



# THE CONCEPT OF AFFECTIVITY

in Early Modern  
Philosophy

GÁBOR BOROS · JUDIT SZALAI · OLIVÉR ISTVÁN TÓTH (eds)

## The Concept of Affectivity in Early Modern Philosophy



*Gábor Boros – Judit Szalai –  
Olivér István Tóth (eds)*

# The Concept of Affectivity in Early Modern Philosophy

© Authors, 2017  
© Editors, 2017

ISBN 978-963-284-820-4 [online]



[www.eotvoskiado.hu](http://www.eotvoskiado.hu)

---

Executive Publisher: The Dean of the Faculty of  
Humanities of Eötvös Loránd University  
Cover: Ildikó Csele Kmotrik  
Layout: Manzana Bt.



# Table of contents

Introduction .....	7
I    Ádám Smrcz: A Neo-Stoic Theory of Natural and Unnatural Affections .....	11
II   Maximilian Kiener: The Compulsion to Believe Something: On Affectivity of Indubitable, Clear and Distinct Perceptions in Descartes .....	24
III  Jan Forsman: Descartes on Will and Suspension of Judgment: Affectivity of the Reason for Doubt .....	38
IV   Hanna Vandebussche: Descartes and Pascal on Imagining Love and Imaginary Love .....	59
V    Judit Szalai: Medicine, Emotion Management and Mind-Body Interaction in the Later Descartes and Early Cartesianism .....	75
VI   Davide Monaco: A New Account of the Objective-Formal Distinction in Spinoza's Parallelism Theory .....	89
VII  Filip Buyse: A New Reading of Spinoza's <i>Letter 32</i> to Oldenburg: Spinoza and the Agreement between Bodies in the Universe .....	104
VIII Keith Green: Spinoza on Reflexive Affects and the Imitation of Affects .....	124
IX   Gábor Boros: Life as Death in Spinoza .....	142
X    Olivér István Tóth: Is Spinoza's Theory of Finite Mind Coherent? – Death, Affectivity and Epistemology in the <i>Ethics</i> .....	159
XI   Brian Glenney: Spinoza on Passionate Misperception .....	174
XII  Christopher Davidson: An Affective Aesthetics in Spinoza and Its Political Implications .....	185

XIII	Zsolt Bagi: Emancipation of the Body in Spinoza's Political Philosophy. The Affective Integration .....	207
XIV	Dávid Bartha: The Human Passions and the Purely Active Will of God: An Introduction to Berkeley's Theory of Emotion .....	221
XV	Dan O'Brien: Hume, Sympathy and Belief .....	240
XVI	Hans D. Muller: Hume on Sympathy, Pity and Impartiality .....	257
XVII	Csaba Olay: Alienation in Rousseau .....	271
Contributors .....		284
Index of Names .....		288

## Introduction

There is scarcely any need to argue for the relevance of the concept of affectivity in early modern philosophy. Theories of emotion in this period have been one of the focal topics our research group, centered around Eötvös Loránd University, Budapest (ELTE), has investigated in the past decades. The research group has taken several forms in the past and is now supported by the National Research, Development and Innovation Office (Hungary) in the frame of a research project on theories of emotions (K120375, “Self-Interpretation, Emotions, Narrativity”) and a research project on Descartes (K125012, “The Cartesian Mind between Cognition and Extension”). We work on various areas within early modern philosophy, with minimally one shared goal: deconstructing the time-honored interpretive distinction between “rationalists” and “empiricists”, which we consider untenable when applied to 17<sup>th</sup>-century thinkers. Conceptualizing affectivity in this period provides us with an interpretive frame independent of and cutting across such unfruitful divisions. This was the main motivation behind our decision to commence the series of Budapest Seminars in Early Modern Philosophy with the topic of affectivity, inviting prospective participants to reinforce or to reject our conviction concerning the fruitfulness of this basic concept. In our view, the participants’ general response to this concealed expectation was illuminatingly positive. We are proudly presenting the volume composed of papers based on the contributions presented at the conference. Besides NRDIO, we are grateful to Eötvös University, Budapest for supporting both the event (held in October 2016 at ELTE) and the preparation of the present volume. Special thanks to the Philosophy Department of the National Association of PhD Students for their support of the seminar.

The aim of this volume is not to present a systematic survey of affectivity in early modern philosophy, but rather to foster collaboration among researchers working in different countries and different traditions. Many of the papers published here are already in implicit or explicit dialogue with others. We hope that they will generate more exchange of ideas in the broader early modern scholarship.

**Ádám Smrcz** discusses the influence of Stoic authors on Justus Lipsius' philosophy in general, and on his theory of affections in particular. He presents the intricacies of Lipsius' attempt to Christianize this ancient philosophical school in a way that was adjusted both to his ever-changing religious affiliations and to the rigor of philosophical reasoning. Lipsius' solution was distinguishing natural and unnatural affections and combatting an emerging prejudice concerning Stoic ethics.

Kiener and Forsman both discuss Descartes' concept of dubitability and indubitability. **Maximilian Kiener** takes a fresh look at the age-old problem of the Cartesian circle in Descartes' *Meditations*. In his view, while most commentators have considered certainty, dubitability and compulsion to believe as psychological phenomena, they should rather be treated as epistemic concepts that might have bearings on psychological notions such as emotions and affectivity. Thus, he urges us to revise our understanding of early modern affectivity from an epistemic point of view. He also presents a new interpretation of the indubitability of clear and distinct ideas. **Jan Forsman** defends an indirect doxastic voluntarist reading of Descartes against direct doxastic voluntarist interpretations. He argues that on Descartes' view we are not capable of rejecting, accepting or suspending judgments based on our volitions; rather, our free will can choose to focus on particular reasons for forming a judgment. This view allows Forsman not to attribute to Descartes the seemingly counter-intuitive view that we can willfully doubt logical truths and accept judgments against all reasons.

**Hanna Vandenbussche** compares Descartes' and Pascal's discussions of love. She shows that while both distinguished imaginary love from intellectual love, they had different views on the role of the imagination. For Descartes it is one faculty of the soul, while for Pascal it is a source of moral corruption. Still, both saw imagination as a source of distortion, and especially as a source of a distorted image of oneself. Surprisingly, this does not prevent them from including the imagination in their accounts of the love of God.

**Judit Szalai** focuses on the philosophical and physical descriptions of the operation of the emotions in the later Descartes and early Cartesians. She argues that contrary to the current mainstream interpretation, Descartes' correspondence with Elizabeth does not represent a break within the Cartesian corpus, but rather shows one aspect of his theory more prominently. She then presents the way in which Descartes' philosophical and physical views map onto the medical practice of the 17<sup>th</sup>-century.

**Davide Monaco** provides a new interpretation of the formal-objective distinction in Spinoza. By tracing back the history of the distinction to the *Timaeus* via Descartes and Suarez, he argues that Spinoza in his mature works does not simply accept the Scholastic or Cartesian version of the distinction. He argues that Spinoza's parallelism doctrine implies the replacement of the objective being of ideas with the formal being of ideas, thus turning the formal-objective distinction into a formal-formal distinction.

Both Buyse and Green discuss the consequences of Spinoza's conatus doctrine for his theory of affects. **Filip Buyse** focuses on Spinoza's *Letter 32* to Oldenburg, in which he answers the question concerning the coherence of finite bodies in the universe. Buyse argues

that Huygens' discovery of the synchronization of pendulum clocks might have inspired Spinoza in his views on the ratio of motion and rest of bodies in the universe. **Keith Green** discusses the problem presented by reflexive affects in Spinoza, especially self-love and self-hatred. He tries to account for them in terms of imitating the affects of others. This, however, proves difficult in the cases of love and hate, where the definitions offered by Spinoza seem to presuppose external causes. Philological inconsistencies in Spinoza's text reveal a deeper philosophical difficulty: Spinoza is unable to define what it is to be internal and external to the subject. Green reconstructs this difficulty in light of contemporary theories of emotion.

Boros and Toth both focus on the problem of death in Spinoza. **Gábor Boros** argues that commentators miss an important layer of Spinoza's text as long as they understand by death simply the decay of the physical body. Placing Spinoza in the double context of the European traditions focused on both the physical death of the body and on the spiritual death of the individual required for a spiritual rebirth, he shows that Spinoza is also heir to this latter tradition and that some of his remarks on death can be understood only in light of the Pauline way of thinking. **Oliver Istvan Toth** approaches death from Spinoza's identity doctrine. He argues that contrary to many interpretations Spinoza's affective understanding of ideas does not make him less of an ethical intellectualist. By focusing on Spinoza's inability to prove that humans are necessarily mortal, he shows that Spinoza either has to give up a core tenet of his philosophy of mind (the conatus doctrine, the identification of epistemic value with epistemic autonomy, or universal intelligibility), or has to embrace the claim that humans might become immortal.

Similarly to Toth, **Brian Glinney** also focuses on the relationship between affectivity and epistemology in Spinoza, addressing the example of the idea of the sun. He shows that inadequate perceptual ideas are both affects and beliefs and assesses possible ways of accounting for this fact. He arrives at a graded interpretation of belief, according to which we continuously revise our beliefs in light of the complete set of ideas we have.

**Christopher Davidson** provides a reconstruction of Spinozist aesthetics. He aims at presenting a historically situated interpretation with political implications for contemporary society, which can plausibly be applied to actually existing cultures. In his view, art is a form of imagination that fosters community and structures the whole of societies from the times of Moses and the prophets to contemporary sub-cultures defined by musical genres.

**Zsolt Bagi** revisits the problem of the physical definition of essence in Spinoza's philosophy. He argues that based on the physical description of the integrity of individuals the affective integration of human beings into a political community can provide a recipe for political emancipation.

**Dávid Bartha** focuses on Berkeley's theory of emotions and argues against those interpretations that present him exclusively as a metaphysician interested in the philosophy of religion, detached from everyday concerns. In Bartha's view, emotions have an important role to play in Berkeley's moral philosophy and the real aim of his philosophy is fostering the right kind of passions.

O'Brien and Muller focus on Hume's theory of affectivity and his ethics. **Dan O'Brien** shows that according to Hume's account belief acquisition is closely related to sympathy: people often accept the beliefs of those with whom they feel sympathy. This raises the questions of whether these beliefs are justified or not. Focusing on this question, O'Brien argues against the mainstream evidentialist interpretation of belief acquisition and for a more associationist interpretation in Hume. **Hans D. Muller** discusses the challenge that Hume's sentimental theory of ethics faces when trying to incorporate the principle of impartiality. In order to meet the challenge, Hume invokes the metaphors of the anatomist and the painter. Discussing these metaphors, Muller presents a way in which Hume is able to ward off the Kantian critique of partiality and present an impartial sentimental ethical theory.

Finally, **Csaba Olay** focuses on the notion of alienation in Rousseau. He shows that Rousseau's concept of alienation is not quite the same as the one used by the Marxist tradition. Olay shows that contrary to the Marxist possession–disappropriation–reappropriation triade, Rousseau only posited a possession–disappropriation dialectic. As a consequence, Rousseau was not able to identify a non-alienated condition either in his theory or in actual society.

# I

*Ádám Smrcz*

## A Neo-Stoic Theory of Natural and Unnatural Affections<sup>1</sup>

### 1. Introduction

According to a well-known claim, the most important intention of Lipsian philosophy was to Christianise Stoicism (Oestreich 1982, 15). Writing about Justus Lipsius, Charles Taylor reaffirms this view, but he also adds that when the two seemed to be in conflict with each other, Lipsius leaned to the Stoic rather than to the Christian side. The two major points where Christianity and neo-Stoicism contradicted each other according to Taylor was (1) the question of divine grace, what the Stoics regarded as unnecessary for human salvation, and hence, took a quasi-Pelagian position; and (2) the question of *apatheia*, what the Stoics regarded as a desirable condition to be achieved by humans. Since Taylor regarded Lipsius as the precursor of Deist metaphysics and an early advocate of secular ethics due to his reliance on reason instead of revelation, Lipsius is credited with an important role in the process of secularisation (Taylor 2007, 115–124).

The problem of *natural* and *unnatural affections* in Lipsian philosophy may help to illuminate whether Lipsius really distanced himself from Christianity. Here we will use the word, Christianity in a narrow sense<sup>2</sup>, and will be referring only to Calvinism, the most relevant confession in Lipsius' case. The question at hand is whether Lipsius advocated the case of *apatheia* (something that could be regarded as alien to Christianity), or rather endorsed *eupatheia* as the proper aim of the human soul. This latter would be in harmony with the requirements of Christianity, and, according to my thesis, Lipsius did in fact choose this direction. While most of our affections originate from the body (and are, hence, unnatural from the perspective of the soul), the soul also has some inherent affections which clearly should not be eradicated.

---

<sup>1</sup> The paper was funded by the NKFI/OTKA K 125012 project.

<sup>2</sup> Although Lipsius scholarship generally does not specify what exactly should be taken by the term Christianity (see Oestreich 1982; Brooke 2012, 12–37).

Although the terms *natural* and *unnatural* affections do not occur explicitly in any Lipsian text, the author often refers to some movements of the soul as *unnatural*.<sup>3</sup> Also, it is well-known that Lipsius was primarily preoccupied with providing remedies for affections originating in the external world, as the subheading *in publicis malis* or *in external evils* also suggests.<sup>4</sup> Affections are defined by the author as movements of the soul, which have to be blocked the soul's firmness (*robur animi*). Since external movements are transmitted to the soul through the body, from the body's perspective they have to be regarded as unnatural. The problem is that although Lipsius provided his readers with an elaborate critique on *unnatural* affections of the soul, he did not discourse on the *natural* ones in a similarly comprehensive way. Still, his fragmentary remarks on the question make it clear, that he did not reject their existence.

This is the point where the Lipsian dualism of substances gains importance. While Platonists and Peripatetics generally held that the human soul had separate parts, ancient Stoics were said to regard it as a homogeneous entity. This obviously entailed entirely different psychological attitudes in each case. While there were those who claimed the existence of separate parts (each of which was partly responsible for "balancing" the power of affections originating in the other), the ancient Stoics had to regard affections in a disjunctive manner: either there should be such movements in the soul, or there should none be at all.

Such an interpretation of Stoicism is – at least at first sight – reinforced by the fact that Lipsius' most important ethical work, the *De Constantia*, is intended to provide one with remedies against external evils (*publicis malis*) or foremost. As a consequence, Lipsius constantly speaks of affections that are harmful, thus to be avoided by all means.<sup>5</sup> Still, this does not completely contradict the claim that I advocate for here, since – according to my interpretation – it is only the problem of external evils in the *De Constantia* which confines Lipsius to the discussion of *unnatural affections* almost exclusively.

It is important to note that neither of Lipsius' philosophical works was intended to be original, but merely a faithful recapitulation of ancient Stoic doctrines. Moreover, he for the most part expressly discloses when he differs from his forerunners, or regards their doctrines as untenable (e.g. *De Constantia*, 55–64). Taken from this perspective, if Lipsius was really advocating the case of eupatheia, he was doing so due to his interpretation of ancient sources. Furthermore, if he regarded his ancient forerunners as the advocates of eupatheia, than his interpretation is consistent with Graver's, who recently claimed the same concerning ancient Stoicism (Graver 2007).

This paper consists of four major parts: I will highlight some Calvinist demands which are relevant from the perspective of a theory of affections. Based on the Calvinist topos of the heart of steel, I will show that Lipsius was doing his best in order to adhere to the Calvinist requirements in this field (parts 1 and 2). Later, I will show that the Lipsian definition of constancy did not exclude affections in general, hence meeting Calvinist re-

---

3 "If these [affections of the soul] were really natural [*a natura*], as you claim, why are not they diffused in everyone and according to the same measure?" Lipsius 1584, 24.

4 Oestreich 1982, 13–28.

5 His clearest example is fear, an affection of external origin, which can be elicited contingently, whenever some cause triggers it.

quirements (part 3). However all this advocacy of eupatheia has to be grounded in a dualist framework of substances, because otherwise, the soul – regarded as a homogenous entity – would not be able to block our unnatural affections (part 4).

## 2. Lipsius the Calvinist

Justus Lipsius was born and raised in Catholic Louvain, but early in his life, he moved to Jena where, due to the city's Lutheranism, he decided to reject his earlier confession. Only after moving to Leiden did he convert to Calvinism, which he followed most of his life (Oestreich 1982, 16–17). Since his most important philosophical work the *De Constantia* was written and published during his Calvinist period, the claim that Lipsius' major concern was the Christianisation of Stoicism can also be understood in a stricter sense, according to which what Lipsius had in mind was to make Stoicism consistent with Calvinism.

The task was not at all an easy one, since Calvin – albeit a former admirer of Seneca – directed harsh criticisms towards the Stoics (Bouwsma 1989, 132). According to some interpreters, Calvin's early admiration of Seneca was in fact due to the latter's literary style, not to his philosophy at all (Hugo 1969, 3–63). In any case, by the time he formulated the term *neostoicism* in 1536 (Sellars 2017) – which later became a word of slander<sup>6</sup> – he was already a major opponent of the school.

One of Calvin's major concerns was the Stoics' theory of affections (more precisely, one that he attributed to them): he claimed that the Stoic theory of *apatheia* was intended to deprive humans of their proper nature, since the feelings of joy and sadness, all being divine gifts to humans, had to be regarded as essentially belonging to humans. This is manifest, for instance, from his commentary on the *Acta Apostolorum* as well, where he examines the feeling of grief that the believers felt when Apostle Stephen died. As he says:

They began to mourn [...]. Earlier, a sorrowful and mournful event could make many people abandon the case that they had previously liked. In turn, these people [viz. the Apostles] declare, instead of mourning Stephen, that they are not the least frightened by his death, [and] hence, they hold on further to approving to their case: meanwhile, they consider how much loss the church suffered due to one man's death. We need to abandon this infuriating philosophy, which rather requires man to be stupid, than to think. The Stoics, who kept every passion away from people, must have been short of common sense. And today some fanatics would even bring this folly back inside the church (Calvin 1863, 175–176).

Calvin's insistence on this anti-Stoic standpoint might be – of course – due to his teleological view on creation, according to which nothing can come into being without a purpose.

---

<sup>6</sup> See: Cudworth 1996; Chapell 1999; Leibniz 1999, 1384–1388.

And, since God's gifts cannot be regarded as purposeless, this claim evidently renders the existence of affections meaningful. On the other hand, Calvin seems to make a sharp distinction between spiritual affections and their opposites: a recurring example of the former is spiritual joy (*gaudium spirituale*), a passion enjoyed by believers when they actually possess their faith in salvation. It is also present in such cases as when people feel sadness over the loss of someone, but simultaneously feel delight over his salvation. Here people can have conflicting passions, which could never be the case according to the Stoic theory, which regards passions as judgements, and hence, no difference can exist between rational and emotional responses to an impulse Gosling (1987, 179–202). According to Calvin this spiritual delight is available for humans all the time, when

everything, that we possess, is regarded by us as a freely given gift [by God]. Since those can easily swallow any hardship, who turn their minds [towards God]. This is why, those, who do not possess even a taster from spiritual goods [*nullum spirituum bonorum*], are oppressed by the sadness of their souls. While those, who live through any pain, as if it was a trial necessary for their salvation, not only *elate*, but also turn themselves towards an occasion of joy (Calvin 1863, 211).

Calvin argues here that feelings of pain and sorrow should not be repressed, but have to be endured and taken as an occasion to discover spiritual delight. Those who oppose this view are often referred to as “iron hearted” by Calvin:

Since [the Stoics] demand, that everyone's heart should be made of iron, they regard [everyone] as possibly the weakest and most effeminate person (Calvin 1863, 175–176).

### 3. Did Lipsius follow Calvin?

With such recurring claims about having hearts of iron, Calvin obviously refers to the question of Stoic *apatheia*. What we need to examine is whether this claim applies to Lipsian Neostoicism. Could Lipsius meet the Calvinist standards when trying to Christianise Stoicism?

The *De Constantia* tells the story of a man called Lipsius fleeing his country due to civil war. While trying to escape the calamities caused by this war, he comes across *Carolus Langius*, who happens to be a Stoic sage. Shortly after beginning a conversation, Lipsius tells him that he is tormented by the *perception* of calamities, and fleeing seems to be the only way for him to avoid perceiving them. He sums it up this way:

my heart is not coated by some kind of steel [*nec chalybs aliquis mihi circa pectus*] (Lipsius 1584, 2).

The imagery might easily invoke that of Calvin's, but it is important to note that Lipsius, the interlocutor of the dialogue, is not yet a Stoic here, hence, his words cannot yet refer to the Stoic state of mind. This claim anticipates the conclusion of the whole work: without the doctrines of Stoicism, one has to have a heart of steel in order to endure external evils. Hence, of major importance, Stoicism and having a heart of steel cannot be identical.

Lipsius the interlocutor, sees only two options: feeling or having a heart of steel. However, Langius the sage later tells him that what he'd outlined was only a false dilemma, since the affections evoked by the perceptions are already stored in Lipsius, he doctrines of Stoicism thus providing him a third option.

At first, Langius identifies the body as the storehouse of affections, but later specifies it as "the outermost layer of the soul [*summa animorum cute*]" (Lipsius 1584, 6). However obscure the phrasing is, it is a crucial one, as we shall see later. But for now, Langius concludes his line of argument by forming the most famous claim of the treatise as follows: "we should flee from our affections, not from our country [*itaque non patria fugienda, Lipsi, sed adfectus sunt*]" (Lipsius 1584, 6).

This phrasing also provides us with an insight into the difference between the Calvinist and Lipsian claim of having hearts of steel: fleeing from our country was previously regarded as the opposite of having a heart of steel (since Lipsius had to choose from either perceiving calamities *or* having a heart of steel). Following this train of thought, we might say that fleeing from our affections has to be regarded as the opposite of having a heart of steel as well. Hence, what the Lipsian phrasing suggests here is that the Stoic theory does not transform one's heart into steel, but the lack of it does.

However, this whole line of argumentation concerning the topos of heart of steel proves thus far that the Lipsian program of constancy was intended to be consistent with the Calvinist requirements, at least on a rhetorical level. Our question will be thence whether the rest of Lipsius' work is consistent with this approach.

## 4. Is *constancy* identical with *apatheia*?

First of all, it should be clarified what constancy is in fact. Lipsius famously defined it as a lack of movement:

I mean by constancy the right and motionless weight of the soul [*rectum et immotum animi robur*], which does not *elate* itself by external and fortuitous things, and also does not get depressed (Lipsius 1584, 10).

It is important to note here that this definition does not claim that a soul should not elate itself in any case, since it only claims that external things should not be the cause of its elation. But even more interestingly, Lipsius goes on with this definition in a quasi-Calvinist manner, when he identifies constancy with the foundation of hope in salvation:

I called you, Lipsius, to turn yourself towards constancy, where the certain hope of your salvation lies (Lipsius 1584, 10).

As we remember, Calvin identified *gaudium spirituale* with the state of being certain in our salvation. Lipsius, hence, seems simply to swap *gaudium spirituale* for constancy, since the two are credited with invoking the same emotion in humans.

