As Stadler argues, this approach can also be found in North-American pragmatism, and it is thus possible to "conclude that pragmatic philosophy was already present in Austria and Germany in parallel, but not explicitly under this notion and American label" (p. 14). In fact, "Mach had already claimed pragmatic positions in epistemology and methodology before his reading of Peirce and direct encounter with James" (*ibid*.). The audience was therefore ready to receive James's 1907 book on *Pragmatism*, which has been immediately translated, supported and popularized in Austria by Wilhelm Jerusalem, to whom Stadler, Uebel and Ferrari all pay particular attention (see e.g. pp. 11, 22 ff. and 91).

The above quoted papers give an important contribution to the volume, for they emphasize some mostly neglected elements pertaining to the late nineteenth- and early twentieth-century history of philosophy. But this does not exhaust the aims of the book, which also collects some chapters more theoretically oriented, e.g. those written by Donata Romizi (*Classical Pragmatism and Metaphysics: James and Peirce on Scientific Determinism*) and Giovanni Rubeis (*Beyond Realism and Antirealism? The Strange Case of Dewey's Instrumentalism*). As the titles of these chapters suggest, they deal with the views of leading figures of classic pragmatism on some fundamental issues of the philosophy of science, namely scientific determinism and realism. Romizi especially focuses on the issue of scientific determinism, and argues that "despite its being 'metaphysical' and thereby 'nonsensical' according to the Vienna Circle's 'scientific world conception', [that issue] bothered philosophers like William James and Charles Peirce" (p. 43). The investigation on these authors is carried on through a contextualization of their views in the late nineteenth-century French anti-deterministic tradition. Romizi deals with the ideas developed by Charles Renouvier, Henri Poincaré, Henri Bergson, and Émile Boutroux, in order to show that, within this French tradition, determinism is criticized for being "the product of a rationalist perspective which neglects many aspects of reality and focuses on quantities" (p. 49). According to Romizi, the anti-determinism which

both James and Peirce defended can be compared to that view. For example, the Jamesian criticism of "intellectualism" converges in many ways with the ideas of Boutroux and Bergson. Furthermore, the debate concerning the problem of necessity – and, therefore, determinism – provoked by Peirce on *The Monist* between 1891 and 1893, involves some aspects that Boutroux stressed, too (p. 57 f.).

This is only a selection of the elements that Romizi takes into account, in order to argue that it would be necessary to reconsider the pragmatist anti-deterministic attitude, and not to conceive it as a merely "metaphysical or irrationalistic reaction against science" (p. 63). On the contrary, according to Romizi "most arguments against scientific determinism put forward ... by Renouvier, Boutroux, Poincaré and Bergson, as well as by James and Peirce, ... were rooted in an empiricist attitude, which emphasized the value of experience, observation and practice against a deductivist, rationalistic and theoretical standpoint" (*ibid*.). Following this line of thought, it would thus be possible to reconceive the very meaning of metaphysical questions, and especially to compare the pragmatist and the logical empiricist view of them. According to the logical empiricists, an issue such as the contraposition between determinism and indeterminism is fundamentally nonsensical. From the pragmatist standpoint, we can rather make sense of some metaphysical questions, for they appear "to have consequences not only for practical life in general, but also ... for scientific practice" (p. 64). Therefore, there seems to be room for a new entrée of metaphysics, even though pragmatically (re)conceived.

Something similar can be argued of John Dewey, as Rubeis's paper suggests. Rubeis explores selected aspects of Dewey's instrumentalism and tries to show that, contrary to what one could expect, it may be compatible with a moderate form of realism. Rubeis especially considers some interpretations of Deweyan instrumentalism, e.g. Hans Reichenbach's, according to which one must interpret it as an anti-realism; Peter Godfrey-Smith's, claiming that "Dewey's philosophy of science is an unorthodox form of realism" (p. 69); David L. Hildebrand's, who considered Dewey's philosophy "as an attempt to avoid some of the classic dualisms like realism/idealism or mind/body by dissolving rather than solving the problems in question" (p. 70). A closer investigation on Dewey's natural empiricism allows Rubeis to argue that his instrumentalism should be considered as a view *beyond* realism and anti-realism, for "the crucial point of Deweyan philosophy [consists in] his re-thinking the theory-practice relation" and in his attempt to overcome that dualism (pp. 79-80). These observations once more allow us to compare the approaches of classic pragmatists and (especially late) logical empiricists. The former, in particular, attempted to focus on the concrete applications of concepts and theories, on the practical consequences of scientific inquiry instead of on its ontological implications (p. 80). In doing this, the classis pragmatist thinkers showed that (and how) an empirically oriented philosophy can be compatible with metaphysics, but also that to defend an anti-metaphysical view does not necessarily involve the complete rejection of fundamental metaphysical issues pertaining to the Western philosophical tradition. When approached pragmatically, these issues in fact make sense, although their ontological value must be limited and, consequently, re-defined.