But does all this imply *apatheia*? As we have previously seen, affections can get hold of the “outermost layer” of the soul. But this claim also implies that only that part of it can be reached by affections, while the rest may remain untouched by them. The motionless weight prescribed by Lipsius, thus can be regarded as applying only to that part.

Furthermore, as Margaret Graver observed, the word *constantia* was first used by Cicero as the Latin equivalent of several Greek terms, including *eustathia* [meaning: stability], *homologia* [meaning: assent] or *eupatheia*, but *apatheia* is just not among these (Graver 2007, 51). This is consistent with the way Cicero once used the term:

When the soul is moved peacefully and *constantly*, we call it delight [*gaudium*]; but when the soul foolishly and confusedly leaps, that can be called excessive and extravagant joy [*laetitia*], what is defined as the irrational elation of the soul (Cicero 1853, 332).

Here, the notions of *gaudium* and *laetitia* are distinguished, based on whether they are elicited from a soul disposed to produce peaceful and constant movements, or from some other kind. It also turns out from another passage by Cicero that only this peaceful and constant disposition of the soul can produce rational desires, which otherwise would be completely irrational:

Happiness is constituted by the desire for what we presume to be good, since desire is moved excitedly and in flames towards what we regard as good; and since happiness arises and exults, when we have obtained something desirable. Nature, hence, follows whatever is seen to be good, and averts the contraries [...]. When this happens in a constant [constanter] and prudent way, this way of appetite is called *boulesin* by the Stoics, and we call it will (Cicero 1853, 343).

The phrasing *boulesin* is an unmistakable allusion to the distinction Aristotle had made in his *Rhetorics* between *horme*, an irrational desire, and *boulesin*, a rational one. As Aristotle says irrational desires are those of the body, which lack any previous deliberation (like the desire to eat), while rational desires are based on rational persuasion:

[I call things] *in accordance with reason* what people long for on the basis of persuasion; for they desire to see and possess many things after hearing about them and being persuaded [that they are pleasurable] (Kennedy 2007, 88).

Constancy, according to Cicero’s interpretation, can hence be regarded as a certain disposition of the soul, which is a necessary prerequisite of rational decisions.

This is also consistent with Lipsius' case: taking the example of the *De Constantia*, if one has once obtained this disposition, he will feel the urge to flee from a civil war only if he is convinced to do so by rational reasons.

As Margaret Graver has pointed out, emotional judgements in ancient Stoa were – even since the times of Chrysippus – based on the distinction between occurrent and dispositional judgements (Graver 2007, 37). According to this thesis, a dispositional judgement of the *De Constantia* might be reconstructed as follows: “war is an evil thing” (a claim one is disposed to make on a regular basis). An occurrent judgement, on the contrary, might be that “I am staying in a military zone” (a claim one makes only at times). Hence, that conclusion would evidently follow that “I am surrounded by evil things”.

As we have seen it, Lipsius the interlocutor, before being acquainted with Stoical ethics, intended to attack the *occurrent judgement*, since by fleeing, he wanted to achieve not staying any further in a military zone. What we see is that Langius the sage, wants the opposite: namely that his student should revise his dispositional judgement and regard war as something evaluatively insignificant.

For Lipsius the author, the problem is that one grows accustomed to certain dispositions by the repetition of certain impulses which are often independent of rational deliberation. This might be the reason why he defines ungrounded public opinion as the major source of erroneous emotions.

Four major emotions arise from a common source, which keep humans as prisoners, and waste their lives: desire [*desire*] and delight [*gaudium*] on the one hand, while fear [*metus*] and pain [*dolor*] on the other. From these, the former ones refer to, and originate from something presumably good [*opinabile bonum*], while the latter ones refer to, and originate from something presumably bad (Lipsius 1584, 19).

Such affections have to be controlled or blocked by the soul somehow, which is only made possible by a dualistic account of the mind-body problem.

## 5. Lipsian dualism

We have already seen how the body can affect the soul, but the way in which the soul affects the body still has to be revealed.

One of the most important Lipsian innovations – compared to most of the adherents of ancient Stoa – was the introduction of a mind-body dualism. While ancient Stoic thinkers regarded the universe as a systematic organisation of the *pneuma* or *spiritus*, and claimed that this *pneuma* constituted souls or bodies according to its tension (or *tonos*), Lipsius claimed that two qualitatively different substances existed. Lipsius was a dualist who claimed that the human soul had quasi-divine cognitive power, overshadowed only by

the bodily dispositions of humans. Unfortunately, in the *De Constantia*, Lipsius provides us with almost no information regarding either the nature of these substances or their connection, but in his *Physiologia Stoicorum* (1604), he partly makes up for this deficiency. Here, he claims that two principles exist in the world: *God* and *matter* (Lipsius 1610, 69). Soul is regarded by him as an emanation of the divine spirit, which possesses a human body as its vehicle (*receptaculum*). Lipsius never clarifies the exact relationship between the two substances, but still he claims that the two principles cannot be traced back to the same cause. However – as we have already seen – there does exist a kind of mind-body interaction, since the body is capable of affecting the other substance with unnatural affections.

## 5.1 On whether the mind affects the body

still, it is not clear whether the opposite claim – namely the mind’s capacity to affect the body – can be regarded as Lipsius’ stance. Scholars have long investigated whether Lipsius had one single, or two separate accounts of mental causation. Advocates of the first claim say that the ones outlined in the *De Constantia* and the *Physiologia* are practically the same (Sellars 2014, 653–674), while others argue that they are essentially different. In my view, Lipsius had in fact three distinct accounts regarding mental causation: the first two (outlined in the *De Constantia*, the *Politica* and the *Monita et Exemplaria Politica*) could be called *volitionalist* accounts, while the latter (explained in the *Physiologia Stoicorum*) can be regarded as an *occasionalist* one. However, from the perspective of Calvinism the latter work is mostly irrelevant due to the fact that by the time of writing and publishing, Lipsius’s intellectual milieu had changed with his return to the Catholic city of Louvain. Hence, our focus here will be on the first two accounts.

Volitionalism, in the sense in which it is used here, can be defined as the theory that a causal relationship takes place merely between two mental events. A volitionalist account therefore deliberately neglects the question of how mental events affect material or corporeal ones, and observes agency only in terms of mental causes. This is evidently observable in Lipsius’ case, since he seems to allow the body to act on the soul, yet does not attribute the same capacity to mental events:

[...] fate is like a master of ceremonies, which holds the strings during the dance in which the whole world takes part: but in a way that our parts should be able to will and not will [certain things]. But we do not have more power [*vis efficiendi*] than this, since we were given only the opportunity, to be free, to be reluctant and to struggle against God [*reluctari et obniti*]; but power [*vis*] was not given by which we could do that [italics are mine – A. S.] (Lipsius 1584, 20).

Consistently enough, it is only God – elsewhere defined as supreme power (*vis supera*) – whom Lipsius endows with power at all. What the author offers here is a radical departure from the traditional notions of agency: by drawing a distinction between will (*voluntas*)

and power (*vis*), he intends to challenge the very premises of scholastic accounts of the matter. When discussing the will (considered either as something acting freely or moved by God), Thomas Aquinas used the terms *vis activa* or *vis passiva* in defining it (Aquinas 1947, 422). Francisco Suárez in turn regarded a *potentia activa* as a prerequisite of agency, which nonetheless still entailed the power to initiate action in the material world (Suárez 1732, 434). In distancing itself from the above mentioned traditions, the Lipsian distinction was therefore intended to be a revolutionary one.

We have previously seen that Lipsius defined constancy as the motionless weight of the soul (or *robur animi*). Such an account of mental causation might shed some light on why he offered such a definition: since he regarded it as something incapable of producing effects on matter, the only operation it can be capable of is producing effects on itself.

How then should the Lipsian theory of agency be considered? As we have previously seen, people were, according to the *De Constantia*, endowed with the capacity to be reluctant against the divine will, but they were nevertheless incapable of actually revolting against it. This is a claim even Lipsius might have found unsatisfying, since a number of events could not be explained by it: although it is conceivable that my hands do not rise when they are commanded to do so (e.g. they may be physically paralysed or prevented by something), the opposite, namely that my hands raise themselves despite being commanded to stay down, is mostly counter-intuitive. Still, the first Lipsian account seems to allow such instances.

This might be the reason why Lipsius revised his original theory shortly after the publication of the *De Constantia*. In the *Politica seu Civilis Doctrina* (1587) he already provides a more elaborate account of why events during which one's body acts independently of one's mental command cannot occur. Here, he envisions a twofold series of events: (1) physical, occurring independently of one's will; (2) mental, occurring in harmony with the command of God.

Hence, according to this later Lipsian account, no event, either mental or corporeal, can come to pass without divine power, and the chain of events is necessitated so much that

no deliberation or help is enough to avoid or change [the course of these events]. The fatal disposition of divine providence [*diuinae prouidentiae fatalis dispositio*]<sup>7</sup> cannot be subverted or reformed either by prudent deliberation or wise remedies (Lipsius 1604, 26–27).

Since the book is intended to teach rulers how to act prudently (and thus is mostly and rightly read as a mirror for princes), one might wonder how prudence can be considered at all upon such deterministic grounds. According to the Lipsian answer, one must will what

---

7 This formulation makes the passage even more interesting, since earlier, the *De Constantia* Lipsius made considerable efforts to make a sharp distinction between providence and fate (the first one being the deliberation of the divine intellect, the second being the product of his will), while here he seems to be combining the two, or at least gloss over the difference. See: Sellars 2014, 653–674; Smrcz 2016, 177–194.

God has decreed, and if the object of his will corresponds to the very event carried out, his will can be regarded as prudent (although no causal relationship was established between the two). This is further highlighted by the claim, according to which

that higher intellect takes away the reason [*mentem*] of men: he corrupts the deliberations [*consilia*] of those, whose fortune he wants to change (Lipsius 1604, 26–27).

As it can be seen, God can and sometimes even does interfere with mental events and can, according to his will, modify them. In the *Monita et Exempla Politica* (1605), a somewhat later work of his, Lipsius still seems to be adhering to the same position. While considering the question of why one should investigate the content of the divine intellect, he comes to two conclusions: (1) the foreknowledge of upcoming events will render their endurance less painful; (2) by such foreknowledge, prudent men can adjust their deliberations to the facts (either present or to be fulfilled). As he says:

The other benefit of inquiring into [the mysteries of] fate in a careful and modest way, is, that one might see which way that supreme power [*vis illa supera*] draws all events. And we have to adjust ourselves and also our deliberations [*consilia adaptare*] to this [fatal] inclination [*inclinazione*]. All this has a great benefit since, due to the signs preceding upcoming events, or those being attached to them, the fate of the public [*publica fata*], and the ups and downs [of fortune] can not be hidden from the eyes of a prudent man (Lipsius 1605, 21).

Since divine intervention into mental events can harmonize our will with external occurrences, the second, modified account of the Lipsian theory of causation can eliminate some rather counter-intuitive cases, as we have seen above. But, in turn, it can also be seen as an explicit endorsement of determinism, neglecting the question of moral responsibility altogether. It is not our concern here to investigate this question in detail, but if this is the case, then it can probably be assumed that Lipsius intended to adjust the Stoic theory of causation to Christianity according to Calvin's observations, which claimed that God intervened in each and every event of the universe<sup>8</sup>.

## 6. Lipsius: an advocate of apathy?

But does the Lipsian theory of affections also fulfill the Calvinist requirements? As we've seen, according to Lipsian volitionalism, one does not have control over physical events; however, instrumental causal relationships do, in fact, produce effects on each other.

---

<sup>8</sup> See: Calvin 1559, 64; Smrcz 2016, 177–194.

Hence, humans would be capable of avoiding – at least in theory – the occurrence of many such events.

Properly speaking this is what agency means for Lipsius. As we have seen, matter is regarded by Lipsius as something capable of obtaining certain long lasting dispositions, which constitute the origins of the “unnatural” affections of the soul. Since the soul has no power over matter, once these material dispositions have been obtained, there is absolutely no chance to modify them. According to Lipsius since physical events are determined by eternal causes, the soul can never hinder the formation of such bodily modifications.

Lipsius classifies these corporeal dispositions into three major groups (oddly enough, he calls them affections as well, but their definitions reveal that they are absolutely different entities). These dispositions are: *simulatio*, *caritas* and *pietas*. The most important of them is *simulatio*, since this is the constitutive disposition of the other two. Lipsius supposedly took the notion of *simulatio* from Cornelius Tacitus; but while the ancient historian regarded it as a voluntary act of copying other people’s affections, Lipsius takes it as an inherent mechanism of matter, which is completely independent from the will. This disposition, hence, automatically copies the affections of others, because we might easily regard the affections of others as our own.

This account of *simulatio* thus constitutes the foundations of the Lipsian critique of empathy, which is often regarded as an appeal for *apatheia*. Since our bodily dispositions are constituted by *simulatio*, and our bodily dispositions can so easily make us believe that the affections of others in fact belong to us, the consequence will necessarily follow, that when feeling empathy for others, we in fact only empathize with ourselves. This is why Lipsius could claim that *caritas* was identical with *amor*, and thus, neighbourly love was equal with self love.

You are afraid, Lipsius, but why are you so? Because pestilence and decay go hand in hand with war. But whom is this pestilence threatening? Although it is now threatening others, it can one day reach you as well. [...] But just as when a thunder stroke others, even those were frightened, who were only standing nearby: among great devastations which affect everyone, damnation reaches only few people, while fear reaches everyone (Lipsius 1584, 24).

## 7. Natural affections

Therefore, one should avoid affections of bodily origin, since they invoke unnecessary disturbances of the soul. But one should not forget about the fact that even if an affection is not produced by *simulatio*, but by a direct impulse, it also must be borne with a constant soul, since one has no control over external events or over their mental representations. All this means that the rational disposition of the soul would hinder the grievance over all external occurrences. Still, this seeming apathy concerns only “the outermost layer of the soul”.

## 8. Conclusions

As we have seen, the Lipsian account of affections never disqualifies the possibility of the existence of natural affections. Moreover, constancy was defined as a certain disposition of the soul, which enables it to direct its affections reasonably. The Lipsian theory should not, hence, be labelled as one advocating *apatheia*, but as one leaving considerable room for *eupatheia*. Besides arguing on behalf of the above mentioned claim, the aim of this paper was also to highlight some of the possible factors that might have served as a motivation for Lipsius to revise many of the claims – either rightfully or erroneously attributed by his contemporaries – to the ancient Stoics. The major factor of motivation presented in this paper was Calvinism. As I intended to illustrate, Calvin's concerns on some Stoic doctrines were also shared by Lipsius, and it was his endeavor to harmonize Stoicism with Christianity, rather than simply opt for the first one over the other.

## Bibliography

- Aquinas, Thomas. 1947. *Summa Theologica*. New York: Benziger Bros.
- Bouwsma, William J. 1989. *John Calvin – A Sixteenth-Century Portrait*. New York: Oxford University Press.
- Brooke, Christopher. 2012. *Philosophic Pride: Stoicism and Political Thought from Lipsius to Rousseau*. Princeton: Princeton University Press. 12–37.
- Calvin, Johannes. 1863. Commentarius in Acta Apostolorum. In: Cunitz, Edouard – Baum, Johann-Wilhelm – Reuss, Eduard Wilhelm Eugen (eds.): *Joannis Calvini opera quae supersunt omnia*. Brunsvigae: C.A. Schwetschke. Tom. 48. 175–176.
- Calvin, Johannes. 1863. Commentarius in Epistola Petri. In: Cunitz – Baum – Reuss: *Joannis Calvini opera quae supersunt omnia*. Tom. 55.
- Calvin, Johannes. 1559. *Institutio Religionis Christianae libri quatuor*. Geneve.
- Cicero, Marcus Tullius. 1853. *Tusculanarum disputationum libri quinque*. Ed. Raphaël Kühner. Ienae [Jena]: Typis et sumptibus Friderici Frommann. 333–343.
- Gosling, Justin. 1987. The Stoics and ἀκρασία. *Apeiron* 20 (2): 179–202.
- Graver, Margaret R. 2007. *Stoicism and Emotion*. Chicago: The University of Chicago Press.
- Hugo, André Malan. 1969. Introduction. In: Calvin, Jean – Battles, Ford Lewis – Hugo, André Malan: *Calvin's Commentary on Seneca's De Clementia*. Leiden: Brill. 3–63.
- Kennedy, George A. (trans.) 2007. Aristotle: *On Rhetoric: A Theory of Civic Discourse*. Oxford: Oxford University Press.
- Leibniz, Gottfried Wilhelm. 1999. Sentiments de Socrate Opposés aux Nouveaux Stoïciens et Epicuriens. In: Leibniz, Gottfried Wilhelm: *Sämtliche Schriften und Briefe*. Reihe 6. Philosophische Schriften. Bd. 4. Berlin: Akademie Verlag. 1384–1388.
- Lipsius, Justus. 1605. *Monita et Exempla Politica*. Mons Divi Hilarii: Petrus Chevalier.
- Lipsius, Justus. 1610. *Physiologia Stoicorum libri tres*. Antwerp: Johannes Moretus.
- Lipsius, Justus. 1584. *De Constantia libri duo*. Antwerp: Christophorus Plantinus.

- Lipsius, Justus. 1604. *Politicorum sive Civilis Doctrinae libri sex: Qui ad Principatum Maxime Spectant*. Licha: Konrad Neben.
- Oestreich, Gerhard. 1982. *Neostoicism and the Early Modern State*. Cambridge: Cambridge University Press.
- Sellars, John. Neo-Stoicism. *Internet Encyclopedia of Philosophy*. Accessed 22 July, 2017. <http://www.iep.utm.edu/neostoi/#H1>
- Sellars, John. 2014. Stoic Fate in Justus Lipsius's *De Constantia and Physiologia Stoicorum*. *Journal of the History of Philosophy* 52 (4): 653–674.
- Sellars, John. 2017. Stoicism. In: Sgarbi, Marco (ed.): *Encyclopedia of Renaissance Philosophy*. Dordrecht: Springer. 1–15.
- Smrcz, Adam. 2016. The Problem of “Stoic Fate” – Or whether Herbert of Cherbury was a Lipsian. *Kultura i Edukacja* 2 (112): 177–193.
- Suárez, Francisco. 1732. Liber Primus de Prologomenis. In: Suárez, Francisco – Noël, Francisco (eds.): *Summa seu Compendium. Pars prima*. Coloniae: Sumptibus Fratrum de Tournes.
- Taylor, Charles. 2007. *A Secular Age*. Cambridge: Harvard University Press.

## II

*Maximilian Kiener*

# The Compulsion To Believe Something: On the Affectivity of Indubitable, Clear and Distinct Perceptions in Descartes

## Introduction

Descartes' method in the *Meditations* opens with a clear statement about doubt and truth: if there is only the slightest chance that one could doubt something, one is not entitled to assume its truth. Unfortunately, the inversion of this conditional – which concerns the relation between the absence of doubt (i.e. indubitability) and truth – is not explained in the same detail. As a result, both the nature of indubitability (whether indubitability is a psychological or epistemic notion) as well as whether Descartes embraced an account which reduces truth to terms of indubitability has remained unclear. These aspects become particularly crucial with regard to the so-called Cartesian Circle as various interpretations defend Descartes by presenting such a reductive account of truth, either in psychological or epistemic terms.

Looking at the debate on the Cartesian Circle, I will make two claims: firstly, the extent to which Descartes' notion of indubitability can be considered a 'psychological' notion is considerably limited, and, secondly, Descartes' notion of truth cannot be reduced to indubitability in any sense which either assumes a psychological quality of indubitability or denies that 'truth' consists in the correspondence between thought and reality. I will argue for these claims in Part Two of this paper, after having presented the accusation of circularity as well as a psychological interpretation of Descartes in Part One. Tackling a psychological reading of Descartes' notion of indubitability is particularly interesting when it comes to the topic of affectivity. Most places where affectivity is discussed in early modern philosophy centre on emotions and related capacities of the mind. The current debate follows this emphasis and gives short shrift to the question as to what role epistemic aspects such as knowledge, truth, and their relation to the mind play. This essay aims to contribute to the debate on affectivity by investigating affectivity from a genuinely epistemic angle.

I want to specify what kind of affectivity the Cartesian ‘compulsion’ to believe something displays, namely how the irresistible appeal of indubitable, clear and distinct perceptions on the mind can be captured. Moreover, I also aim to analyse how such an appeal relates to an objective notion of truth. These aspects will then add greater breadth to the debate and also help situate another aspect of Descartes’s philosophy with greater precision.

## PART ONE: The Accusation of Circularity

### Arnauld’s Objection and Descartes’ Reply

From the Third Meditation onwards, Descartes’ statements are puzzling to his readers. On the one hand, he claims that “certainty and truth of all knowledge [*scientiae*] depends uniquely on my awareness [*cognitione*] of the true God, to such an extent that I was incapable of perfect knowledge [*perfecte scire*] about anything else until I became aware of him” (CSM II 49 / AT VII 71). However, on the other hand, he still relies on a “clear and distinct idea of (...) God” (CSM II 37 / AT VII 53) as well as on the assumption that “it is manifest by the natural light that there must be at least as much reality in the efficient and total cause as in the effect of that cause” (CSM II 28 / AT VII 40) in order to demonstrate God’s existence. Therefore, it is not surprising that Arnauld, one of Descartes’ contemporaries, highlighted a problem of circularity (see AT VII 124–125 / CSM II 88–89). Formally construed, if P is the proposition “What one clearly and distinctly perceives is true”, and Q is the proposition “God exists”, Descartes’ argument (extracted from the above quotations) seems to embrace both claims:

- (1)  $K(P) \rightarrow K(Q)$  One knows P only if one first knows Q.
- (2)  $K(Q) \rightarrow K(P)$  One knows Q only if one first knows P.

Hence, Descartes’ aim to justify P and Q seems to fail, since (1) and (2) are caught in a circle and therefore cannot justify either P or Q.

Weighty as this accusation appears, Descartes himself was nevertheless quite unimpressed and replied:

when I said that we can *know* [*scire*] nothing for certain until we are aware that God exists, I expressly declared that I was speaking only of *knowledge* [*scientia*] of those conclusions which can be recalled when we are no longer attending to the arguments by means of which we deduced them. Now *awareness* [*notitia*] of first principles is not normally called ‘knowledge’ [*scientia*] by dialecticians. When someone says “I am thinking, therefore I am, or I exist”, he does not deduce existence from thought by means of a syllogism, but *recognizes it as something self-evident* by

*a simple intuition of the mind* [*per se notam simplici mentis intuitu agnoscit*] (CSM II 100 / AT VII 140. Emphasis added).

I made a distinction between *what we in fact perceive clearly* and *what we remember having perceived clearly on a previous occasion*. To begin with, we are sure that God exists because we attend to the arguments which prove this; but subsequently it is enough for us to remember that we perceived something clearly in order for us to be certain that it is true. This would not be sufficient if we did not *know* [*sciremus*] that God exists and is not a deceiver (CSM II 171 / AT VII 246. Emphasis added).

Descartes admits that before the demonstration of God's existence he does not possess any "knowledge", that which he calls "scientia", or denotes with the verb "scire". However, what he can rely on is what he calls "notitia", "mentis intuitu[s]" and later on several times – which becomes the central term – "cogitatio". Descartes considers *scientia* to be a body of knowledge, as something which is stable and exists over time. *Cogitatio* in contrast means single flashes of perceptions without temporal extension in any way. Rather, cogitatio is restricted to the instantaneous moment of being thought or perceived. As Descartes said, it lasts "quandiu cogito" (AT VII 27 / CSM II 18).<sup>1</sup>

Descartes' reply contains two crucial distinctions: Firstly, Descartes distinguishes between *present* clear and distinct perceptions and *past* clear and distinct perceptions (respectively between *single* present and clear perceptions, and the *general link* between clear and distinct perceptions and truth).<sup>2</sup> Secondly, he argues that different epistemic relations apply to present and past perceptions, namely cogitatio and scientia.

On the basis of these distinctions, Descartes claims that in his demonstration of God's existence he employed only present clear and distinct perceptions which are encountered through cogitatio and are not devoid of justificatory potential on their own. They are "self-evident", and as such are reliable premises in his demonstration.

It is important to note that Descartes does not hold that there is a certain class of (present) perceptions which can always be recognised as true while other present perceptions cannot. He only claims that *at the moment when* he perceives something clearly and distinctly (without further qualification of what this "something" is about), he can be instantly assured of its truth.<sup>3</sup> However, it would be false to say that the Cartesian Meditator knows that "whatever he presently clearly and distinctly perceives is true" from the outset.

1 See also AT VII 25 (CSM II 17) where Descartes uses the word "quoties".

2 For the purpose of this paper I do not need to decide whether the so-called *Memory Defence* which Doney (1955) argued for, or whether the so-called *General Link Defence*, which Frankfurt (1970), Kenny (1970), Van Cleve (1979) argued for, is correct.

3 Nelson (1964), Broughton (1984), and Cottingham (1998) nevertheless assume that there is a class of demon-proof perceptions. However, only according to my *temporal* interpretation, which is similarly advocated by Gerwitz (1941) and Della Rocca (2005), can one make sense of Descartes saying both that *all* perceptions are subject to doubt (AT VII 21 / CSM II 21), and that nevertheless *some* perceptions are not subject to doubt (see Descartes' aforementioned reply). Whether something is subject to doubt depends on whether it is before one's mind (as a clear and distinct perception) at a given moment.

He does not claim any such second-order knowledge about the truth of present perceptions like K(P). Nevertheless, there is still an ongoing debate on luminosity here, namely, whether the Cartesian Meditator has to be at least aware that he is currently perceiving something clearly and distinctly, or whether just the act of perceiving something clearly and distinctly is sufficient to be aware of its truth (See Van Cleve 1979, 66–74).

A few further clarifying comments on Descartes' notion of "clear and distinct perception" may be required.

Firstly, Descartes seems to embrace a strong kind of internalism. Defining a clear and distinct perception in the *Principles*, he says that a perception is clear when it is "present and accessible to the attentive mind" (*Principles*, Part I, 45. CSM I 207 / AT VIII A 22) and distinct when "it is so sharply separated from all other perceptions that it contains within itself only what is clear" (*Principles*, Part I, 45. CSM I 208 / AT VIII A 22). Thus, the grounds for both clearness and distinctness arise *only from within* the Meditator's perspective. The question as to whether this account presents a genuinely *epistemic* internalism remains, however, to be seen.

Secondly, a clear and distinct perception is an intellectual presentation of something being so-and-so without being identical to believing. Not only does Descartes repeatedly talk about cases in which he used to believe something but nevertheless did not clearly and distinctly perceive it, a clear and distinct perception is also considered a state of mind which – if it occurs – is prior to believing. This point was made clear in the *Discourse* when Descartes formulated his goal in the following way: "to include nothing more in my judgement than what presented itself to my mind *so* clearly and distinctly *that* I had no occasion to doubt it" (CSM I 120 / AT VI 18).<sup>4</sup> Saying that something "presented itself (...) *so* (...) *that*" indicates the logical order. Adopting a belief is subsequent to the evaluation of clearness and distinctness, no matter how quickly the transition may happen and no matter how compelled one feels to adopt the belief when facing a clear and distinct perception. This very transition from clearly and distinctly perceiving to believing will be discussed in the next section.

Lastly, it is worth being aware that Descartes was not clear on whether propositions<sup>5</sup> or ideas (for instance AT VII 53 / CSM II 37) are the object of clear and distinct perceptions, nor was he clear on whether it is the perception (*Principles* Part I, 30, 45–46: AT VIII A 17, 22 / CSM I 203, 207–208) that is clear and distinct or the proposition/idea (AT VII 46 / CSM II 31–32). For the rest of the paper, I will not draw sharp distinctions between these different possibilities but rather assume that the idea of God and respectively the proposition "God exists" are such that one can clearly and distinctly perceive it.

To come back to Arnauld's objection: if Descartes is entitled to claim that a present clear and distinct perception is justified in its own right, he can successfully refute the accusation of circularity. The proof of God's existence is not required for the validity of cogi-

4 The original reads as follows: "si clairement et si distinctement (...) que je n'eusse aucune occasion de le mettre en doute".

5 For instance AT VII 35 (CSM II 25): "sum certus *me esse rem cogitantem*". I take the grammatical structure to indicate a proposition here. See also AT VI 32–33 / CSM I 127.

tations that are used as premises. The proof is only required for facilitating the transition from “cogitatio” to “scientia” and thereby creating a stable body of knowledge out of these single and disconnected flashes of perceptions. However, Descartes’ stance is unsatisfactory in the sense that it significantly begs the unanswered question: *how* can one be assured of the truth of present clear and distinct perceptions in the first place? This is where the discussion on the Cartesian Circle reaches its peak.

## A Psychological Interpretation of Descartes

The psychological interpretation gives a twist to everything that Descartes said in reply to the accusation of circularity. Rubín and others start by claiming that Descartes was, after all, not in the business of presenting a deductive argument beginning from premises and inferring a general theorem. Instead, Descartes’ general rule that “[a]ll which I clearly and distinctly perceive is true” is supposed to “record the observation that (as a matter of *psychological facts*) he is unable to doubt that which he has clearly and distinctly apprehended” (Rubín 1977, 197. Emphasis added). Descartes’ ultimate aim is to arrive at a state in which it is “*psychologically* impossible to doubt beliefs based on clear and distinct perceptions” (See Loeb 1992, 201. See also Larmore 1984, 61). There is consensus among proponents of this interpretation that the way in which we are *psychologically* compelled contrasts sharply with the way in which we abide by an *epistemic* and *rational* norm. It is not about what one (normatively) ought to do but rather just what one (descriptively) does. Rubín identifies the “psychological” effect as a “causal” effect and contrasts it with that which is “evidential” (See Rubín 1977, 207).

Descartes’ argument is seen to depart from a situation in which some present clear and distinct perceptions, such as the “*cogito, existo*” are psychologically indubitable in such a way. We simply cannot withhold our assent to these clear and distinct perceptions *while* we are attending to them. Rubín takes Descartes to be talking about this specific psychological compulsion when he refers to present clear and distinct perceptions and says “it is *not possible* for me not to believe” (AT VII. 70 / CSM II 48; Rubín 1977, 199. Emphasis in Rubín), “we *can never doubt* them” (AT VII 145 / CSMK III 103–104; the same again in AT VII 146 / CSM II 104; Rubín 1977, 198), “we *can never think* them without believing them to be true” (AT VII 146 / CSM II 104; Rubín 1977, 198), “we *cannot possibly doubt* their truth” (Principia, 43; Rubín 1977, 200), they “have the power of *persuading* [us] entirely” (AT VII 68 / CSM II 47; Rubín 1977, 200; Rubín’s translations).

However, as soon as one stops attending to a clear and distinct perception, one can easily entertain doubts about the truth of these perceptions. One can think of the possibility of having a poorly designed mind, or being systematically deceived by an evil demon, which could bring about (from this perspective of not currently attending to a clear and distinct perception), that even what seems most evident may turn out to be false.

What Descartes therefore needs to do in this interpretation is to render these doubts psychologically impossible as well. He can do so by abandoning the grounds on which we

have so far been able to doubt the truth of beliefs based on clear and distinct perceptions. These grounds consisted of our ability to assume that we have a systematically flawed mind, or that we are being systematically deceived. Confronted with Descartes' demonstration of God's existence and veracity, we shall now be deprived of the ability to embrace such possibilities. Rubin says:

the proposition 'There is a veracious God' has an epistemological property which other propositions (...) lack: If our understanding compels our will to affirm it, then the causal mechanism of our minds makes it impossible for us to have doubts about the truth of the propositions "All which I have clearly and distinctly apprehended is true" (Rubin 1977, 206).

Therefore, the subjective compulsion which the Meditator had with regard to believing that *single present* clear and distinct perceptions are true, now develops into a more general subjective compulsion which also comprises believing that *all* clear and distinct perceptions are true (See also Larmore 1984, 71).

This interpretation finally leads to a special take on the notion of truth. According to Rubin, Descartes takes propositions to be *true* in the sense of being psychologically indubitable. Descartes considers something true if, "[b]ecause of the design of his mind, he cannot prevent himself from admitting them all as true and certain" (Rubin 1977, 206). Thus, the truth of propositions is ascertained by means of a closed question argument. By rendering the Cartesian Meditator psychologically unable to entertain worries about our perceptions failing to correspond to reality (after the demonstration of God's existence), no separate question about truth itself is conceivable anymore. According to Rubin, this is what Descartes means when he says:

What is it to us [who have understood the argument of the *Meditations*] though perchance some one supposes that, of the truth we are so firmly persuaded [*firmiter sumus persuasi*], appears false to God or to an Angel, and hence is, absolutely speaking false? What heed do we pay that absolute falsity [*falsitatem absolutam*] when we by no means believe that it exists *or even suspect* its existence [*nec vel minimum suspicemur*]? (AT VII 145 / CSM II 103)<sup>6</sup>

In summary, the psychological interpretation accounted for what Descartes himself stressed: the distinction between present and past perceptions as well as the attempt to validate only the latter and thereby establish a general link between truth and clear and distinct perceptions. However, in order to defend Descartes this interpretation reduces his aim to presenting a psychological observation as opposed to presenting an argument. Since any accusation of circularity is applicable only to *arguments* and not to *observations*, Descartes' critics went astray.

---

<sup>6</sup> The translation and emphasis is Rubin's. Latin phrases were added. See Rubin 1977, 205. See also Larmore 1984, 71.

# PART TWO: Critique of the Psychological Reading

## CLAIM ONE: The extent to which Descartes' notion of indubitability can be considered a 'psychological' notion is considerably limited

The psychological interpretation connected the notion of indubitability with Descartes' comments on a subjective compulsion to believe something. Therefore, in order to evaluate to what extent indubitability is and could be considered a psychological notion, as opposed to an "epistemic" or "rational" notion, we need to scrutinise exactly what "psychological" means in this context. It seems that there are two aspects present in the psychological interpretation.

Firstly, a subjective compulsion may be called "psychological" and not "epistemic" if its way of affecting us is not necessarily connected to our awareness of there being good reasons and evidence. Many of our standard psychological reactions as human beings seem in fact not to depend on the awareness of reasons in any way. For instance, at a very early stage, babies start to smile as a reaction to the smiling of their parents and other people. This standard reaction, which adults also have, is not linked to the awareness of reasons in any way. It is worth noting that Rubin's account at least partly suggests such a reading. He focuses on the claim that "Descartes conceived of belief and doubt in *causal* terms" (Rubin 1977, 197. Emphasis added) and goes on to set out a contrast between a causal effect and something that is "evidential" (Rubin 1977, 207), as well as between "psychological" and "logical" (Rubin 1977, 205–206).

Secondly, we may also consider indubitability as "psychological" not because the compelling force of belief operates apart from any awareness of reasons but rather because it merely *happens to us* without us being able to exert control over it. Larmore, Loeb, and also Rubin embraced this aspect of "psychological" (See Larmore 1984, 61; Loeb 1992, 202). They admit that "our *understandings* compel our wills to affirm" (Rubin 1977, 199, 202, 203. Emphasis added) and that it is "*evidence* which, when attended to, compels assent" (Larmore 1984, 64. Emphasis added). Nevertheless, one is still "compelled" and cannot exert control over this compulsion. *Therefore*, indubitability is a "psychological" notion. A genuinely *epistemic* understanding of indubitability would require having greater control over the adoption of beliefs.

With the aforementioned first reading of "psychological" in mind, let us now focus on how Descartes' account relates reasons or evidence on the one hand to the compulsive force of beliefs on the other hand. Descartes' discussion is divided into compulsion concerning:

- a) the immediate adoption of a belief, and
- b) the abandonment of a belief.

Concerning the immediate adoption of a belief, in the Fourth Meditation, Descartes asks whether he can be considered free if he sometimes cannot but assent to a proposition. He then goes on to claim that the reason why he is drawn to assent is because he *recognises something as so evidently true* [*quia rationem veri & boni in ea evidenter intellego* (AT VII 57–58 / CSM II 39–40)]. The terms *ratio*, *verus*, *evidenter*, and also *intellegere* are not merely “psychological” notions in this first and narrower sense.

Moreover, in the Fifth Meditation, Descartes said: “even if I am no longer attending to the *arguments* [*rationes*] which led me to judge that this is true, as long as I remember that I clearly and distinctly perceived it, there are no *counter-arguments* [*ratio contraria*] which can be adduced to *make me doubt* it [*me ad dubitandum impellat*], but on the contrary I have true and certain knowledge of it” (CSM II 48 / AT VII 70). Here, Descartes explains the situation in which he could not but assent with reference to “rationes”.

However, even beyond present clear and distinct perceptions, reasons play a crucial role. Concerning the abandonment of a belief, let us re-examine the quotation from the *Fifth Meditation*. When Descartes talks about what can make him doubt something [*me ad dubitandum impellat*] towards the second part of the sentence, and thereby attempts to abandon a dubitable belief, he refers to “ratio” as well. Moreover, this connection between being compelled to doubt and reasons is also confirmed in his reply to the Seventh Set of Objections when Descartes says that there are “reasons which are strong enough to compel us to doubt” [*sunt satis validae rationes ad cogendum nos ut dubitemus* (AT VII 473 / CSM II 318–319)]. Therefore, subjective compulsion concerning both a) the adoption of belief and b) the abandonment of a belief, operates clearly on the basis of *reasons*.<sup>7</sup> Both dubitability and indubitability require the presence of reasons.

What Descartes claims here is very close to what is nowadays called “transparency” of belief (See Edgley 1969, 90. See also Evans 1982, 225–226; Moran 2011, 60). If one asks oneself whether to believe *p*, this question reduces to, or is *transparent* to, the question of whether *p* is true, in other words: whether there are sufficient reasons for the assumption of this claim’s truth. And once we have answered the question of *p*’s truth in the affirmative, we cannot but assent to *p* – *mutatis mutandis* with the abandonment of belief. Hence, we can indeed make philosophical sense of this interpretation. A compulsion to believe can be explained with reference to reasons and considerations of truth.

Although these passages rule out the first reading of “psychological”, the second reading is still possible. Such a reading allows for reasons but requires that something merely happens to us without us being able to exert any control over it. However, such an account is also at odds with Descartes’ remarks. He does not think that compelled assent is something that just happens to us. Those cases in which the Cartesian Meditator is most compelled to assent – because something is shown to him/her by the natural light – are explicitly contrasted with a compulsion by an *impetus* (AT VII 38 / CSM II 26–27), some

7 There is a great variety of further passages in which Descartes requires one to have a *reason* to doubt. See also: *Meditations*: AT VII 18, 21, 36, 70 (CSM II 12, 14, 25, 48). *Objections and Replies*: AT VII, 171, 460, 473–474 (CSM II 121, 308–309, 318–319); IXa 204 / CSM II 270. *Discourse*: AT VI 29, 38–39 (CSM I 125, 130).

kind of inclination or force which simply happens to someone. Descartes argues that the more he feels compelled to believe something which is being shown by the natural light, the more *freely* he *chooses* to believe [“quo magis in unam propendeo (...) tanto *liberius* illam *eligo*” (AT VII 57–58 / CSM II 39–40. Emphasis added)]. Descartes talks about his *own choice* and not about himself as being the mere target of some force.

However, this might only be a partial rejection of the second reading of “psychological”. Larmore for instance allows for freedom in the sense just described, but argues that for Descartes’ account to be “nonpsychologistic” (Larmore 1984, 67) the Cartesian Meditator would need to have another kind of freedom or control over beliefs, namely one that stems from the ability to be *indifferent* to considerations of truth. Experiencing subjective compulsion together with the aspect of a choice [*eligo*] is sufficient to render Descartes’ account “psychologistic”. It is worth noting, however, that Larmore further qualifies this interpretation and stresses that it does not imply that Descartes

shared the other doctrines of those psychologistic philosophers whom Frege attacked at the end of the 19<sup>th</sup> century. For example, Descartes did not believe that the domain of mathematics consisted in psychological states or acts, but attributed to it much the same “objectivity” which Frege did. His appeal to psychological fact seems to have been restricted to the relation between evidence and assent (Larmore 1984, 61).

I agree with such a narrow understanding of “psychologistic” if it only comprises “transparency” and the Cartesian *eligo*. I disagree, however, with every attempt which aims to expand the “psychological” aspect beyond this interpretation. Descartes’ account is certainly not psychological in the two senses I described above, attributed to Rubin. Indubitability and being compelled to believe something, both display clear epistemic aspects in the sense of being dependent on reasons and evidence – hence epistemically normative – as well as in the sense of stemming from a free doxastic choice (*eligo*) rather than from an event which is merely causally forced upon us.

Having analysed the notion of indubitability, I will now investigate Descartes’ notion of truth and its relation to indubitability.

## **CLAIM TWO: Truth Does Not Reduce to Indubitability**

The arguments from *Claim One* do not address the following two positions concerning Descartes’ notion of truth. The first is the position held by the psychological reading according to which there is nothing but indubitability at the level of cogitationes, and that Descartes cannot *know the truth* of perceptions at this level. The truth of propositions can only be known after the demonstration of God’s existence and is to be understood in terms of indubitability. The second is the position offered by Frankfurt. According to him, Descartes’ aim consists in abandoning not the psychological but rather the rational, or

epistemic grounds (viz. reasons) for doubt. However, when Descartes then arrives at a state in which he can no longer doubt the truth on rational grounds, Frankfurt does not think Descartes achieves a notion of truth itself in a strong sense. Descartes arrived at a notion of truth *not* in the sense of correspondence between thought and reality, but rather truth in the sense of coherence among the deliverance of reason. If something is indubitable, it is only *true* to the extent that the epistemic grounds, which are the deliverances of reason itself, cannot be turned against themselves. Inner coherence in this sense is still compatible with being false *absolutely* speaking (See Frankfurt 1970, 170–180). I mention these views because they both claim that indubitability is the primary notion, and that truth has to be understood in terms of indubitability.

In what follows I want to elaborate on three aspects in Descartes which provide evidence that he clearly opposes what Frankfurt and the psychological reading propose as Descartes' account of truth and indubitability. I want to stress that Descartes is a non-reductionist about truth, which means that Descartes holds that truth cannot be reduced to indubitability (insofar indubitable, clear and distinct perceptions could still be false).

## 1. Correspondence Theory of Truth

Descartes explicitly thought of truth in the sense of correspondence to reality. In a letter to Mersenne, Descartes writes that “the word truth, in the strict sense, denotes the *conformity of thought with its object*” (“ce mot vérité, en sa propre signification, dénote la *conformité de la pensée avec l’objet*”) (AT II 597 / CSMK III 139. Emphasis added. Translation by Kenny in Descartes 1970, 66). And he adds that because this notion of truth is “very simple and naturally known” (AT II 597 / CSMK III 139), defining and redefining truth would only “obscure (...) and cause confusion” (AT II 597 / CSMK III 139). In the *Principles* Descartes says: “Given, then, that our efforts are directed solely on the search of truth, our initial doubts will be about the existence of the *objects* of sense-perception and imagination” (Principles I 4. CSM I 193–194. AT IXb 26. Emphasis added).<sup>8</sup>

## 2. Lumen Naturale

Moreover, the extent to which Descartes employs this notion of truth is striking, particularly in a group of passages in which he focuses on what he calls the “natural light” [*lumen naturale*] (AT VII 15, 38, 40, 42, 44, 47, 49, 52, 60, 82, 107, 108, 119, 134–136, 148, 161, 238, 240–241 / CSM II 11, 26, 27, 29, 30, 32, 33, 35, 41, 56, 77, 78, 84, 96–97, 105, 113, 166, 167–168). Most of these passages appear in the *Third Meditation*, which is when Descartes presents the demonstration of God’s existence. Although Descartes’

---

<sup>8</sup> See also AT VII 37 (CSM II 25–26). Here, Descartes says error would not be possible if ideas did not refer to anything else beyond themselves.

explanations of what exactly the natural light is are very sparse, we are told that if there is anything like a criterion [*règle*] of truth, it is the distinctness of perceptions yielded by the natural light (Letter to Mersenne [16 October 1639], AT II 597–598 / CSMK III 139). Thus, we are dealing with an extraordinary faculty of recognition. The natural light is explicitly employed with regard to *cogitationes* and *notitiae*,<sup>9</sup> among them the premise of the demonstration of God’s existence, that a cause has to have at least as much reality as its effect. Thus, if we accept Descartes’ explanation of truth found in the letter to Mersenne, we can assume that Descartes thought that at least some perceptions *cannot* be false anymore once the Meditator, resorting to the *lumen naturale*, recognises them as indubitable.

### 3. Descartes On “Scire” in the Conversation with Burman

The same kind of stronger truth aspirations appear in Descartes’ conversation with Burman. Here, Descartes said that “he knows [*scit*] that he is not deceived with regard to them [the premises of the proof of God’s existence], because he is actually paying attention to them” (CSMK III 334 / AT V 148). Later in the same conversation, he adds: “If we did not know that all truth has its origin in God, then however clear our ideas were, we would not know [*sciremus*] that they were true, or that we were not mistaken – I mean, of course, when we were not paying attention to them” (CSMK III 353 / AT V 178). What renders these passages striking is Descartes’ repeated use of the verb *scire*. In the first passage, he mentions it directly with regard to how we *know* that we are not deceived through present and clear perceptions. In the second passage, we may also conclude – through a reasoning *e contrario* – that we can *know* the *truth* of ideas which are present on our minds at a given moment. One might feel inclined to think of this passage as almost a slip of the tongue, because *scit* and *sciremus* do not fit when Descartes actually talks about *cogitationes*, for it was Descartes himself who distinguished *cogitatio* from *scientia*. However, it still seems that the conversation with Burman confirms what the passages on the natural light together with the letter to Mersenne stated. Descartes is concerned with truth to a considerable extent, and these truth aspirations apply at the level of *cogitationes*. To use *scire* is therefore not to confuse terminology but rather perhaps to (over-)stress the fact that one can recognise the truth of some present, clear and distinct perceptions, even if this “recognition” or “knowledge” does not ultimately qualify as part of *scientia* in the strict sense.