The last paper which is worth mentioning in this review is the one written by one of the editors of the volume, Sami Pihlström (*On the Viennese Background of Harward Neopragmatism*). The paper is inspired by the same interest in the relationship between pragmatism and logical empiricism that characterizes Ferrari's, Stadler's, and Uebel's papers, and focuses on the purely theoretical issue of metaphysics in pragmatism and neopragmatism. According to Pihlström, "metaphysics has forcefully returned to the center of mainstream analytic philosophy, and pragmatists need to reflect on their ways of reacting to such developments (p. 142). Pihlström's investigation on Ludwig Carnap's influence on Putnam (pp. 143 ff.) and on the role played by Ludwig Wittgenstein's views in the emergence of neopragmatism (pp. 153 ff.) are thus aimed at dealing with that issue. But these connections also reveal how strongly the Viennese background influenced the further development of classic pragmatism, thus contrasting "the thesis that pragmatism ... was 'eclipsed' by logical empiricism (and later by analytic philosophy)" (p. 141). On the contrary, Pihlström argues that "not only does the pragmatic maxim ... resemble the logical empiricist's verificationist theory of meaning ..., [but] also the resolute rejection of unempirical metaphysical speculation, as well as the link between scientific progress and social progress, can be regarded as points of contact between the two traditions" that lasted also after the second generation of both pragmatists and logical empiricists (*ibid*.). The never abandoned critical attitude towards metaphysics shows in particular that "neopragmatism still remains committed to important logical-empiricist ideas" (p. 142) and that these two philosophical approaches should be studied in parallel. Finally, Pihlström argues that, given these correspondences and the general compliance of the two philosophical perspectives explored, it would be possible to develop a new research program resembling both the approaches.

According to Pihlström, an "integrated pragmatist approach" that would lead us back "to the original rendezvous of pragmatism and logical empiricism" (p. 158) can be found in Morton White's *holistic pragmatism*. This epistemological position that "can be extended to a holistically pragmatist ontology of culture as well as to a metaphilosophical account of what is correct and incorrect in both metaphysics and anti-metaphysics" (p. 158) is explored in the final part of Pihlström's paper, leaving the room for further analysis. Although not thoroughly examined, White's view contributes to the general picture that Pihlström aims to draw. When combined with the observations on Putnam's neopragmatism indebitedness to Carnap's ideas, and with the various pragmatist themes that can be found in Wittgenstein's scholarship, White's holistic pragmatism sheds light on what pragmatism should be – and, moreover, on how it should be conceived. As Pihlström conclusively remarks, "pragmatism ought to speak – and also *listen* – not only to the well-known philosophical orientations today, such as analytic philosophy or phenomenology, but also to the marginalized, for-

gotten, and eclipsed ones, whether or not pragmatism itself was ever truly eclipsed by logical empiricism or analytic philosophy" (p. 162).

As a final remark, it is possible to say that the reviewed volume provides a good introduction to an issue which is worth investigating because of its importance for the history of the philosophy of science. In fact, both pragmatism and logical empiricism put the basis for the contemporary approach to the aims and character of the scientific inquiry, and their convergence on some fundamental topics is revealing. As shown above, the papers collected offer a multifaceted view on that issue, thus outlining a picture which is detailed both on the historical side and on the theoretical plane, and (hopefully) stimulating a further debate on that matter.

Pietro Gori
*Universidade Nova de Lisboa*
DOI: 10.5281/zenodo.3903567