What these passages then show is that Descartes is not only interested in a correspondent notion of truth, but that he also attempts to employ truth at the level of *cogitationes*. Descartes assumes that the employment of the natural light can yield certain clear and distinct perceptions which are known as true beyond a state of indubitability. Therefore, these passages oppose the claims of the psychological reading and suggest that *being justifi-*

9 AT VII 15 (CSM II 11): Descartes talks about “*luminis naturalis ope cognitas veritates*” (Emphasis added). In AT VII 38 (CSM II 26), he connects the natural light to the “*cogito, existo*”. In: AT VII 44 (CSM II 30). Descartes uses the term “*notum*”. In: AT VII 40 (CSM II 27–28) he employs the natural light to the claim that a cause has to have at least as much reality as its effect.

*ably assured of the truth* of some perceptions, and to the extent one follows Descartes' use of *scire* in the conversation with Burman, also *knowing* the truth, becomes possible even *before* the demonstration of God's existence, namely, at the level of *cogitationes*. Hence, when Descartes is assured of truth at this level, assurance of truth can only mean (psychological) indubitability, as well as other potential reductionist accounts, like Frankfurt, which deny a "correspondence" account of truth.

## Conclusion

I aimed to make two claims in this essay. I wanted to argue, firstly, that there are limitations in understanding Descartes' notion of indubitability in psychological terms and, secondly, that Descartes' notion of truth cannot be reduced to indubitability in any way proposed by the interpretations I have discussed. Although an interpretation of the notion of truth and indubitability in Descartes may be interesting in its own right, it gives us a particularly good grasp on the discussion of the Cartesian Circle. If my interpretation is correct, Descartes claims that even before the demonstration of God's existence, one is in a position to recognise and so know the *truth* of present, clear and distinct perceptions. As a result, any interpretation which limits Descartes' aim to anything weaker than this must fail on the basis of the arguments and text passages I have presented. The real challenge then is to further explain how such a recognition or knowledge of truth is possible when Descartes himself did not give anything close to a comprehensive account of it. Hence, one must either present an explanation which does not contradict Descartes' writing, or deny that Descartes has the means to pursue successfully this kind of internalist project. The real challenge for Descartes in this debate is not circularity, but over-ambition. These results are also particularly interesting with regard to affectivity. There is an epistemic kind of affectivity in Descartes, consisting of a compulsion to believe something created by the "transparency of belief", but still being compatible with the aspect of a particular kind of epistemic choice [*eligo*]. This kind of affectivity is genuinely epistemic to the extent that it does not go beyond "transparency of belief" in spelling out the particular impact it exerts on the person being subject to it. No additional emotional strings are attached. This form of affectivity is also special in terms of how Descartes connects the compulsion to believe something with freedom in believing. Although he may not be able to resist believing something, or be able so simply to choose to believe something, the capacity of appreciating the validity of what is to be believed grounds what in contemporary terminology may be called doxastic freedom. Descartes' account is very close to what D. Owen today calls an appreciation-based account of doxastic freedom (Owens 2000). Descartes here offers a way of thinking about doxastic freedom, compulsion and affectivity that goes contrary to the majority in the debate on doxastic freedom which followed B. Williams' influential work (Williams 1973, 136–151) and almost exclusively focuses on a will-based account of doxastic freedom. Nevertheless, Descartes' influence can be found in the contemporary debate

at places where the idea that our doxastic freedom may be grounded in our ability to appreciate evidence has received interest. Hieronymi (2006, 2008), Steup (1982, 2000), Ryan (2003), Owens (2000), Shah (2002, 2003), and McHugh (2014) all point to our ability to appreciate evidence as the reason for some kind of doxastic freedom. Hence, investigating Descartes not only contributes to the breadth of the debate on affectivity in early modern philosophy, but may also facilitate a deeper understanding of accounts in contemporary epistemological debates.

## References

- Broughton, Janet. 1984. Skepticism and the Cartesian Circle. *Canadian Journal of Philosophy* 14: 593–615.
- Cottingham, John. 1998. *Descartes*. Oxford: Oxford University Press.
- Della Rocca, Michael. 2005. Descartes, the Cartesian Circle, and Epistemology Without God. *Philosophy and Phenomenological Research* 70 (1): 1–33.
- Descartes, René. 1970. *Descartes. Philosophical Letters*. Translated and edited by Anthony Kenny. Oxford: Clarendon Press.
- Descartes, René. 1984. *The Philosophical Writings of Descartes. Volume II*. Translated by John Cottingham, Robert Stoothoff, Dugald Murdoch. Cambridge: Cambridge University Press.
- Descartes, René. 1985. *The Philosophical Writings of Descartes. Volume I*. Translated by John Cottingham, Robert Stoothoff, Dugald Murdoch. Cambridge: Cambridge University Press.
- Descartes, René. 1991. *The Philosophical Writings of Descartes. Volume III*. Translated by John Cottingham, Robert Stoothoff, Dugald Murdoch, Anthony Kenny. Cambridge: Cambridge University Press.
- Doney, Willis. 1955. The Cartesian Circle. *Journal of the History of Ideas* 16 (3): 324–338.
- Edgley, Roy. 1969. *Reason in Theory and Practice*. London: Hutchinson.
- Evans, Gareth. 1982. *The Varieties of Reference*. Oxford: Oxford University Press.
- Frankfurt, Harry. 1970. *Demons, Dreamers, and Madmen: The Defense of Reason in Descartes's Meditations*. New York – London: Garland.
- Gewirtz, Alan. 1941. The Cartesian Circle. *The Philosophical Review* 50 (4): 368–395.
- Hieronymi, Pamela. 2006. Controlling Attitudes. *Pacific Philosophical Quarterly* 87: 45–74.
- Hieronymi, Pamela. 2008. Responsibility for Believing. *Synthes* 161 (3): 357–373.
- Kenny, Anthony. 1970. The Cartesian Circle and the Eternal Truths. *The Journal of Philosophy* 67 (19): 685–700.
- Larmore, Charles. 1984. Descartes' Psychologistic Theory of Assent. *History of Philosophy Quarterly* 1 (1): 61–74.
- Loeb, Louis. 1992. The Cartesian Circle. In: Cottingham, J. (ed.): *The Cambridge Companion to Descartes*. Cambridge: Cambridge University Press. 200–235.
- McHugh, Conor. 2014. Exercising Doxastic Freedom. *Philosophy and Phenomenological Research* 88 (1): 1–37.

- Moran, Richard. 2012. *Authority and Estrangement. An Essay on Self-Knowledge*. Princeton, NJ: Princeton University Press.
- Nelson, John. 1964. In Defence of Descartes: Squaring a Reputed Circle. *Dialogue* 3: 262–272.
- Owens, David. 2000. *Reason Without Freedom. The Problem of Epistemic Normativity*. London: Routledge.
- Rubin, Ronald. 1977. Descartes' Validation of Clear and Distinct Apprehension. *The Philosophical Review* 86 (2): 197–208.
- Ryan, Sharon. 2003. Doxastic Compatibilism and the Ethics of Belief. *Philosophical Studies* 114 (1/2): 47–79.
- Shah, Nishi. 2002. Clearing Space for Doxastic Voluntarism. *The Monist* 85 (3): 436–445.
- Shah, Nishi. 2003. How Truth Governs Belief. *The Philosophical Review* 112 (4): 447–482.
- Steup, Matthias. 1988. The Deontic Conception of Epistemic Justification. *Philosophical Studies* 53 (1): 65–84.
- Steup, Matthias. 2000. Doxastic Voluntarism and Epistemic Deontology. *Acta Analytica* 15 (1): 25–56.
- Van Cleve, James. 1979. Foundationalism, Epistemic Principles, and the Cartesian Circle. *The Philosophical Review* 88 (1): 55–91.
- Williams, Bernard. 1973. Deciding to believe. In: Williams, B.: *Problems of the Self. Philosophical Papers 1956–1972*. Cambridge: Cambridge University Press. 136–151.

# III

*Jan Forsman*

## Descartes on Will and Suspension of Judgment: Affectivity of the Reasons for Doubt

In this paper, I join the so-called voluntarism debate on Descartes's theory of will and judgment, arguing for an *indirect doxastic voluntarism* reading of Descartes, as opposed to a classic, or *direct doxastic voluntarism*. More specifically, I examine the question whether Descartes thinks the will can have a direct and full control over one's suspension of judgment.

Descartes was a doxastic voluntarist, maintaining that the will has some kind of control over one's doxastic states, such as belief and doubt (e.g. AT VII, 22 & 59–60; CSM II, 15 & 41).<sup>1</sup> According to a long-held reading, the control that the will has over doxastic states in Descartes's theory is *direct*; the doxastic states are affected by the mere act of will. This reading, called direct doxastic voluntarism (DDV) or direct voluntarism (DV) for short, states that we are capable of assenting, rejecting and suspending a judgment based only on our will to do so. Thus, these actions would be utterly and merely volitional. DV can be divided into two further positions, *direct positive voluntarism* (+DV) and *direct negative voluntarism* (-DV). Direct positive voluntarism deals with the act of forming judgments, maintaining that one can accept or deny a proposition wilfully and either merely believe or not believe something voluntarily. Direct negative voluntarism deals with the suspension of judgment, maintaining that it can likewise be accomplished by a simple act of will (cf. e.g. Newman 2008, 343; Vitz 2010, 107–108 & 2015a, 73–74; Schüssler 2013, 148–150).

However, I support an alternate account of Descartes's voluntarism, which is called indirect doxastic voluntarism (IDV) or indirect voluntarism (IV) for short. By this account, the will is capable of affecting a doxastic state indirectly by making one concentrate on

---

1 When referring to Descartes, I use the standard style of reference, where AT stands for the 12-volume edition of original texts by Adam and Tannery, CSM stands for the 2-volume English translations by Cottingham, Stoothoff and Murdoch, and CSMK stands for the translations of correspondence (in the third volume of the latter edition) by Cottingham, Stoothoff, Murdoch and Kenny.

essential tasks for forming that state, such as gathering up and paying attention to *strong reasons* and *evidence*. IV is also possible to divide into *indirect positive voluntarism* (+IV) and *indirect negative voluntarism* (-IV). Per indirect positive voluntarism the will needs to pay attention to *reasons for accepting or denying* some proposition. Likewise, by indirect negative voluntarism, in order to suspend judgment the will needs to direct this attention to the *reasons for doubt*. By attending to these reasons, the will also comes face-to-face with its own freedom (AT VIII 6; CSM I, 194. Cf. AT VII, 57; CSM II, 40). This feeling of freedom can be described as *affectivity* of the reasons for belief (assent) and doubt (suspension).

My main goal in this paper is to defend indirect voluntarism over direct voluntarism, especially concentrating on voluntarism concerning suspension of judgment (-IV). Even though indirect voluntarism has gained some attention in the literature (e.g. Kenny 1998; Cottingham 1988 & 2002; Della Rocca 2006; Newman 2008), the systematic defense it deserves has not been attempted yet. My own reading is intended to fill this gap.<sup>2</sup> I put forward three kinds of evidence for indirect voluntarism: 1) All the textual evidence that seems to support DV can just as easily be read to support IV. 2) IV is a philosophically and psychologically more convincing account of the judgment-forming process in humans than DV. 3) -IV is a more coherent and less conflicting reading of the general suspension of judgment by the Method of Doubt in the *First Meditation*.

The paper is divided into four parts. In Part 1, I start with a short introduction to Descartes's account on judgments and beliefs and especially on the role of the will in forming them. In Parts 2 and 3, I introduce direct and indirect voluntarism and bring forward the textual and non-textual evidence supporting them. Finally, in Part 4, I compare Descartes's theory of will and judgment with the suspension of judgment in the *First Meditation*, defending indirect (negative) voluntarism and laying out the evidence for my own reading.

---

2 However, DV does not need to hold that the will can directly suspend from judging or deny *any proposition*. Many readings consider that *clear and distinct* perceptions are utterly irresistible for the will. When confronted by a clear and distinct perception, the will would always accept it. Cf. e.g. Frankfurt (2008, chapter 11), Kenny (1998, 150–159), Curley (1975, 177), Williams (2015, 165–167), Wilson (1978, chapter 3), Rosenthal (1986, 431), Newman (2008, 338–342) and Carriero (2009, chapter 4). However, Descartes's stance on this is quite ambiguous (cf. especially the Letter to [Mesland], 9 February 1645: AT IV, 173–175; CSMK, 244–246). Because of this, not all commentators take the irresistibility of clear and distinct perception at face value and some view Descartes as retaining the will's independent power of choice even in these cases. Indeed, there's an interesting debate going on in the current literature about this topic. Cf. esp. Alanen (2003, chapter 7; 2013), Newman (2008), Shapiro (2008), Carriero (2009, chapter 4), Schüssler (2013) and Wee (2014). Even though this debate is certainly important, in this paper I will mostly stick to the voluntarism debate on perceptions that are not clear and distinct. However, even when it comes to clear and distinct perception, Descartes seems to maintain IV: clear and distinct perception is evidence that is *so strong* that the will has no choice but to accept it. Despite this, Descartes seems to want to preserve the freedom to suspend judgment on them at least in an absolute or ideal sense (cf. Schüssler 2013, 163).

# 1. Will in Descartes's *Theory of Judgment*

To get a better understanding of Descartes's doxastic voluntarism it is necessary to first lay out the role of the will as Descartes sees it. According to Descartes, thinking (*cogitatio*) is divided into two modes: *perceptio* (perception) and *volitio* (will). Descartes also calls *perceptio* understanding (*intellectus*). In the *Principles of Philosophy* (1644, henceforth the *Principia*) he further explains the topic as follows:

All the modes of thinking that we experience within ourselves can be brought under two general headings: perception [*perceptio*], or the operation of the intellect [*operatio intellectus*], and volition [*volitio*], or the operation of the will [*operatio voluntatis*]. Sensory perception, imagination and pure understanding are simply various modes of perception; desire, aversion, assertion, denial and doubt are various modes of willing (*Principia* I, §XXXII: AT VIII A, 17; CSM I, 204).

For Descartes then, the will is a faculty, moreover a *free* faculty:

[T]he will simply consists in our ability to do [*facere*] or not to do [*non facere*] something (that is, to affirm or deny, to pursue or avoid); or rather, it consists simply of the fact that when the intellect puts something forward for affirmation or denial or for pursuit or avoidance, our inclinations are such that we do not feel we are determined by an external force"<sup>3</sup> (Med. IV, 8.: AT VII, 57; CSM II, 40).<sup>4</sup>

Beliefs and opinions, on the other hand, are judgments. In the famous example of the *Second Meditation*, Descartes's meditator examines first a piece of wax and then people walking outside the window:

We say that we see the wax itself, if it is there before us, not that we judge it to be there from its colour or shape; and this might lead me to conclude without more ado that knowledge of the wax comes from what the eye sees, and not from the scrutiny

---

3 This quote is particularly relevant to another discussion on Descartes's understanding of the will. Does Descartes follow earlier (direct) voluntarists like Duns Scotus, Occam and Suárez and consider the will's freedom to be essentially indifference (hence, the two-way power to do otherwise), like the first part of the quote suggests, or spontaneity (hence, being self-caused and undetermined by anything external), like the latter part suggests? See e.g. Ragland (2006), Alanen (2013), Schüssler (2013), Cunning (2014) and Wee (2014). I will mostly skip this discussion here. However, I view that Descartes holds both to be part of the will's freedom. In the case of clear and distinct perceptions, the will experiences spontaneous freedom, as it feels strongly inclined towards them. When perception is not clear and distinct, one can reach a state of equilibrium related to the reasons for assent and non-assent. In such a state, the will experiences indifferent freedom, as neither side is stronger or more inclined than the other (AT VII, 22 & 57–58; CSM II, 15 & 40).

4 When referring to Descartes's work *Meditations on First Philosophy* (1641–1642, henceforth *Meditations*) I also add the number of Meditation (Med.) and paragraph (p.) of the text I am referring to.

of the mind alone. But then if I look out of the window and see men crossing the square, as I just happen to have done, I normally say that I see the men themselves, just as I say that I see the wax. Yet do I see any more than hats and coats which could conceal automatons? I *judge* that they are men. And so something which I thought I was seeing with my eyes is in fact grasped solely by the faculty of judgment which is in my mind (Med. II, 13.: AT VII, 32; CSM II, 21. Emphasis in the original).

This was clearly a difficult part for the authors of the *Sixth Objections*, and so Descartes had to explain it further in his *Replies*:

For example, when I see a stick [...] rays of light are reflected off the stick and set up certain movements in the optic nerve, in the brain, as I have explained at some length in the *Optics* [1637]. This movement in the brain, which is common to us and the brutes, is the first grade of sensory response. This leads to the second grade, which extends to the mere perception of the colour and light reflected from the stick [...] Nothing more that this should be referred to the sensory faculty, if we wish to distinguish it carefully from the intellect. But suppose that, as a result of being affected by this sensation of colour, I judge that a stick, located outside me, is coloured; and suppose that on the basis of the extension of the colour and its boundaries together with its position in relation to the parts of the brain, I make a rational calculation about the size, shape and distance of the stick; although such reasoning is commonly assigned to the senses (which is why I have here referred to the third grade of sensory response), it is clear that it depends solely on the intellect (AT VII, 437–438; CSM II, 295).

Descartes thus differentiates between sensing as a bodily function (first grade) and sensing as a representation (second grade), viewing the latter as the mental part of sensation. Besides these, he further differentiates a third grade, which is related to the understanding, being judgments about the mental representation.

However, Descartes does not base judgment solely on understanding. In *Notes on a Certain Broadsheet* (1648), he heavily criticises Regius for dividing understanding into perceiving and judging.

I saw that over and above perception, which is a prerequisite of judgment, we need affirmation and negation to determine the form of the judgment, and also that we are often free to withhold our assent, even if we perceive the matter in question. Hence I assigned the act of judging itself, which consists simply in assenting (i.e. in affirmation or denial) to the determination of the will rather than to the perception of the intellect (AT VIIIIB, 363; CSM I, 307).

Judgments are not mere acts of understanding but come about by the cooperation of understanding and the will. To be precise, making a judgment is an act of the will.

Allow me to elucidate this with an example. When I happen to turn my attention from writing this paper and look outside the balcony window, I gain a sense impression of a tree growing next to the railway track. I see that the tree exists and that it has certain qualities, such as size, figure and the colour of its leaves (or that it is completely leafless, like now). My belief of the existence of the tree is a judgment, which is formed when a sense impression of a tree is conveyed to my understanding and I accept it to be real and existing by my will. Because my will is free and has the independent freedom of choice, I may voluntarily also deny said proposition as false (the tree does not exist) or alternatively suspend my judgment on the existence of the tree altogether. If I for example consider that my perception of the tree is too dim and obscure, I can decline the judgment that the tree exists (it merely is a telephone pole which I mistake for a tree). Alternatively, if I consider that I might presently only be dreaming of seeing a tree (and even of writing the paper, perhaps), I can wilfully suspend judgment on whether the tree exists or not. As Descartes describes in the *Principia* (I, §VI): “[W]e [...] experience within us the kind of freedom [*libertatem esse experimur*] which enables us always to refrain [*abstinere*] from believing things which are not completely certain and thoroughly examined. Hence we are able to take precautions against going wrong on any occasion” (AT VIII 6; CSM I, 194).<sup>5</sup>

Belief then follows from accepting or alternatively denying a given proposition. It should also be noted that acceptance does not simply mean shrugging one’s shoulder regarding the proposition in question. Accepting a proposition conveyed to the understanding means that one *truly believes* in it (cf. e.g. Williams 2015, 161). However, a question follows: if my beliefs come from judgments and my judgments are formed by my will, which for Descartes is free (no outside force can make me accept the existence of the tree), does this mean that according to Descartes I can wilfully believe whatever I want? Even if I very well knew I am not dreaming in this instance (say, I perform a test by pinching myself), can I nevertheless suspend my judgment on the tree’s existence? Better yet, to have an even more radical example, can I believe that there is a warm summer in Finland, even if all my senses tell me that it is winter?

## 2. Direct Voluntarism

Do I then have direct, easy and unproblematic voluntary access to my beliefs? Even so direct, easy and unproblematic that I am capable of believing in any proposition I can think of? As an example, let’s say that I find the winter period in Finland completely hostile and in order to improve my mood, I decide to believe it is in fact summer. Even though all the

---

5 For a more detailed description of the role of willing and judging in Descartes’s philosophy, see e.g. Kenny (1998), Rosenthal (1986), Della Rocca (2006), Newman (2008), Shapiro (2008), Kambouchner (2008) & Naaman-Zauderer (2010). Schüssler (2013) pays close attention to the role of doxastic voluntarism in late-scholastic discussions, especially among the Jesuits, which also motivated Descartes’s stance on the issue.

evidence I can gather alludes to a freezing cold winter in Finland, can I nevertheless believe that it is summer?<sup>6</sup>

According to direct positive voluntarism, Descartes would answer affirmatively. My will has a direct, uncomplex relation to my doxastic states, and even though all the evidence my understanding gathers points towards winter, my will has an independent ability to either accept or deny this proposition. If I end up denying the winter proposition, I can perform another direct act of the will to accept another, in this case the summer proposition. The evidence supporting this might not be as strong, but according to +DV my will has a direct and independent power to both accept and deny propositions, thus I can just as easily accept that it is summer (cf. Curley 1975; Grant 1976).

Direct negative voluntarism works the same way for suspension of judgment. When the meditator states in the beginning of the *First Meditation* that earlier acquired knowledge is full of falsehoods and because of this decides to overthrow everything previously learned (Med I., 1.: AT VII, 17–18; CSM II, 12.), according to -DV, general overthrow of all opinions would already follow from this mere decision. In this so-called *provisional suspension of judgment*, the meditator would suspend her judgment on earlier beliefs and opinions merely because she *decides* to suspend them. The resolution to overthrow all beliefs is already in itself the act of overthrowing. No further steps would be required. Suspension of judgment is utterly voluntary and one can suspend a judgment by merely deciding so. Because of this, the provisional suspension of judgment by which the *First Meditation* begins would be *distinct* from the following skeptical scenarios (madness, dream, deceiving God, origin by faith or chance, malicious demon). Suspension of judgment would then occur even before consideration of these scenarios (cf. Frankfurt 2008, 24–31).

This reading can be supported by the *Fourth Meditation*, where Descartes's meditator states: "If [...] I simply refrain from making a judgment in cases where I do not perceive the truth with sufficient clarity and distinctness, then it is clear that I am behaving correctly and avoiding error". This seems to indicate complete affective control over our doxastic states and an ability to suspend judgment merely by an act of will. The same seems to be suggested by the comment of God-given "freedom to assent or not to assent" in those cases where there is no clear and distinct perception (Med. IV, 12. & 15.: AT VII, 59–61; CSM II, 41–42). Another way to support the reading is to look at the use of the *Method of Doubt*. In the *First Meditation*, the meditator concludes that "[I]n [the] future I must withhold my assent from these former beliefs just as carefully as I would from obvious falsehoods, if I want to discover any certainty". Later on, she adds: "In view of this, I think it will be a good plan to turn my will in completely the opposite direction and deceive myself, by pretending for a time that these former opinions are utterly false and imaginary" (Med. I, 10–11.: AT VII, 21–22; CSM II, 15. Cf. e.g. Della Rocca 2006, 148; Newman 2008, 344). Likewise, in the letter to Clerselier (12 January 1646), Descartes states:

---

6 The inspiration for this example comes from an article by Brian Grant titled "Descartes, Belief and the Will" (1976).

[S]ince making or not making a judgment is an act of will [*action de la volonté*] [...] it is evident that it is something in our power. For, in order to get rid of every kind of preconceived opinion, all we need to do is resolve not to affirm or deny [*ne rien assurer ou nier*] anything which we have previously affirmed or denied until we have examined it afresh (AT IXA, 204; CSM II, 270).<sup>7</sup>

Based on these remarks, many commentators have taken the will to have a direct effect on the suspension of judgment. For instance, Janet Broughton considers Descartes as meaning that “it is here and now within my power to suspend judgment about the truth of anything I believed” (Broughton 2002, 58). Harry G. Frankfurt likewise views that the suspension of judgment “results directly from a decision or an act of will”. A person suspends judgment merely by resolving that his judgments are suspended (Frankfurt 2008, 25). Even Michael Della Rocca, though inclining more towards indirect voluntarism, agrees that “[d]irect control is what Descartes’ talk in *Meditation IV* of ‘the freedom to assent or not to assent’ most naturally suggests...” (Della Rocca 2006, 148).<sup>8</sup>

It is useful to notice that this view of Descartes’s theory of judgment and will is upheld not only by contemporary scholars. Already in the *Fifth Objections*, Gassendi asks: “[W]hy did you not make a simple and brief statement to the effect that you were regarding your previous knowledge as uncertain so that you could later single out what you found to be true?” (AT VII, 257; CSM II, 180). In other words, if suspending judgment is indeed this easy, why doesn’t Descartes simply state the suspension without further decorum? Why should one even bother with the skeptical scenarios when the suspension can be accomplished directly, with a single strike of the will?

Descartes however denies Gassendi’s statement that the suspension can be accomplished in a superficial way:

Is it really so easy to free ourselves from all the errors which we have soaked up since our infancy? Can we really be too careful in carrying out a project which everyone agrees should be performed? But not doubt you wanted to point out that

---

7 This letter, and the “Author’s and Translator’s notes concerning the Fifth Set of Objections” (*Avertissement de l’auteur touchant les cinquièmes objections & Avertissement du traducteur touchant les cinquièmes objections*) preceding it, were published as part of the first edition of the *Meditations*’s translation into French (1647). The letter along with the author’s note are translated in CSM as the *Appendix to the Fifth Set of Objections and Replies*.

8 Other readings supporting +DV have been offered by e.g. Wilson (1978, 144–146), Williams (2015, 161–163) and less evidently MacArthur (2003, 166). In recent literature, Vitz (2010; 2015a, chapter 6) has most prominently defended -DV. Schüssler on the other hand views Descartes holding both direct and indirect (negative) voluntarism. When it comes to clear and distinct perception, Descartes is on the side of -IV (though is keen on preserving the possibility of direct suspension of judgment, at least ideally). With perceptions that are not clear and distinct Descartes would hold -DV, except with entrenched and habitual opinions, which can only be dislodged indirectly (Schüssler 2013, 173). My reading differs from Schüssler in that I view Descartes holding indirect voluntarism for both habitual/ deeply ingrained and newly made fresh beliefs.

most people, although verbally admitting that we should escape from preconceived opinions, never do so in fact, because they do not spend any care or effort on the task... (AT VII, 348; CSM II, 242).

Yet this does not have to be a problem for -DV. Frankfurt emphasises the following example when defending reading of the suspension of judgment as “an uncomplicated act of will” (Frankfurt 2008, 26): Mr. X decides to stop smoking at noon of the 1<sup>st</sup> of January. This is an uncomplicated act of will, which can be done very easily by Mr. X. Yet, based only on this, can we say that Mr. X really did stop smoking at noon of the 1<sup>st</sup> of January? If Mr. X smoked a cigarette at 12:30 of the same day, the answer would obviously be no. But if by the end of December the next year Mr. X had not smoked one cigarette, we could be able to say yes – Mr. X did indeed stop smoking. Frankfurt sees the meditator to be in a similar position to Mr. X. She can make the decision to suspend her judgment on all beliefs (and by Frankfurt’s account, this decision is something she can do very easily and without effort), but if she would immediately go to accept another belief, she wouldn’t have suspended her judgment after all (Frankfurt 2008, 29–30).

This example can be supported by the *Objections and Replies*. In the already mentioned letter to Clerselier Descartes comments:

[N]o matter how much we have resolved to assert or deny anything, we easily forget our resolution afterwards if we have not strongly impressed it on our memory [*fortement imprimée en sa memoire*]; and this is why I suggested that we should think about it very carefully [*pensait avec soin*] (AT IXA, 204; CSM II, 270).

This seems to also be implied by the following comment at the end of the *Fourth Meditation*:

[E]ven if I have no power to avoid error in the first way [...], which requires a clear perception of everything I have to deliberate on, I can avoid error in the second way, which depends merely on my *remembering* [*recorder*] to withhold judgment on any occasion when the truth of the matter is not clear. Admittedly, I am aware of a certain weakness in me, in that I am unable to keep my attention fixed on one and the same item of knowledge at all time; but by attentive and repeated meditation I am nevertheless able to make myself remember it as often as the need arises, and thus get into the habit of avoiding error (Med. IV, 16.: AT VII, 61–62; CSM II, 43. Emphasis added).

The skeptical scenarios then do help with the suspension of judgment, but only by reinforcing the resolution to suspend judgment and thus helping to steer clear of forming new beliefs (cf. Frankfurt 2008, 29–30; Broughton 2002, 58). I will henceforth be referring to this as the *memory-argument*.

Frankfurt also suggests that the *Seventh Objections and Replies* provides another supportive paragraph for his reading:

Suppose he [Bourdin, the author of the *Seventh Objections*] had a basket full of apples and, being worried that some of the apples were rotten, wanted to take out the rotten ones to prevent the rot spreading. How would he proceed? Would he not begin by tipping [*rejiceret*] the whole lot out of the basket? And would not the next step to be to cast his eye over each apple in turn, and pick up and put back in the basket only those he saw to be sound, leaving the others? In just the same way, those who have never philosophized correctly have various opinions in the minds which they have begun to store up since childhood, and which they therefore have reason to believe may in many cases be false (AT VII, 481; CSM II, 324).

This *basket-of-apples* analogy suggests that Descartes's project of overthrowing his opinions has two phases. In Phase 1, the "basket" is tipped over and judgment is suspended on all earlier beliefs. In Phase 2, the earlier beliefs are closely examined to see which of them can be reinstated. According to Frankfurt, the skeptical scenarios would belong to Phase 2, being used to examine the earlier beliefs and opinions. If doubt can be cast, e.g. on the reliability of sensory perceptions, this belief would not be reinstated but left suspended. By Frankfurt's account, this analogy suggests that "emptying one's mind is a rather headlong and indiscriminate affair", while evaluating the former opinions requires careful argumentation (Frankfurt 2008, 27–28). Suspension of judgment would therefore be an easy task and could be done with a simple and direct act of willing the suspension.<sup>9</sup>

### 3. Indirect Voluntarism

Is the act of the will truly this simple? If I now used all of my energy for wanting it to be summer in Finland instead of winter, would my belief actually change? No matter how much I try, I don't seem to be able to affect my belief in winter in any way. To be sure, I don't even know what mental apparatus to use. Should I imagine daisies? By imagining daisies, I might momentarily come to the conclusion it is now summer, the sun is shining and the grass outside is blooming with daisies. But once my thoughts get distracted, or if I just happen to look outside the window, I once again can only assert: "By Plato's beard, it's a cold winter!". If the directly voluntaristic reading is accurate, Descartes seems to hold a view that is quite problematic both philosophically and psychologically.

---

9 Broughton also views that the meditator can suspend her judgment merely by deciding to not accept or deny, and that the only obstacle for this action lies in the difficulty of remembering the resolution, instead of in any difficulty or inability while trying it (Broughton 2002, 58). However, she disagrees with Frankfurt that suspension of judgment would occur before considering the skeptical scenarios (*ibid*, 55, note 20). More on this below, in note 15.

However, it seems that DV simplifies Descartes's account in an unwarranted way. Indeed, the opposing theory, indirect voluntarism, has in recent discussions gained growing support.<sup>10</sup> IV denies directly voluntaristic affective control over doxastic states and views this to be indirect. The will can affect belief-forming in a *mediate* or indirect way by affecting the essential tasks required for forming doxastic states, such as directing the attention towards additional evidence for the desired state – to be more precise, by attending to the *reasons* to assent or suspend. Per indirect positive voluntarism, accepting a proposition requires one to attend to the reasons for accepting, while denying that proposition requires one to attend to the reasons for denying it. Likewise, per indirect negative voluntarism, doubting the proposition at hand requires one to attend to the reasons for doubt (reasons that question said proposition), while suspension of judgment requires one also to attend to the reasons for suspending judgment (cf. e.g. Newman 2008, 343; Vitz 2010, 107–108 & 2015a, 73–74; Schüssler 2013). By considering these reasons, one likewise gets to feel the *freedom of will*, thus feeling the *affectivity* of the reasons for making a free choice.

Like direct voluntaristic readings, readings supporting IV can refer to both the use of the *Method of Doubt* in the *First Meditation* and to the role of the will in the *Fourth*. Right at the beginning of the *First Meditation*, the meditator agrees to the following maxim:

I should hold back my assent from opinions which are not completely certain and indubitable [*indubitata*] just as carefully as I do from those which are patently false. So for the purpose of casting aside [*rejiendas*] all my opinions, it will be enough to find in each of them at least some reason for doubt [*rationem dubitandi*] (Med. I, 2.: AT VII, 18; CSM II, 12. Emphasis added. Translation altered).<sup>11</sup>

Descartes then argues that to suspend judgment, the meditator first requires *reasons for doubt*. These reasons are provided by the different skeptical scenarios, which the meditator ponders in the *First Meditation* and which make her question her earlier beliefs. For example, finding particular sense perception occasionally deceiving does not yet make the meditator become convinced that the senses are generally not trustable for concluding that I am here sitting by the fire, writing on this piece of paper (or in the balcony, writing on my laptop) (Med. I, 4.: AT VII, 18; CSM II, 12–13). The same applies to every opinion one suspends. After the meditator has stated that “in future I must withhold my assent from these former beliefs just as carefully as I would from obvious falsehoods” and that to succeed in this it is a “a good plan” to turn one’s will in the opposite direction completely, and

10 Commentators supporting IV include e.g. Kenny (1998), Cottingham (1988; 2002), Della Rocca (2006, 149), Newman (2008, 343–346), Perin (2008, 60–62), Carriero (2009, 262–263), Naaman-Zauderer (2010, 120) and possibly Rosenthal (1986, 411). Even Curley notes that perhaps Descartes is not quite the direct voluntarist he seems, but does not take the thought further (Curley 1975, 176).

11 CSM translates *rejiendas* as “rejecting”. However, it is not a conjugation of the word *rejectio* but of the verb *rejicere*, which literally translates as “throwing back”. Descartes uses the same term with the basket-of-apples analogy, referring to removing apples from a basket (*ex corbe rejiceret*), where he is not talking of complete rejection of the apples. The meaning of the word seems indeed to be less strong than *rejection* would implicate. I have thus opted for translating it as “casting aside”.

deceive oneself to pretend those opinions are utterly false, Descartes immediately comes up with another skeptical scenario (the malicious demon) by which the meditator could succeed in this self-deception (Med. I, 11–12.: AT VII, 22; CSM II, 15).

Descartes also emphasises the reasons for doubt in two other instances. In an unfinished dialogue, *The Search for Truth* (henceforth the *Search*), one of the debating characters (Epistemon)<sup>12</sup> states: “But you are not ignorant of the fact that the opinions first received in our imagination remain so deeply imprinted [*imprimées*] there that our will *cannot erase* [*effacer*] them on its own, but can do so only by calling on the *assistance of powerful reasons* [*puissantes raisons*]” (AT X, 509; CSM II, 406. Emphasis added). In the letter to Clerselier, Descartes also comments:

Nevertheless, I did say that there was some difficulty in expelling from our belief everything we have previously accepted. One reason for this is that before we can decide to doubt we need some *reason for doubting* [*raison de douter*] and that is why in my *First Meditation* I put forward the *principal reasons for doubt* [*les principales*] (AT IXA, 204; CSM II, 270. Emphasis added).

Descartes then clearly holds that one requires strong reasons to suspend judgment and likewise to doubt. Therefore, the skeptical scenarios precede the suspension of judgment, because the former is a necessary requirement for the latter.<sup>13</sup> As the reply to Gassendi also indicates, Descartes denies the ease of suspending judgment and holds that suspension requires more than a solitary act of the will. This also indicates Descartes holding indirect effect on the doxastic states (cf. Della Rocca 2006, 149; Newman 2008, 344; Cottingham 1988).

The *Fourth Meditation* can likewise be read as support for IV. In stating “[i]f I [...] simply refrain from making a judgment in cases where I do not perceive the truth with sufficient clarity and distinctness, then it is clear that I am behaving correctly and avoiding error”, Descartes would mean that in such cases our responsibility is to concentrate and attend to the reasons for doubt and suspension. Similarly, the statement of God having given us “the freedom to assent or not to assent” to perceptions is not in contradiction with IV. God could have given us such an ability, without that ability being a direct act of the will (Med. IV, 12. & 15.: AT VII, 59–61; CSM II, 41–42).

It is important then to notice that every textual evidence which seems to support DV can just as easily be read as supporting IV. However, this is clearly not yet sufficient for

---

12 In the dialogue, Epistemon (gr. ἐπιστημων, ‘knowledgeable’) represents a spokesperson for scholasticism and for this reason does not generally present opinions which Descartes himself holds. However, there are reasons to view Descartes as agreeing with this statement. I return to these reasons in Part 4.

13 However, the skeptical scenarios – that is to say the reasons for doubt – do not by themselves suffice as reasons for suspending judgment. Suspension of judgment similarly requires a reasoning, which is not constituted by the reasons for doubt. I might find some reason to doubt the landing on the moon to have taken place (perhaps, it was a hoax) but this by itself does not yet make me suspend my judgment on whether man has been to the moon. Despite this, for the suspension to succeed, it requires the aid of the reasons for doubt. For this, see also Perin (2008).

establishing indirect voluntarism as the better and stronger interpretation. Indeed, some commentators step outside the text, asking what kind of position Descartes *should* be holding as an intelligent philosopher (e.g. Newman 2008, 345–346). In other words, they ask what would be a philosophically and psychologically convincing account of our ability to accept, deny or suspend judgment.

As the earlier example of winter in Finland demonstrates, evidence does not seem particularly strong for DV. Instead, there are strong reasons to consider it a psychologically problematic position. Direct voluntary acts simply don't seem to have much to do with my beliefs or opinions. Most commentators supporting a directly positive voluntaristic reading of Descartes also admit this. Williams comments that “if this is what assent is, it is far from clear how assenting is even dependent on the will, let alone a mode of it”, asking: “are there not a very large number of things that one just cannot believe, and others that one cannot help believing?” (Williams 2015, 161). Wilson adds that this sort of view on the role of the will seems questionable since “of course we can't just decide to believe or assent to something, and forthwith believe or assent to it”. This, as she states, can also be discovered phenomenologically (Wilson 1978, 145). Broughton likewise considers direct negative voluntarism as “false to the psychology of human intellectual activity”, maintaining that the meditator “can no more suspend judgment by willing to do so than [s]he can believe by willing to do so” (Broughton 2002, 59). However, each of the above commentators reads Descartes's theory of judgment and will as directly (positively or negatively) voluntaristic. If DV is so philosophically unconvincing and psychologically problematic, wouldn't this also speak for IV? At least this would be the case, if we were to presume that Descartes is aiming towards a theory which is as convincing as possible, both philosophically and psychologically.<sup>14</sup>

Thus far, I have argued for indirect voluntarism being a better reading than direct voluntarism for two reasons: 1) Texts that support DV can just as well be read as supporting IV; and 2) IV is both philosophically and psychologically more convincing an account of our ability to accept, deny and suspend judgment. However, these reasons do not yet tip the scales for IV. Merely observing the textual evidence as supporting both sides does not validate one over the other. Similarly, referring to an account that Descartes *should* be holding does not mean that he actually holds it. Newman (2008, 345) especially builds on Descartes's intelligence, arguing that we should not be too hasty in putting a problematic theory like DV into his mouth. However, as Schüssler (2013, 172) has also pointed out, if evidence for DV can be found in Descartes's texts, Newman's argument does not hold. Thus, in the next part, I will be demonstrating what I call a 'knock-out argument' for why indirect (negative) voluntarism is a better reading than direct (negative) voluntarism: -IV is a more coherent and less self-contradictory reading of the general suspension of judgment in the *First Meditation* than -DV.

---

<sup>14</sup> However, DV might not in fact be as impossible a position as it has sometimes been portrayed. Cf. esp. Vitz 2015a, appendix & 2015b.

## 4. Will and Suspension of Judgment in the *First Meditation*

In Part 2, I went through the way in which a supporter of -DV can read the suspension of judgment in the *First Meditation*. Suspension of judgment occurs effortlessly, with a simple act of the will – the decision to suspend is the act of suspension itself and the only related difficulty comes from remembering this decision (the so-called memory-argument) (cf. Frankfurt 2008, 29–30; Broughton 2002, 58).<sup>15</sup> This account is usually defended by the *Letter to Clerselier* and by the end of the *Fourth Meditation*. The end of the *First Meditation* seems to also support this: “But it is not enough merely to have noticed this; I must make an effort to remember [*recorder*] it” (Med. I, 11.: AT VII, 22; CSM II, 15). Other suggested evidence for this reading includes the basket-of-apples analogy, which would make the general overthrow of opinions to have two phases, the skeptical scenarios constituting the second phase (cf. Frankfurt 2008, 27–28).

However, this account is problematic in that it leaves us with a great deal of incoherence on Descartes’s part.<sup>16</sup> First, it suggests the suspension of judgment to be far easier than Descartes himself describes it to be. In fact, Descartes goes to great lengths to emphasise how *difficult* and *unnatural* this general overthrow of opinions actually is. Note for example the comment at the beginning of the *First Meditation*: “But the task looked an enormous one, and I began to wait until I should reach a mature enough age to ensure that no subsequent time of life would be more suitable for tackling such inquiries” (Med. I, 1.: AT VII, 17; CSM II, 12). The difficulty of the task is already implied in this comment on its *enormity* – why would the meditator otherwise have to wait for a “mature enough age” that would be more suitable for engaging in the inquiry. At the end of the *First Meditation*, she also describes its result as an “arduous [*laboriosum*] undertaking” (Med. I, 12.: AT VII, 23; CSM II, 15).

In his reply to Gassendi, Descartes likewise questions the idea of the ease to free ourselves from preconceived opinions and common-sense experience. He also emphasises the required effort for the suspension of preconceived opinions in here as well: “But no doubt you wanted to point out that most people, although verbally admitting that we should escape from preconceived opinion, never do so in fact, because they do not spend any *care and effort* [*studium aut laborem*] on the task” (AT VII, 348; CSM II, 242. Emphasis added). Based on these comments, it seems when discussing suspension

---

15 It should be noted though that Broughton disagrees with Frankfurt on the issue of Descartes employing suspension of judgment for the meditator *before* she has considered the skeptical scenarios (Broughton 2002, 55, note 20). However, since Broughton also sides with -DV, it is not clear why the meditator could (would) not suspend on judgment merely by willing so, especially since she considers suspension to be in one’s power “here and now” (ibid, 58). Broughton acknowledges some of the difficulties for her position (ibid, 58–59).

16 It is to be noted that Frankfurt is fully aware of this, bringing up some textual problems for his reading (ibid, 25–31).

of judgment, Descartes has in mind something else than a simple and easy act of the will.

However, a proponent of DV does not have to commit to the ease of an act of the will. That voluntariness is difficult does not mean that it cannot be direct. Therefore, referring to the difficulty of suspension is not enough for a convincing argument against -DV.<sup>17</sup> Let's then return to the reasoning behind indirect (negative) voluntarism. In the *Search* (through the mouth of Epistemon), Descartes dictates that the will is not enough to overthrow the earlier opinions on its own and needs the assistance of "powerful reasons" (AT X 509; CSM II, 406). It is true that Epistemon, who in the dialogue represents Aristotelian scholastics (and who can therefore be justifiably seen as holding "the opinions first received" coming from the senses), does not generally represent Descartes's own stand on matters.<sup>18</sup> However, Eudoxus<sup>19</sup>, who in the dialogue is almost purely Descartes's alter ego, does not challenge the view, but instead goes on to *provide these reasons* (though not to Epistemon but to the third character in the discussion, Polyandros)<sup>20</sup> (AT X, 509; CSM II, 406–407). This last paragraph demonstrates well why the overthrow is so difficult. To be able to overthrow one's preconceived opinions, one has to have powerful reasons for the overthrow. Descartes then unequivocally denies that an act of the will would in itself be enough for the suspension.

The second incoherence in -DV comes from the order of the suspension of judgment and the skeptical scenarios. Descartes's own account becomes very clear in the Letter to Clerselier:

[S]ince making or not making a judgment is an act of will (as I have explained in the appropriate place) it is evident that it is something in our power. For, after all, in order to get rid of every kind of preconceived opinion, all we need to do is resolve *not to affirm or deny* anything which we have previously affirmed or denied until we have examined it afresh. But this does not entail that we cease to retain all the same notions in our memory. Nevertheless, I did say that there was some difficulty in expelling from our belief everything we have previously accepted. One reason for this is that before we can decide to doubt, we need some *reason for doubting*; and *that is why* [*c'est pourquoi*] in my *First Meditation* I put forward *the*

---

17 Many of the proponents for -DV clearly do imply easiness (e.g. Frankfurt 2008 & Broughton 2002). This, however, is not necessary.

18 The scholastic status of Epistemon can be questioned by referring to the many other viewpoints during Descartes's time (e.g. atomism and materialism) which Descartes likewise wanted to overturn. However, Epistemon is described as having "a detailed knowledge of everything that can be learned in the Schools [*escholes*]", apparently referencing the scholastic school-system (AT X, 499; CSM II, 401). Thus, it is justifiable to consider Epistemon as the spokesperson for Aristotelian scholasticism.

19 Eudoxus (gr. εὐδοξός, 'famous' or 'one of good belief') represents an enlightened Cartesian philosopher in the dialogue and speaks for Descartes's own views.

20 Polyandros (gr. πολυανδρός, from πολυς ἀνήρ, 'everyman') represents a person who lacks tutoring but has untutored common sense and by this, according to Descartes, is more embracing of the overthrow of earlier opinions than someone with an Aristotelian education.

*principal reasons for doubt.* Another reason is that no matter how much we have resolved not to assert or deny anything, we easily forget our resolution afterwards if we have not strongly impressed it on our memory; and *this is why* [*c'est pourquoy*] I suggested that we should think about it very carefully (AT IXA, 204; CSM II, 270. Emphasis added).

The reason why Descartes presents the meditator with the skeptical scenarios in the *First Meditation* is that in order to be able to suspend judgment, one first requires *reasons for doubt*. As Descartes clearly places the skeptical scenarios *before* the suspension of judgment, it would be incoherent on his part if the suspension of judgment occurred before them, as a distinct act. Frankfurt views this to be a mistake on Descartes's part, claiming him to be confusing the two phases of his project as explained by the basket-of-apples analogy (Frankfurt 2008, 26 & 30–31). This is unconvincing to me for two reasons. First, it is insufficient to explain all the consequent incoherence in Descartes's writings. Second, and more importantly, the basket-of-apples analogy is not related to the discussion in the *First Meditation*. It is a reply to Bourdin's objection to the *Second*.

After concluding in the *Second Meditation* that she exists (Med. II, 3.), the meditator goes on to consider what she is (Med. II, 4–5). First of all, she ponders on what she previously thought she was before beginning the project to overthrow her opinions: "What then did I formerly think I was? A human [*hominem*]" (AT VII, 25; CSM II, 17. Translation altered).<sup>21</sup> Here Bourdin objects and asks if Descartes isn't now referring to something that was already found to be false. How can one refer to things which were already rejected and overthrown from the mind? (AT VII, 479; CSM II, 323.) The basket-of-apples analogy would be Descartes's reply to this objection: to examine preconceived conceptions it is necessary to tip the "basket" and go through what is in it one by one. However, turning the basket over does not mean throwing the apples away. Instead, they stay on the side for a later inspection (AT VII, 481–482; CSM II, 324).<sup>22</sup> In other words, the second phase of the 'two-phase project' happens in fact in the *Second Meditation*, as the meditator begins to examine her previous sense-based (Aristotelian) conceptions of herself and the world. The first phase, the general overthrow of opinions, happens in the *First Meditation* and there is no reason to assume its occurrence distinctly from the skeptical scenarios. These paragraphs seem to provide enough evidence for preferring a reading that views the suspension of judgment as coming after the skeptical scenarios, as a result of them.

In the beginning of the *First Meditation* (Med. I, 1.), Descartes has the meditator unequivocally devote herself to the "general overthrow [*eversio*] of [her] opinions" (AT VII,

---

21 CSM translates *hominem* as "a man". As I refer to the meditator with the female pronoun, I have opted for the more literal translation "human".

22 This is likewise supported by Descartes's comment to Clerselier: "But this does not entail that we cease to retain all the same notions in our memory [*sinon après (...) quoy qu'on ne laisse pas pour cela de retenir toutes les mesmes notions en sa memoire*]" (AT IXA, 204; CSM II, 270).

18; CSM II, 12. Translation altered).<sup>23</sup> But the overthrow itself does not happen here – the commitment here is merely *preliminary*.<sup>24</sup> It is a dedication to the *effort* required for the overthrow. This general overthrow of opinions requires careful attention and considerable mental effort, which the meditator must promise to adhere to. If the meditator was like Gassendi, acknowledging the need for the suspension but not taking the skeptical scenarios *seriously*, according to Descartes she would not be able to genuinely suspend her judgment.<sup>25</sup> Suspension of judgment requires commitment that is *serious* (sincere) and *free* (without reservation) (*seriò tandem & libere*). As the next paragraph demonstrates, this commitment requires one to concentrate on the reasons for doubt (*rationes dubitandorum*) (AT VII, 17–18; CSM II, 12).<sup>26</sup>

What about the memory-argument? According to it, suspension of judgment is direct, accomplished by a simple act of the will. However, since the decision to suspend is difficult to recall, skeptical scenarios are needed for the suspension to be properly and enduringly possible. Thus, the scenarios help with the suspension. When armed with the memory argument, -DV seems to check both of the required boxes: suspension of judgment is mentally difficult, and the skeptical scenarios are a requirement for it. Merely referring to the insufficiency of the will alone or to the suspension resulting from the skeptical scenarios is not enough to respond to its challenge for -IV.

But why would it be so difficult for the meditator to remember her earlier decision? If the suspension of judgment comes from a simple act of the will, it is not easy to explain why it would be difficult to recall this. Take for example the case of Mr. X quitting smoking. Mr. X quits smoking with a single mental act and decides at the same time never to smoke another cigarette again. According to a proponent of direct voluntarism, why would this decision be so difficult to recall? It seems doubtful that neither Mr. X nor the meditator would be a person with especially poor mnemonic abilities. (If Mr. X happens to be a person with an especially bad memory, the analogy does not really work.) Descartes makes it clear that the suspension of judgment is *generally* difficult for humans psychologically. However, nowhere else does he seem to consider memory to be the stumbling block of the human psyche.<sup>27</sup>

23 CSM translates *eversio* as “demolition”. However, I consider demolition to be misleading, as I don’t see the doubt as *rejecting* the former beliefs. For this, see note 11. Thus, I find the more fitting translation to be “overthrow”.

24 Cf. the *Discourse on the Method*, where the call to abandon all prior opinions comes in Part Two, but the radical doubt itself does not follow until Part Four (AT VI, 13–15 & 31–31; CSM II, 116–118 & 126–127). See also Broughton 2002, 5, note 7.

25 Cf. the *Fifth Objections and Replies* (AT VII, 257–258 & 348–351; CSM II, 180 & 241–243).

26 My reading therefore follows the tradition of seeing the *Meditations* as a *mental exercise*. Cf. esp. Alquié (1950, 176), Gilson (1951, 186), Guerout (1953, 39, note 16), Gouhier (1978, 110–112) & Frankfurt (2008, 20). Recent literature emphasising this aspect include e.g. Schüssler (2013, 172) & Vitz (2015a, chapter 2).

27 Cf. esp. the *Conversation with Burman* (ATV, 148; CSMK, 334).

Let's take a closer look at the evidence for the memory-argument. In the *Fourth Meditation*, the meditator declares:

[E]ven if I have no power to avoid error in the first way just mentioned, which requires a clear perception of everything I have to deliberate on, I can avoid error in the second way, which depends merely on my remembering to withhold judgment on any occasion when the truth of the matter is not clear. Admittedly, I am aware of a certain weakness in me, in that I am *unable to keep my attention fixed* on one and the same item of knowledge at all times [*ut non possim semper uni & eidem cognitioni defixuss inhaerere*]; but by *attentive and repeated meditation* [*attentâ & fraepius iteratâ meditatione*] I am nevertheless able to *make myself remember* it as often as the need arises, and thus get into the habit of avoiding error (Med. IV, 16.: AT VII, 61–62; CSM II, 43. Emphasis added).

According to Descartes then, remembering the decision is difficult because the attention span of the mind is restricted. The mind's grasp of the meditation easily loosens and for this reason the suspension is difficult to retain. Suspension of judgment is therefore *temporarily restricted*. However, when the meditation is replicated often enough with sufficient attention, one becomes more competent in the suspension and it can be accomplished for longer periods of time. This also becomes clear at the end of the *First Meditation* (Med. I, 11.):

But it is not enough merely to have noticed this; I must make an effort to remember it. *My habitual opinions* [*constuetae opiniones*] *keep coming back, and, despite my wishes, they capture my belief, which is as it were bound over to them as a result of long occupation and the law of custom* [*longo usu & familiaritatis*]. [...] In view of this, I think it will be a good plan to turn my will in completely the opposite direction and deceive myself, by pretending for a time these former opinions are utterly false and imaginary. I shall do this until the weight of preconceived opinion is counter-balanced and the distorting influence of habit no longer prevents my judgement from perceiving things correctly (AT VII, 22; CSM II, 15. Emphasis added).

The difficulty to remember is not because of the memory itself but rather the mind's natural propensity for believing in habitual opinions (e.g. that sensory perception is reliable). It is not easy to diverge from this inclination (and the inclination is in a sense justifiable), and for this reason the mind's grip from the suspension slackens. Suspension of judgment is cumbersome, not because it is difficult to remember, but because it is mentally laborious. For this same reason, it is also difficult to retain in memory. Recalling the suspension is then specifically paying attention to the reasons for doubt against the reasons for belief. By doing so, the vitality of the suspension is recalled and one can once again vigorously concentrate on it, while also being faced with the affectivity of the will's freedom. This is also

what Descartes means in the *Letter to Clerselier* by impressing the resolution not to affirm or deny anything strongly on memory (AT IXA, 204; CSM II, 270).

It is therefore crucial to note that the memory-argument in no way contradicts -IV. Instead, -DV seems to lead Descartes inescapably to incoherence and self-contradiction. I maintain then that the burden of proof lies with the supporters of -DV. If Descartes considers that the suspension of judgment is executed with a simple act of will, and the only difficulty therein lies in remembering, why would the suspension come about only after consideration of the skeptical scenarios, resulting from them?<sup>28</sup>

## 5. Conclusion

I have presented three kinds of evidence for indirect voluntarism in this paper: 1) All textual evidence for direct voluntarism can just as easily be interpreted for indirect voluntarism; 2) indirect voluntarism is a more convincing reading both philosophically and psychologically; and, what I consider the ‘knock-out argument’, 3) indirect negative voluntarism is a more coherent and less self-contradictory reading of the *First Meditation* and the suspension of judgment therein.

By this account, even Descartes does not argue that I can just decide to believe it to be summer, when all the evidence suggests winter. However, the situation can be different if I can find some evidence that it might be summer (say, even though there is snow on the ground and -12 degrees, my calendar informs me that it is in fact June), or at least some evidence that it might not be winter (and a reason to suspend my judgment on it being winter). In such a case, my will can deny the winter-proposition and affirm the summer-proposition, or at least suspend the belief in winter, while noticing the feeling of its own freedom. Therefore, it seems that those commentators reading Descartes’s judgment theory as directly voluntaristic have in their criticisms – if I may borrow a colloquial phrase – been barking up the wrong tree.<sup>29</sup>

To suspend judgment on the existence of the tree, I need something more than just will and the motivation for suspension. I also need to find reasons to consider my experience of the tree to be in some way in error or disconnected from the way the reality truly is

---

28 It should be noted though that perhaps not all criticism on Descartes’s theory of judgment and will is undue. After all, Descartes clearly is a doxastic voluntarist and views the will to have at least some kind of control over our doxastic states, be as it may that it is indirect. It is anything but clear whether voluntariness has something to do with our beliefs or suspension of them, a point that many commentators have likewise paid attention to (e.g. Curley 1975, 173–174; Della Rocca 2006, 149). The topic is still hotly debated (see for instance Shah 2002 & Vitz 2015b), though it may be that Descartes is at error here. However, in any case his voluntaristic theory is not as psychologically problematic as most other readings have suggested.

29 Vitz likewise finds the critique of direct voluntarism in Descartes as unjustified, but for different reasons. According to him, this criticism has only been directed towards +DV and not -DV (Vitz 2010; 2015a, chapter 6 & appendix). In this paper, I have pointed out problems for -DV as well.

(e.g. I might be dreaming or deceived by a malicious demon), and a reason to suspend my judgment. In this sense, suspension of judgment as a mental act is not comparable to the act of quitting smoking. The decision to suspend judgment may be easy to make, but actual success is the result of arduous and attentive meditative practice, and even then, one can suspend judgment only temporarily.

The challenging nature of the suspension of judgment alludes also to another intrinsic aspect of the *Method of Doubt*. As Descartes describes the suspension to be arduous and difficult, it suggests that we should read him as being completely serious about the general suspension of judgment.<sup>30</sup> Based on this, the use of the method is not a purely hypothetical mind game. Descartes truly means that we should suspend judgment on all of our opinions and beliefs, as difficult as this may be. Suspension of judgment is therefore meant to be *psychologically real and genuine*<sup>31, 32</sup>

## References

- Adam, Charles – Tannery, Paul. (eds.) 1899–1908; 1963–1967. *Oeuvres de Descartes I–XII*. Paris: Vrin/C.N.R.S. (AT: 1899–1908; 1963–1967.)
- Cottingham, John – Stoothoff, Robert – Murdoch, Dugald. (eds.) 1984–1985. *The Philosophical Writings of Descartes*. Volumes I–II. Cambridge: Cambridge University Press. (CSM: 1984–1985.)
- Cottingham, John – Stoothoff, Robert – Murdoch, Dugald – Kenny, Anthony. (eds.) 1991. *The Philosophical Writings of Descartes*. Volume III – The Correspondence. Cambridge: Cambridge University Press. (CSMK: 1991.)
- Alanen, Lilli. 2003. *Descartes's Concept of Mind*. Cambridge: Harvard University Press.
- Alanen, Lilli. 2013. The role of will in Descartes' account of judgment. In: Detlefsen, Karen (ed.): *Descartes' Meditations: A Critical Guide*. Cambridge: Cambridge University Press. 176–199.
- Alquié, Ferdinand. 1950. *La découverte métaphysique de l'homme chez Descartes*. Paris: Presses Universitaires de France.
- Broughton, Janet. 2002. *Descartes's Method of Doubt*. Princeton: Princeton University Press.
- Carriero, John. 2009. *Between Two Worlds: A Reading of Descartes's Meditations*. Princeton: Princeton University Press.
- Cottingham, John. 1988. The Intellect, the will, and the passions: Spinoza's critique of Descartes. *Journal of the History of Philosophy* 26 (2): 239–257.
- Cottingham, John. 2002. Descartes and the voluntariness of belief. *The Monist* 85 (3): 343–360.
- Cunning, David. 2014. The First Meditation: divine omnipotence, necessary truths, and the possibility of radical deception. In: Cunning, David (ed.): *The Cambridge Companion to Descartes' Meditations*. Cambridge: Cambridge University Press. 68–87.

---

30 Cf. e.g. the *First Meditation* (Med. I, 1.: AT VII, 18; CSM II, 12) and the Preface (AT VII, 9; CSM II, 8).

31 Rosenthal (1986, 429–430) comes to a similar conclusion. For this, see also MacArthur (2003).

32 I would like to thank the audience and the organizers of the First Budapest Seminar in Early Modern Philosophy, as well as the editors of this volume. Special thanks to Olivér István Tóth.

- Curley, E.M. 1975. Descartes, Spinoza & Ethics of Belief. In: Freeman, Eugene – Mandelbaum, Maurice (eds.): *Spinoza: Essays in Interpretation*. La Salle: Open Court. 159–189.
- Della Rocca, Michael. 2006. Judgment and Will. In: Gaukroger, Stephen (ed.): *The Blackwell Guide to Descartes' Meditations*. Oxford: Blackwell Publishing. 142–159.
- Frankfurt, Harry G. 2008; 1970. *Demons, Dreamers, and Madmen: The Defense of Reason in Descartes's Meditations*. Princeton University Press, Princeton; New York – London: Garland.
- Gilson, Étienne. 1951. *Études sur le rôle de la pensée médiévale dans la formation du système cartésien*. Paris: J. Vrin.
- Gouhier, Henri. 1978. *La pensée métaphysique de Descartes*. Paris: J. Vrin.
- Grant, Brian. 1976. Descartes, Belief and the Will. *Journal of the Royal Institute of Philosophy* 51 (198): 401–419.
- Gueroult, Martial. 1953. *Descartes: selon l'ordre des raisons*. Vol. I. L'âme et Dieu. Paris: Aubier Éditions Montaigne.
- Kambouchner, Denis. 2008. Liberté et structure de l'âme. In: Kambouchner, Denis: *Descartes et la philosophie morale*. Paris: Hermann. 25–75.
- Kenny, Anthony. 1998; 1972. Descartes on the Will. In: Cottingham, John (ed.): *Oxford Readings in Philosophy: Descartes*. Oxford: Oxford University Press. 132–159.
- MacArthur, David. 2003. The Seriousness of Doubt and Our Natural Trust in the Senses in the First Meditation. *Canadian Journal of Philosophy* 33 (2): 159–181.
- Naaman-Zauderer, Noa. 2010. *Descartes' Deontological Turn: Reason, Will, and Virtue in the Later Writings*. Cambridge: Cambridge University Press.
- Newman, Lex. 2008. Descartes on the Will in Judgment. In: Broughton, Janet – Carriero, John (eds.): *A Companion to Descartes*. Oxford: Blackwell Publishing. 334–352.
- Perin, Casey. 2008. Descartes and the Legacy of Ancient Skepticism. In: Broughton, Janet – Carriero, John (eds.): *A Companion to Descartes*. Oxford: Blackwell Publishing. 52–65.
- Ragland, C.P. 2006. Is Descartes a Libertarian? In: Garber, Daniel – Nadler, Steven M. (eds.): *Oxford Studies in Early Modern Philosophy*. Vol. III. Oxford: Oxford University Press. 57–90.
- Rosenthal, David M. 1986. Will and the Theory of Judgment. In: Rorty, A.O. (ed.): *Essays on Descartes' Meditations*. Berkeley: University of California Press. 405–434.
- Schüssler, Rudolf. 2013. Descartes' Doxastic Voluntarism. *Archiv für Geschichte der Philosophie* 95 (2): 148–177.
- Shah, Nishi. 2002. Clearing Space for Doxastic Voluntarism. *The Monist* 85 (3): 436–445.
- Shapiro, Lisa. 2008. Turn My Will in Completely Opposite Direction: Radical Doubt and Descartes' account of Free Will. In: Hoffman, Paul – Owen, David – Yaffe, Gideon (eds.): *Contemporary Perspectives on Early Modern Philosophy: Essays in Honour of Vere Chappell*. Peterborough: Broadview Press. 21–39.
- Vitz, Rico. 2010. Descartes and the Question of Direct Doxastic Voluntarism. *Journal of Philosophical Research* 35: 107–121.
- Vitz, Rico. 2015a. *Reforming the Art of Living: Nature, Virtue, and Religion in Descartes's Epistemology*. Switzerland: Springer.
- Vitz, Rico. 2015b. Doxastic Voluntarism. *Internet Encyclopedia of Philosophy* (<http://www.iep.utm.edu/doxa-vol/>).

- Wee, Cecilia. 2014. The Fourth Meditation: Descartes and the libertarian freedom. In: Cunning, David (ed.): *The Cambridge Companion to Descartes' Meditations*. Cambridge: Cambridge University Press. 186–204.
- Wilson, Margaret. 1978. *Descartes*. London: Routledge.
- Williams, Bernard. 2015; 1978. *Descartes: The Project of Pure Enquiry*. London: Routledge.

# IV

*Hanna Vandenbussche*

## Descartes and Pascal on Imaging Love and Imaginary Love

What is the relationship between one's estimation of someone and love for that person, if any?<sup>1</sup> For instance: do love and esteem exclude each other or does love need to be based on reasonable esteem? In this contribution, this question will be engaged by means of an analysis of the respective views of love and imagination by René Descartes and Blaise Pascal.<sup>2</sup> The Cartesian account of love entertains numerous similarities with Blaise Pascal's approach of love. Some authors have stressed that both Descartes and Pascal adopt the same definition of love. For instance, according to Alberto Frigo (cf. Frigo 2016, 19), Pascal's definition of charity (the love of God) was directly inspired by the Cartesian doctrine of love.<sup>3</sup> For both Pascal and Descartes, (true) love indeed corresponds to a movement of the will in the direction of an object with which the soul desires unification.

In this paper, however, I will focus not upon the similarities in their definitions of love, but upon their respective views of the relationship between love and the imagination. While at first sight, these thinkers appear to adopt a radically opposed definition of

- 
- 1 An earlier version of this paper was presented at the *First Budapest Seminar in Early Modern Philosophy*. I am grateful for the organizers and participants for their comments and suggestions. Also, I am grateful to Dennis Vanden Auweele for his suggestions for improvement of the paper.
  - 2 References to Descartes follow the standard edition of Adam et Tannery: Adam, Charles – Tannery, Paul. (éd.) 1996. *Œuvres de Descartes*. XI volumes (I refer to this edition as follows: AT ..., ...). References to Pascal follow the standard edition of Louis Lafuma: Lafuma, Louis. 1963. *Blaise Pascal. Œuvres complètes* and Sellier: Sellier, Ph. (éd.) 2011. *Blaise Pascal. Pensées. Opuscules et lettres* (I refer to this edition as follows: L.../ S...). For Descartes' *Passions of the soul*, I used the translation of Stephen Voss: Voss, Stephen. 1989. *Descartes. Philosophical Letters*. For Descartes' Correspondence, I used the recent translation of Antony Kenny: Kenny, Antony. 1970. *The Passions of the Soul. René Descartes*. For Pascal's *Pensées*, I used the translation of W.F. Trotter: Trotter, W.F. 1978. *The Thoughts of Blaise Pascal*. The original French texts always appear in the footnotes.
  - 3 Frigo 2016, 19: "c'est en pensant l'amour et en le pensant à partir de la définition cartésienne, que Pascal parvient à décrire l'économie de l'ordre de la charité."

the imagination (as a faculty of the soul for Descartes and as a corrupted puissance for Pascal), Descartes' approach of the relation between love and the imagination reveals some similarities with Pascal's anthropological interpretation. First, both philosophers warn of the danger when the faculty of imagination intertwines with the passion of love through which it falsifies the representation of the beloved object. Towards the end of the *Passions of the Soul*, Descartes warns his readers about the dangerous effects of some passions; he simultaneously refers to the deceptive and seductive role of imagination:

What I think I can set down here as the most general remedy for all the excesses of the Passions and the easiest to put into practice is this: when one feels the blood stirred up like that, one should take warning, and recall that everything presented to the imagination tends to deceive the soul, and to make the reasons for favoring the object of its Passion appear to it much stronger than they are, and those for opposing it much weaker (Voss 1970, 134).<sup>4</sup>

Descartes ascribes a foundational role to the imagination when it comes to love. For instance, love as a passion urges the human soul to represent its object in an imaginary and often idealized way. Indeed, Descartes describes to Chanut how a soul that is driven by a passionate love easily imagines "lovable qualities in objects in which, at another time, it would see nothing but faults" (Kenny 1970, 210).<sup>5</sup> Such a mutual interpenetration between love and imagination can have immoral consequences. For instance, a passionate love easily leads to other, often immoral desires such as when the soul does not love the object as such, but only desires the possession of this object. Such possessive love leads to many of the egregious effects that love may have.

Despite his theological background and moral (thus non-scientific) approach to the passions in general, Pascal is equally aware of the potential dangers when a collaboration ensues between the passion of love and the imagination. In several fragments of his *Pensées*, he shows how the overwhelming force of the imagination can turn the most futile qualities – such as the length of a nose – into a primary cause of love.<sup>6</sup> Imagination is deceptively capable of providing material images and representations to which the soul, driven by a corrupted self-love, blindly consents. One of these representations will receive particular attention, namely the representation of the human self. According to Pascal, the

---

4 AT X, 487: "je pense pouvoir mettre ici comme le remède le plus général et le plus aisé à pratiquer contre tous les excès des passions, c'est que, lorsqu'on se sent le sang ainsi ému, on doit être averti et se souvenir que tout ce qui se présente à l'imagination tend à tromper l'âme et à lui faire paraître les raisons qui servent à persuader l'objet de sa passion beaucoup plus fortes qu'elles ne sont."

5 AT IV, 603: "cela fait que l'âme imagine des qualités aimables en des objets, où elle ne verrait que des défauts en un autre temps."

6 Of course, this is a reference to the fragment discussing the nose of Cleopatra as a cause of love. Cf. Laf. 197, S.163. In this fragment Pascal stresses the disproportion between, on the one hand, the causes, and on the other hand, the effects, of the passion of love. The most futile quality, such as the length of the nose, would have caused in Caesar his violent passion of love for Cleopatra and as a consequence the whole world has changed.

self appears to be an image motivated and reinforced by the passion of self-love. This passion as fundamental inclination of the human will, appeals to the imagination in order to represent a desirable image of the self. This leads to a fundamental form of self-deception, springing from a close collaboration between the imagination and the passion of love.

Does this then imply that Pascal and Descartes consistently associate love with the imagination negatively? Not necessarily: near the end of this paper, I show that there is a very specific form of love that does not face this negative evaluation, namely the love of God. Despite their seemingly opposite convictions about the human capacity to love God, both Pascal and Descartes invoke the imagination as an instrument to represent the disproportionate relation between the human soul and God.

Before attending to Pascal's and Descartes' views regarding the relationship between love and imagination, some clarification of their concepts is necessary. Pascal and Descartes appear to put forward radically distinct interpretations of the imagination: for Descartes, the imagination is a faculty of the soul that attends to cognitive functions, namely to provide concrete images of abstract concepts; for Pascal (like many other French moralists), imagination is a derailed capacity that overpowers reason. Given these vastly divergent definitions of the imagination, what use is a comparison? This is why the first part of this paper shows that the imagination is, for Descartes, not merely a neutral faculty of the soul, but she can – if entangled with the passion of love – give rise to egregious and dangerous consequences.

## **1. The imagination as a source of deception: the seventeenth-century moralists and Pascal**

By calling her mistress of errors (Pascal), source of folly (Nicole), and principle of human blindness (Malebranche), most of the French moralists and philosophers appear to express a very critical and even hostile attitude towards the imagination. Their negative views on the imagination can be explained by their Augustinian-Jansenist background; these authors explain the depravity of humankind by referring to the original sin: the choice of the first man, Adam, has brought the whole of humankind into a state of corruption. Most moralists consider the overwhelming force of the imagination as the clearest symptom of a corrupted human nature; the imagination is an independent force that opposes rationality (cf. Guenancia 2010, 87). Due to the unsettling force of the corrupted faculty of imagination, any and all proper relationship to truth and falsehood is reversed. For instance, Malebranche describes in *De la Recherche* how imagination is a serious obstacle in our quest for knowledge by being a symptom of the soul's dependence on the body.

A similar hostile attitude towards the imagination can be found in the work of Pascal. Recent commentary has pointed out (building particularly on Laf. 44) that Pascal's views have influenced Malebranche's theory of the imagination (cf. Frigo 2010, 518). While not mistaken, Pascal's diagnosis of the imagination surpasses in significant ways the general moralistic critique of the imagination as enemy of reason. More specifically, Pascal's theory of the imagination not only attends to an epistemological, but also to an anthropological and ontological problem (Bras-Cléro 1994, 11). For this reason, the relationship between the imagination and self-love in Pascal's anthropological thought should be stressed. Self-love is an effect of original *sin*, and refers to an infinite and disordered love for oneself. In our post-Lapsarian stage, the capacity of love is disordered because the First Man (Adam) desired to be 'as God' (*eritis sicut dei*) and resultantly turned himself away from the love of God. The imagination fulfills the role of providing desirable images to which a corrupted soul, driven by self-love, blindly consents. How far are we removed here from Descartes' definition of judgment as it appears in the *Fourth Meditation*? For Descartes, a truthful judgment consists in a movement of the will towards clear and distinct ideas of the intellect, and Pascal claims that human judgments result from a movement of a corrupted will towards confused images delivered by imagination.

At first sight, Pascal's theory of imagination appears to be radically opposed to Descartes' neutral and cognitive interpretation of imagination as a faculty of the soul: the latter does not consider it an independent force or 'puissance' that would meddle with all domains of human existence. In his *Regulae*, (and later in the *Sixth Meditation*), Descartes presents the faculty of imagination as an important expression of the human thinking capacity (*vis cognoscens*). The imagination creates material figures or images out of abstract conceptions from the intellect, by which it shows herself to be dependent on both body and soul. However, the function of imagination has its limits: not every mathematical definition of the intellect can be represented by imagination. In his *Sixth Meditation*, Descartes refers to the example of the chiliagon, a thousand-angled mathematical figure, in order to stress the difference between the intellect and the imagination. While our intellect can understand a chiliagon (we know what the word means), our imagination fails in its capacity to render a thousand-angled figure present to our minds. The abstract conception of the chiliagon can only arise at best as a confused image in our mind. Despite their fundamental distinction, Descartes never opposes the intellect and the imagination to each other as hostile faculties. In the epistemological and metaphysical domains, as far as the human soul is known by the primitive notion of pure thinking capacity (*mens*), imagination fulfills a function complementary to the intellect.

Nonetheless, imagination does threaten at times to drift from its epistemological function by collaborating closely with certain passions.<sup>7</sup> Those passions are confused thoughts which, just as clear and distinct ideas, appear in the human soul. But unlike clear and distinct ideas, the passions do not faithfully represent truth or objects: every representation

---

<sup>7</sup> Imagination thus does not always fulfill a purely positive function. As Frigo puts it: "l'imagination est en même temps utile et dangereuse". Frigo 2010, 518.

of a passion is originally destined to inform the human soul about the good or harmful character of external objects for the body.<sup>8</sup> The representations of the passions only serve a biological purpose, instituted by a reliable God (*Institution de la Nature*): God establishes that the passions inform the human soul about the good or bad character of objects with regard to the human body.<sup>9</sup>

Deception happens when the human soul forgets the original, biological significance of the passions and is happily inclined to consent to the confused representation of a passion, especially because of its vivid and intense character. This is because a passion is always associated with a representation in the soul that presents an object to be more important than it is in reality:

They almost always make both the goods and the evils they represent appear much greater and more important than they are, so that they incite to seek the former and flee the latter with more ardor and more anxiety than is suitable (Voss 1989, 93).<sup>10</sup>

Love is a concrete example of Descartes' view of the dangers coming from a close collaboration between imagination and a passion.

## 2. Descartes on the passion of love and its relation to the imagination

In his letters to Pierre Chanut,<sup>11</sup> René Descartes offers a multi-faceted and complex philosophical view of the essence of love. Regrettably, the complex character of Descartes' account of love is easily missed. He defines love as a movement of the will in the

---

8 This is why Descartes' approach to the passions starts from a physical point of view: every passion has a corporeal origin, namely as a certain movement of the life spirits which moves the pineal gland and transforms itself into a concrete passion in the soul. Cf. AT XI, 353: "Et on peut aisément concevoir que ces images ou autres impressions se réunissent en cette glande par l'entremise des esprits qui remplissent les cavités du cerveau." Voss 1989, 37: "And we can easily understand these images or other impressions to unite in this gland by the mediation of the spirits filling the brain's cavities."

9 AT XI, 357: "(...) ils (les esprits animaux) excitent un mouvement particulier en cette glande, lequel est institué de la nature pour faire sentir à l'âme cette passion." Voss 1989, 39: "Simply in virtue of entering these pores, these spirits excite a particular movement in this gland which is instituted by nature to make the soul feel this passion."

10 AT XI, 431: "elles nous représentent tant les biens que les maux beaucoup plus grands et plus importants qu'ils ne sont, en sorte qu'elles nous incitent à rechercher les uns et fuir les autres avec plus d'ardeur et plus de soin qu'il n'est convenable."

11 Pierre Chanut was an appointed diplomat at the Swedish court and a good friend of Descartes. Unlike many other intellectual correspondents of Descartes, Chanut was not really concerned with scientific issues but expressed a deep interest in morality. For further information, see Clarke (2006).

direction of an object that appears convenient or agreeable to the soul. As a passion, love accomplishes a bodily function, in that love inclines the will towards an object that appears useful. As a result, some might find – especially those who are somewhat romantically-minded – that Descartes’ account of love is somewhat disappointing since he views this passion as a physician: love is rooted in a biological process. In order to explain the biological origin of love, Descartes refers to the very beginning of human life: the most primitive form of love consists in the desire of the unborn fetus to unite with the nutrients in the mother’s milk. With the French philosopher Alain (born as Émile-Auguste Chartier), we could call the Cartesian passion of love “a hymn to mother’s milk” (Alain 1925, 175). Very generally, Descartes’ account of the passion of love tends to be interpreted as nothing more than a mere scientific study of a certain biological phenomenon.

According to Descartes, love entails an incomprehensible connection between an indifferent physical process and a passion in the soul. The passion of love is originally rooted in an indifferent, bodily movement or, more specifically, in a certain stream of life spirits in the body. Therefore, passionate love always appears as a confused thought in the mind: the passion has nothing to do with the intrinsic value of the beloved object. A biographical example from Descartes’ early childhood might be illuminating here: ‘the cross-eyed girls’. As a child, Descartes felt strangely attracted to cross-eyed girls. Without being able to give a reason why, he felt impressed by the singular property of being ‘cross-eyed’. As an adult, Descartes offers a dry, physical explanation for this strange, youthful attraction: the first time he met a cross-eyed girl, he was so impressed that the sight of this girl, it left an impression in his brain. Every time he saw a cross-eyed girl this trace was reactivated, as a result of which poor René unwillingly lapsed into love.

Of course, Descartes’ predilection for cross-eyed girls is quite innocent and certainly does not arouse moral condemnation. Descartes could not control the physical movements triggered by the pleat in his brain. However, the passion of love, rooted in a physical movement, which may appear as a thought in the soul can become excessive and even immoral. By this, it is meant that the soul is inclined to consent blindly to the images this passion evokes. Some passionate lovers spend all their energy and efforts embellishing and idealizing the image of their beloved. Ascribing a disproportionate value to the object of their love, they become obsessed with an image and invent all kinds of qualities and reasons that justify their love. Sometimes the illusion even expels the real person, and one thus falls in love with an image. At this point, love becomes an imaginary passion and therefore causes a serious form of alienation: the fascination for the image of the beloved becomes an obsession ousting every other idea of the soul. This alienation does not necessarily imply a love directed towards a person since objects, whatever their form may be, can give rise to the passion of love as well. Curiously, Descartes juxtaposes moral objects (a father loving his ‘son’/ someone loving his ‘friend’) and immoral objects (sex, alcohol, money):

For, e.g., although the passions an ambitious person has for glory, an avaricious person for money, a drunkard for wine, a brutish man for a woman he wants to violate, a man of honor for his friend or his mistress, they are nevertheless similar in that they participate in Love. But the first four have Love only for the possession of the objects their passion has reference to, and have none whatever for the objects themselves, for which they have only desire mixed with other particular passions (Voss 1989, 63).<sup>12</sup>

However divergent and arbitrary the nature of these objects may be, they all participate in the passion of love. But despite the fact that Descartes contends that all these (moral and immoral) kinds of love have the same essence, he still insists upon a difference with regard to the relation these objects have to the loving soul. A miser desiring money or a rapist longing for a female body do not love the objects (women, money) *in themselves*: they only desire the possession of these objects. In other words, all these lovers believe (wrongly) that they need to usurp or conquer the objects of their love in order to satisfy their desires. Descartes calls this phenomenon a love for possession.

As soon as the passion of love derails and becomes an obsession, it can give rise to even more dangerous consequences. These consequences can even be more dire than those of the passion of hatred. Descartes surprisingly claims that an excessive love can cause the most pernicious evil of humanity. This is not necessarily so because love, unlike hate, conjoins with much strength and audacity (do not all deeds of heroism spring from love?), but because this passion entails a blind and wrong judgment of its object. In other words, if the passion of love corresponds to a union of the soul with an object that is faulty and excessively valorized by the imagination, this love is uncompromising: every other object that can be an obstacle for the unification with the object is destroyed in the most merciless way. The love for one's object brings along hate for an infinite series of other objects. "The greatest and most tragic disasters" according to Descartes, "can be seasoning for a disordered love" (Kenny 1970, 218).<sup>13</sup>

Strikingly, Descartes' characterization of these dangerous effects that arise from excessive and unreasonable love is similar to Pascal's anthropological approach of love. Despite his radically different background – he approaches the passions not as a physician, but as

---

12 AT XI, 388–389: "(...) encore que les passions qu'un ambitieux a pour la gloire, un avaricieux pour l'argent, un ivrogne pour le vin, un brutal pour une femme qu'il veut violer, un homme d'honneur pour son ami ou pour sa maîtresse, et un bon père pour ses enfants, soient bien différentes entre elles, toutefois en ce qu'elles participent de l'amour elles sont semblables. Mais les quatre premiers n'ont de l'amour que pour la possession des objets auxquels se rapporte leur passion et n'en ont point pour les objets mêmes."

13 Cf. AT IV, 616: "Mais l'amour est toujours plus coupable que la haine des maux qu'on attribue à l'amour, pour ce que, si nous aimons quelque chose, nous haïssons, par même moyen, tout ce qui lui est contraire."

a moralist – Pascal also pays attention to the immoral dangers of a cooperation between the imagination and the passion of love.

### 3. Pascal on the imagination and the passion of love

Unlike Descartes, Pascal discusses the nature of love as a moralist and does not provide a physiological description of this passion. Influenced by an Augustinian background further molded by Jansenism, Pascal accounts for the corrupted character of human love by reference to original sin. As a consequence of Adam's choice to become as God, humanity was brought into a permanent state of corruption, and was no longer capable of loving appropriately (Wood 2013, 20). Because of this, error, misery and deception became the main characteristics of a human life.

In Pascal's anthropology, passions never have a useful biological goal but are the representatives *par excellence* of the corruption of human nature. In several of his *Pensées*, Pascal explicitly associates the passion of love with desires for domination. A soul that loves an object or a person does not do so for its real value, but only desires to conquer or absorb the object. Pascal's description of human love is accordingly highly similar to Descartes' descriptions of the love of possession:

Mine, thine. This dog is mine, aid these poor children; 'That is my place in the sun.' Here is the beginning and the image of the usurpation of all the earth (Trotter 1978, 105).<sup>14</sup>

According to Pascal, love expresses itself through the desire to dominate the other and therefore Pascal explicitly associates human love with such things as self-interest and usurpation. Every human being tends to behave as a tyrant by considering himself a god, thereby transgressing the proper ontological and cosmological order.<sup>15</sup> As a consequence, every relation in society seems to hide a fondness of tyrannical desires and hate towards the others:

---

14 Laf. 64, S. 295: "Mien, tien. 'Ce chien est à moi,' disaient ces pauvres enfants. 'C'est là ma place au soleil.' Voilà le commencement et l'image de l'usurpation de toute la terre."

15 Pascal invokes here repeatedly the metaphor of the center. The self functions in his anthropology as a magnetic field that attracts all kinds of objects which could respond to his insatiable desire of possession. The proper value of these objects, however, is never taken into consideration: they only need to fulfill the insatiable desires of possession, proper to the self. Cf. Guenancia 2000, 287: "le moi pascalien fait dépendre sa satisfaction de la consommation des choses qu'il désire, qui ne seront d'ailleurs jamais suffisantes pour produire en lui un contentement ou une satisfaction au sens propre du terme."

In a word, the Self has two qualities: it is unjust in itself since it makes itself the center of everything; it is inconvenient to others since it would enslave them; for each Self is the enemy, and would like to be the tyrant of all others (Trotter 1978, 153).<sup>16</sup>

Strikingly, Pascal also explicitly associates the passion of love with the imagination. He namely claims that human love originates in futile and insignificant qualities which are valorized by the imagination. When, for instance, Pascal discusses the causes and the effects of love, he invokes the length of the nose as a decisive quality on the basis of which love has emerged. Indeed, according to Pascal, the length of Cleopatra's nose would have caused the love of Caesar and, as a consequence, changed the world as we know it.

The close collaboration between the imagination and the passion of love in Pascal's anthropology is clearly illustrated by the way in which a human being relates to himself. Indeed, the human self is nothing more than the expression of a specific (and immoral) form of love, namely self-love. In the fragment 'What is the self', Pascal evokes the scene of someone passing by and wondering if the unknown man at the window is watching him in particular.<sup>17</sup> In other words, this person wants to be seen, a kind of visual spectacle, by some unknown spectator standing at the window. Unfortunately, this does not seem the case and the person needs to conclude that he is not the center of attention and that the one staring at the window is not thinking of him in particular.

This desperate conclusion leads to a restless succession of questions which express the frantic efforts of the soul to be recognized and loved as a singular and remarkable object in the eyes of the other: "And if someone *loves* me for my judgment or my memory, do they love me? Me, myself? (...) And where then is this self, if it is neither in the body nor the soul?" Finally, the fragment ends in *aporia* as the self appears to escape every possible determination, being invisible, unfindable and (even worse) incapable of being loved. The only thing we know for sure about its nature is that the self can only be loved for its perishable and contingent qualities which are valorized by the imagination. According to Pascal, thinking the self is intrinsically allied with a desire of being loved as a self.

---

16 Laf., 597, S. 455: "En un mot le moi a deux qualités. Il est injuste en soi en ce qu'il se fait centre de tout. Il est incommode aux autres en ce qu'il les veut asservir, car chaque moi est l'ennemi et voudrait être le tyran de tous les autres. Vous en ôtez l'inconfort, mais non pas l'injustice."

17 Several secondary commentators stretch the similarity between the scene and the decor evoked in the *Second Meditation* of Descartes. For instance, in his work *L'invention du moi*, Vincent Carraud establishes a very detailed comparison between the self figuring in the fragment 'Qu'est-ce que le moi' and the Cartesian ego of the *Second Meditation*. la Carraud 2010, 61: "C'est peut-être même la lecture de la *Méditation II* et le retournement de situation qu'elle permet qui sert d'arme à Pascal contre ce qui semblait être le premier résultat du *Discours*, faisant rejaillir sur celui-ci et son « ce moi » la critique invalidante de l'ego ille des *Méditations*." Also in the Anglo-Saxon literature, we see a similar tendency to approach these fragments. Cf. Moriarty 2003, 146: "Just as Descartes withholds recognition from the passers-by, Pascal's story of the man at the window is also a story of recognition." Or Wood 2013, 96: "In Descartes, the self is the watcher at the window, the one who melts the wax, (...) In Pascal, the self is watched from the window, and its qualities are progressively stripped away as if it were the wax."

Driven by self-love, every human being evokes a fictitious image of himself by means of which he/she can be recognized in the eyes of the other. In this context, self-love can be understood analogously with what Jean-Paul Sartre calls an imaginary passion, which is a passion that gives birth to an imaginary object out of nothing. The human being constantly embellishes his imaginary self, looking for reasons which justify his self-love. This process can be compared with a person who has fallen obsessively in love and consequently calls to mind all sorts of nice qualities relating to his beloved by means of which he creates an image of her in her absence. But this search for qualities in order to conjure the imaginary self appears to be endless and quite desperate, moreover; it leads to a permanent state of restlessness. Every quality that it attributes to the image of the self is ultimately contingent, perishable, vain and insignificant. Therefore, every human being needs to be permanently looking for other reasons, merits and qualities in order to embellish his imaginary self and in the meanwhile to justify his self-love:

We do not content ourselves with the life we have in ourselves and in our own being; we desire to live an imaginary life in the mind of others, and for this purpose we endeavor to shine. We labor unceasingly to adorn and preserve this imaginary existence, and neglect the real (Trotter 1978, 59).<sup>18</sup>

Do these pessimistic descriptions of a depraved human love accompanied with self-deception imply that for Pascal, no authentic love is possible anymore? At first sight, this seems to be the case since, because of his corrupted nature, the human being seems to be condemned to an existence of vanity where the imagination and the passions are dominant. However, despite his pessimism with regard to *human* love, Pascal also refers to another form of love, namely the love for God. Curiously Pascal appeals to a certain use of the imagination in order to represent this spiritual form of love. More precisely, he refers to a specific image that renders the (disproportionate) relation between the human being and God more concrete: the image of the body full of thinking members.

Strikingly, this function of the imagination to represent the disproportionate relation between the human soul and God, was already present in Descartes' description of the love for God. The love of God is a love that consists in a specific act of thinking, namely the imagining of a whole from which the soul only is a certain part (Frigo 2016, 159). Despite their seemingly opposite convictions about the *human capacity* to love God, both Pascal and Descartes invoke the importance of the imagination in order to represent the disproportionate *relation* between the human soul and God.

---

18 Laf. 806, S. 147: "Nous ne nous contentons pas de la vie que nous avons en nous et en notre propre être. Nous voulons vivre dans l'idée des autres d'une vie imaginaire et nous nous efforçons pour cela de paraître. Nous travaillons incessamment à embellir et conserver notre être imaginaire et négligeons le véritable."

## 4. Descartes and Pascal on the imagination and the love of God

Descartes describes the love of God as “the most delightful and useful passion possible” (Kenny 1970, 212).<sup>19</sup> How is the love of God a passion and what role is played by imagination in this context?

It is generally conceived that the love of God is spiritual and intellectual, which means that God can never be represented truthfully by the faculty of imagination. Descartes defines intellectual love as a movement of the will towards an object from which the soul possesses a complete knowledge. The soul then considers itself and the good as two parts of a single whole.<sup>20</sup> The movement of the will towards the clear and distinct ideas, (as it is shown in the *Fourth Meditation*), could be conceived as the most clear example of an intellectual love independent from every influence of the body. From a metaphysical point of view, the soul is conceived according to the primitive notion of ‘mens’.

At the beginning of his letter to Chanut, however, Descartes claims that, with regard to human life (where the soul is conceived in its union with the body), intellectual love and passionate love necessarily co-exist.<sup>21</sup> Every human soul is in a close union with his body. This means that even when love is based on a reasonable judgment, this always engenders the movement of the life spirits in the body.<sup>22</sup> So even the noblest intellectual love, the love of God, will necessarily co-exist with a passionate form of love. But of course, the love of God first appears as a purely intellectual love in that a human soul can love God by means of pure thinking. God can then be conceived by means of a pure intellectual reflection. In order to acquire this intellectual conception of God, the soul needs to consider God in a specific way. More precisely, the soul is required to represent God as an infinite thinking substance, of which the human soul constitutes a little particular aspect. The joyful contemplation of God’s attributes in which the finite human soul takes part, refers to a relationship of radical disproportion: through the reflection upon God’s infinity, the soul becomes aware of his own finiteness. The awareness of this disproportion gives rise to the

---

19 AT IV, 608: “J’ose dire au regard de cette vie, c’est la plus ravissante et la plus utile passion que nous puissions avoir; et même qu’elle peut être la plus forte.”

20 Kenny 1970, 208: “The first (intellectual love), in my view, consists simply in the fact that when our soul perceives some present or absent good, which it judges to be fitting for itself, it unites itself to it in volition, that is to say, it considers itself and the good in question as forming two parts of a single whole.” AT, IV, 601: “La première est, ce me semble, autre chose sinon que, lorsque notre âme aperçoit quelque bien, soit présent, soit absent, qu’elle juge lui être convenable, elle se joint à lui de volonté, c’est-à-dire, elle se considère soi-même avec ce bien-là comme un tout dont il est une partie et elle l’autre.”

21 He only mentions two (rare and accidental) exceptions: on the one hand, the soul can experience a passionate love without knowing why he loves this particular object and, on the other hand, the soul can esteem an object to a very high extent while not feeling any corporeal excitement or passion for this object. Cf. AT IV, 603.

22 As Denis Kambouchner puts it: “With regard to this life, the soul never loves alone, but rather loves with the body it is joined to.” Kambouchner 2008, ‘Cartesian subjectivity and love.’ 31.

attitude of humility, since the soul forgoes the ambition to replace God through experiencing gratitude for its dependence upon God's perfection. Moreover, the insight in God's infinity and one's own finiteness does not evoke any feelings of powerlessness or anxiety. On the contrary, the soul experiences an intense and extreme feeling of joy because of the realization of its dependence upon God.

The question remains how these intellectual forms of love and joy turn into passions. At this point, Descartes will invoke the faculty of imagination. Of course, the imagination will serve to represent God, by reducing God to an object that can be loved by the soul.<sup>23</sup> While the reflection upon God's infinity excludes every influence of the imagination (the risk of idolatry), the soul nonetheless appeals to the imagination in order to represent his *union* of dependency with God. Indeed, the soul, as being closely united with the body, will spontaneously imagine his disproportionate relation with God. This image of the union between the soul and God in turn gives birth to a mechanical movement in the body from which arises the violent passions of love and joy. The joy that arises from the consideration of one's union with God does not have a pure intellectual character only: the emotion tends to prove the presence of the body. The passions, which arise from this image of the union between the soul and God, tend to surpass their natural and biological finality: they have a pure disinterested nature. In this context, the imagination renders the abstract reflection upon the disproportionate relation between the soul and God more concrete by providing an image of this union. More precisely, it offers a concrete image of the union between the soul and God, as part of an infinite whole.

Strikingly, Pascal also appeals to the imagination, and more precisely to its function of representing a disproportionate relation, in his view, of the love of God. Despite the fact that the image of the union between the human soul and God never gives rise to passions of love and joy for Pascal, the act of imagining the relation of the soul towards God (as a part of a whole) is an important step towards an authentic morality. In the preceding part, we have seen how the interdependency between human love and the imagination expresses a profound pessimism in Pascal's anthropological thought: human love as sustained and reinforced by some images provided by the imagination, necessarily occurs with (self-)deception and an immoral attitude towards the other. Nonetheless, Pascal did write the *Pensées* from an apologetic point of view; this means that his ultimate aim is to allude to the possibility of escaping one's misery next to his emphasis on the miserable and the desperate character of the human condition. Pascal honestly believes that this possibility is offered by the Christian religion. Morality only truly becomes possible by considering the values and the truth of Christianity. However, a real conversion towards Christianity cannot be reached by merits alone since God bequeaths his grace arbitrarily to some elected human beings: "man is only a subject full of error, natural and ineffaceable, without grace. Nothing shows him the truth" (Trotter 1978, 36).<sup>24</sup> The question arises then whether the

---

23 Guenancia 2010, 108: "Dans le cas de l'amour de Dieu, l'imagination ne se fait pas un Dieu à sa mesure pour pouvoir l'aimer, et, du même coup, s'aimer à travers lui."

24 Laf. 45, S. 83: "L'homme n'est qu'un sujet plein d'erreur naturelle, et ineffaçable sans la grâce. Rien ne lui montre la vérité."

love of God is necessary to live a good and happy life, given the fact that one is never sure of receiving divine grace. Pascal deals with this question in a variety of different ways. His most famous answer can be found in the fragment *Infini Rien*, namely in the conclusion of his well known wager. Pascal argues that if one acts according to the rituals and habits of Christianity, one might naturally start believing. In other words, by imitating the acts of a real and converted Christian, one spontaneously will lead a moral and meaningful life.<sup>25</sup>

Besides the wager-argument, Pascal also deals with this question in another way. Namely, he compares the human capacity for loving God with the Paulinian image of the body full of thinking members. This image refers to a whole theological tradition in order to reflect upon the body of Christ as the symbol of the Church.<sup>26</sup> Pascal, largely inspired by this Paulinian metaphor from 1 Corinthians 12 (Wood 2013, 220), appeals to this image in order to explain how one can regulate one's vicious self-love in order to regain an authentic love for God:

To regulate the love which we owe to ourselves, we must imagine a body full of thinking members, for we are members of the whole, and must see how each member should love itself, etc. (Trotter 1978, 158).<sup>27</sup>

Pascal then appeals to a specific use of the imagination in order to render the nature of this love more concrete. The imagination no longer refers, in this context, to an overwhelming force that is capable of producing all kinds of (untruthful) images to which the corrupted will consents. When imagination is no longer in such a close relation to self-love, but in relation to the love of God, it fulfills a positive function:<sup>28</sup> the imagination can provide metaphors to represent an incomprehensible reality. While, however, the imagination does not picture the essence of God (this would be idolatry), she does seem capable of producing an image in order to present the relationship between the human soul and God. This

---

25 "Learn of those who have been like you, and who now stake all their possessions. These are people who know the way which you would follow, and who are cured of an ill of which you could be cured. Follow the way by which they began; by acting as if they believed, taking the holy water, having masses said, etc." (Trotter 1978, 86). Laf. 418, S. 233: "(...) apprenez de ceux, etc. qui ont été liés comme vous et qui parient maintenant tout leur bien. Ce sont gens qui savent ce chemin que vous voudriez suivre et guéris d'un mal dont vous voulez guérir; suivez la manière par où ils ont commencé. C'est en faisant tout comme s'ils croyaient, en prenant de l'eau bénite, en faisant dire des messes, etc. Naturellement même cela vous fera croire et vous abêtira. (...) Or quel mal vous arrivera(-t)-il en prenant ce parti? Vous serez fidèle, honnête, humble, reconnaissant, etc."

26 For a recent study which elaborates the theological background of this image, we refer to the book of Alberto Frigo, *L'esprit du corps*.

27 Laf. 368, S. 474: "Pour régler l'amour qu'on se doit à soi-même il faut s'imaginer un corps plein de membres pensants, car nous sommes membres du tout, et voir comment chaque membre devrait s'aimer, etc."

28 Pierre Guenancia, in his work *Divertissements pascaliens*, explicitly refers to this positive use of imagination. He argues that the imagination fulfills in this context an instrumental role with regard to the intellect by representing the love for God as the relation between a part and a bigger whole (Guenancia 2011, 178).

image is the (spiritual) body full of thinking members. In order to emphasize the distance between 'meaningless material images' and this spiritually significant image, Pascal uses the term 'figure' rather than 'image' (cf. Guenancia 2011, 172).

How can the precise function of this image, *le corps des membres pensants*, be understood? The fragments in the *Pensées* that deal with this metaphor all relate to the issue of regulating disorderly self-love in order to prepare for the love of God. This means that this image should be conceived of as a thought-experiment (cf. Wood 2016, 219) by means of which we resist our innate self-love and diminish our passions. Pascal argues that self-love only disappears by the ready recognition of one's dependence upon God, and because of this one is required to appeal to the image of a body on which every human being depends in terms of membership. Self-love precisely consists in the refusal to recognize one's dependence on the body as a whole. Therefore, Pascal refers to the soul possessed by self-love, the hatred self, by invoking the image of a separated member. This is a soul that is driven by self-love and that takes himself to be a center or a whole upon which everything depends:

The separate member, seeing no longer the body to which it belongs, has only a perishing and dying existence. Yet it believes it is a whole, and seeing not the body on which it depends, it believes it depends only on self, and desires to make itself both center and body (Trotter 1979, 161).<sup>29</sup>

Through the act of imagining of oneself as a member depending on God as a body, one discovers the depraved, unjust and interested character of his self-love. As a consequence, by means of this metaphorical reflection, one starts to hate the interested desires of power and domination that accompany self-love.

The insight that arises from this image of the body full of thinking members thus gives rise to a disposition for authentic and spiritual love. This disposition to love God, however, does not correspond to an active movement of the will, nor to a useful passion as Descartes puts it. Pascal indeed would never speak in terms of an active movement of the will that gives rise to passions of joy and love. On the contrary, for him true love arises from a (passive) disposition that is made possible by the process of destroying our self-love. For Pascal, the love for God only happens through the gift of divine grace. Nonetheless, he refers to a specific use of the imagination in order to prepare the human soul to receive this grace. The image of the 'body of thinking members' is a means to recognize the soul's dependence upon the whole of God. This part-whole metaphor also played an important role in Descartes' view of the love of God.

---

29 Laf. 372, S. 483: "Le membre séparé ne voyant plus le corps auquel il appartient n'a plus qu'un être périssant et mourant. Cependant il croit être un tout et ne se voyant point de corps dont il dépende, il croit ne dépendre que de soi et veut se faire centre et corps lui-même."

# Conclusion

Descartes and Pascal entertain a complex view with regard to the relationship between love and the imagination. Not only has it been shown that both thinkers condemn the close collaboration between a passionate form of love and the imagination, but also it has become clear that both thinkers appeal to the imagination through their approaches of a specific kind of love, namely the love of God.

Both Descartes and Pascal warn their readers of the immoral consequences of an unfettered love in close collaboration with the imagination. We have discovered how Descartes warns of the deceptive and misleading force of the imagination when he deals with the passion of love: a soul often represents his beloved object in an imaginary and ideal way. In his letter to Chanut, Descartes explains how an excessive form of love not only evokes error and deception, but can lead also to evil and catastrophe. Strikingly, Pascal also alludes to the dangers related to a mutual dependency between human love and imagination. For instance, the hatred self incessantly appeals to the force of imagination in order to over-valorize some (vain) qualities he ascribes to himself. From this follows that Pascal, despite the fact that he approaches both the imagination and the passions from a theological point of view (considering them as symptoms of original sin), seems to adopt Descartes' viewpoint on the misleading role of imagination with regard to the passion of love.

What is more, both philosophers analyze the specific case of the love of God by ascribing a positive role to the imagination. For Descartes, the imagination fulfills the role of figuring the disproportionate relation between the finite human soul and the God's infinity. This image leaves in turn an impression in the brain that causes a movement of life spirits in the body. This gives rise to the most violent and useful passions a human soul can experience, namely love and joy. Pascal, however, will interpret the positive function of the imagination with regard to the love of God in a different way. He claims that a human being can never reach the love of God by his own efforts: he can only dispose himself towards the gift of divine grace, which is the ultimate condition of an authentic love. This disposition, however, requires a certain use of the imagination: in order to recognize his dependence upon God, the soul needs to imagine himself as a member relating to a body as a whole. The Paulinian 'metaphor' of "the body of thinking members" serves therefore a very important moral goal: the soul diminishes all the vicious desires and passions that accompany his self-love and disposes himself to be affected by the love of God.

## Bibliography

### Standard editions

Adam, Charles – Tannery, Paul. (éd.) 1996. *Œuvres de Descartes*. XI volumes. Paris: Vrin.

Alquier, Ferdinand. (éd.) 2010. *René Descartes. Œuvres philosophiques*. Tome III. 1643. Paris: Éditions Classiques Garnier.

- Lafuma, Louis. (éd.) 1963. *Blaise Pascal. Œuvres complètes*. Paris: Seuil.
- Sellier, Philippe. (éd.) 2011. *Blaise Pascal. Pensées. Opuscules et lettres*. Paris: Éditions Classiques Garnier.

### Translations

- Kenny, Antony. 1970. *Descartes. Philosophical Letters*. Oxford: Basil Blackwell.
- Trotter, William Finlayson. 1978. *Pascal, Thoughts*. Westport: Greenwood Press.
- Voss, Stephen. 1989. *The Passions of the soul. An English Translation of Les Passions de l'Âme*. Indianapolis / Cambridge: Hackett Publishing Company.

### Secondary studies

- Alain, Emile Auguste Chartier. 1925. *Propos sur le Bonheur*. Paris: Gallimard (Folio Essais).
- Carraud, Vincent. 2010. *L'invention du moi*. Paris: Presses Universitaires de France.
- Clarke, Desmond M. 2006. *Descartes. A Biography*. Cambridge: Cambridge University Press.
- Bras, Cléro – Gérard, Jean-Pierre. 1994. *Pascal. Figures de l'imagination*. Paris: Presses Universitaires de France.
- Frijo, Alberto. 2016. *L'esprit du cœur*. Paris: Vrin.
- Frijo, Alberto. 2016. A very obscure definition: Descartes's account of love in the Passions of the Soul and its scholastic background. *British Journal for the History of Philosophy* 24 (6): 1097–1116.
- Guenancia, Pierre. 1998. *L'intelligence du sensible*. Paris: Gallimard.
- Guenancia, Pierre. 2000. *Lire Descartes*. Paris: Gallimard (Folio Essais).
- Guenancia, Pierre. 2010. *Descartes. Chemin Faisant*. Paris: Éditions Les Belles Lettres.
- Guenancia, Pierre. 2011. *Divertissements pascaliens*. Paris: Éditions Hermann.
- Kambouchner, Denis. 1995. *L'homme des Passions. Commentaires sur Descartes (II)*. Paris: Albin Michel.
- Kambouchner, Denis. 2008. *Descartes et la philosophie morale*. Paris: Hermann.
- Kambouchner, Denis. 2008. Cartesian subjectivity and love. In: De Dijn, Herman – Boros, Gábor – Moors, Martin (eds.): *The concept of love in seventeenth century philosophy*. Leuven: Leuven University Press. 23–42.
- Marion, Jean-Luc. 2013. *Sur la pensée passive de Descartes*. Paris: Presses Universitaires de France.
- Moriarty, Michael. 2003. *Early Modern French Thought. The Age of Suspicion*. Oxford: Oxford University Press
- Wood, William. 2013. *Blaise Pascal on Duplicity, Sin and the Fall. The secret instinct*. Oxford: Oxford University Press.

# V

*Judit Szalai*

## Medicine, Emotion Management and Mind-Body Interaction in Later Descartes and Early Cartesianism<sup>1</sup>

The view that Descartes changed his mind concerning the relationship between mind and body in the period of his correspondence with Elizabeth has been voiced by some prominent Descartes scholars, among them Stephen Gaukroger and Lisa Shapiro.<sup>2</sup> His correspondence with Elizabeth is thought to present us with a different Descartes, one who is both metaphysically more realistic, who acknowledges the fact and implications of embodiment, and is more useful as a health advisor, with a non-mechanist approach to diseases. In this paper, I try to show that there is in fact substantial continuity within Descartes' work concerning the functional aspects of the mind-body relationship, as well as from a medical point of view. Further, I will trace the way in which certain elements of these core views were retained and developed in some first- and second-generation Cartesian authors.

### **I. Descartes' conception of health and decay**

To set the stage, I will give an outline of Descartes' theoretical views on health, which, as it will emerge during this discussion, are pretty consistent (and consistently ambiguous) throughout his work. It is well-known that Descartes ranked the development of medicine very high among the goals of science. In the "Preface to the French edition" of the

---

1 This paper was prepared with the support of OTKA projects K 120375 and 125012.

2 I quote a familiar formulation, by Shapiro: "The line Elisabeth presses shows Descartes that [...] he need not deny that the way we find ourselves in the world will very much affect our thought... In the *Passions*, more than any other of his [works], Descartes respects the fact that we are embodied". Shapiro 1999, 516. Gaukroger claims that "the correspondence brings to light a very significant change of focus in Descartes's thought". Addressing the theme of the passions called for "the appropriate notion of a substantial union needed to account for the source and nature of affective states". Gaukroger 1995, 398–399.

*Principles*, we read that medicine is one of the branches that yield the “principal benefit of philosophy” (AT IX B 14–15 / CSM 1.186). In the *Discourse on the Method*, we are told that Descartes intended to devote his life to its pursuit (AT VI 63 / CSM 1.143 and AT VI 78 / CSM 1.151), which he confirmed in a letter dated much later.<sup>3</sup> Not only did Descartes think it important to advance medicine; he had an almost boundlessly ambitious vision for it: “we might free ourselves from innumerable diseases, both of the body and of the mind, and perhaps even from the infirmity of old age, if we had sufficient knowledge of their causes and of all the remedies that nature has provided” (AT VI 62 / CSM 1.143. Cf. AT XI 223–224 / CSM 1.314).

Descartes’ medicine, insofar as it is concerned with bodily diseases with physical causes, is bound to be mechanistic; no doubt, this is the default case. In the *Discourse*, Descartes talks about deriving rules of medicine from the knowledge of nature (AT VI 78 / CSM 1.151). Since he understands the human body as a machine, it is natural to interpret his conception of health and sickness accordingly, and perhaps even to consider this as extending to mental processes as well.<sup>4</sup>

The human body-machine idea remained with Descartes into his later years. He offers analogies on several occasions between the human body and a machine, and even identifies the human body as a machine:

I suppose the body to be nothing but a statue or machine [...] We see clocks, artificial fountains, mills, and other such machines which, although only man-made, have the power to move of their own accord in many different ways. But I am supposing this machine [the body] to be made by the hands of God, and so I think you may reasonably think it capable of a greater variety of movements than I could possibly imagine in it, and of exhibiting more artistry than I could possibly ascribe to it (AT XI 120 / CSM 1.99).

[...] those who know how many kinds of automatons, or moving machines, the skill of man can construct with the use of very few parts [...] will regard this body as a machine which, having been made by the hands of God, is incomparably better ordered than any machine that can be devised by man [...] (AT VI 55–56 / CSM 1.139).

And let us recognize that the difference between the body of a living man and that of a dead man is just like the difference between, on the one hand, a watch or other automaton (that is, a self-moving machine) when it is wound up and contains in itself the corporeal principle of the movements for which it is designed, together with everything else required for its operation; and, on the other hand, the same watch or machine when it is broken and the principle of its movement ceases to be active (AT XI 330–331 / CSM 1.329–330).

---

3 “The preservation of health has always been the principal end of my studies [...]” *Letter To [The Marquess of Newcastle]*, AT IV 329 / CSMK 275.

4 It has been claimed that Descartes simply extended mechanism to the mental in his “medical physiology”, see Carter 1983, 19.

The primary basis of this analogy or identification is the body's *self-moving* nature. It is not the soul that produces movement in the body, as Aristotelians had thought; that is, although certain movements of the body do depend on the soul, the latter is not the general principle of the movement of the body. That principle is internal to the body: "the heat in the heart is like the great spring or principle responsible for all the movements occurring in the machine" (AT XI 227 / CSM 1.316). Accordingly, the reason why a dead body does not move is not that the soul has departed from it, but that "the heat ceases and the organs which bring about bodily movement decay" (AT XI 330 / CSM 1.329).

However self-contained the operation of the body appears to be in Descartes' view, he also often emphasizes the *intrinsic* connectedness of mind and body. Apart from the sailor-and-the-ship disanalogy (which can be said to be about the way in which we *perceive* the unity of mind and body rather than the way they *are* actually united), Descartes talks about the *incompleteness* of mind and body when viewed from the perspective of their union:

It is also possible to call a substance incomplete in the sense that, although it has nothing incomplete about it *qua* substance, it is incomplete in so far as it is referred to some other substance in conjunction with which it forms something which is a unity in its own right.

Thus a hand is an incomplete substance when it is referred to the whole body of which it is a part; but it is a complete substance when it is considered on its own. And in just the same way the mind and the body are incomplete substances when they are referred to a human being which together they make up (AT VII 222 / CSM 2.157).

If we are to understand the appropriate approach to physiological processes as mechanistic, what could Descartes say about pathologies where the mind, apparently related to the body in an intrinsic manner, is also involved? The extension of mechanism to the mental is of course not a real possibility in Descartes. Although the body does have determinate effects on the mind, in the etiology and treatment of psychosomatic diseases and mental problems, the mind, with its own characteristics, is also involved; but there is a further problem as well, formulated by Lisa Shapiro in the following way:

On Descartes' official position, living bodies are simply machines whose workings can be explained by appeal to the laws of physics. However, these laws do not in and of themselves allow us to distinguish when a machine is working well or badly. [...] There seems to be nothing in the nature of either a clock or an animal-machine, taken on its own, which can tell us whether the machine is in good working order, or an animal is healthy. [...] And so, there seems to be no good way to ground ascriptions of health to a human body by appeal to the body alone (Shapiro 2003, 422–423).

What does it mean to be sick? It is, Descartes writes in the *Meditations*, to have a nature that is “disordered”. However, the sick body also operates according to the laws of nature, just like the healthy one, so “disordered” functioning cannot be “unnatural” functioning.

When we say, then, with respect to the body suffering from dropsy, that it has a disordered nature because it has a dry throat and yet does not need drink, the term ‘nature’ is here used merely as an extraneous label. However, with respect to the composite, that is, the mind united to this body, what is involved is not a mere label, but a true error of nature, namely that it is thirsty at a time when drink is going to cause it harm (AT VII 85 / CSM 2.59).

This is the sense of nature “which is really to be found in the things themselves; in this sense, therefore, the term contains something of the truth” (ibid). Descartes does seem to be saying in this passage that no real norm of the adequate functioning of the body can be found by taking the body separately from the mind, and that the good and the harm of the composite of mind and body is the standard by which adequate functioning is to be measured.

Thus, although we may intelligibly talk about “bodily” and “mental” health in a Cartesian context, health or illness belongs to the composite in the first place. By this I do not deny that for Descartes it is the body rather than the mind, that health (the survival of the composite) mostly *depends upon*. The reason for death is the decay of a fundamental part of the body, Descartes tells us in *The Passions of the Soul*, not the absence of the soul (AT XI 330 / CSM 1.329). But the break-down concerns the whole composite, even if is caused by bodily processes only.<sup>5</sup>

## II. Psychosomatic medicine in later Descartes

The most important evidence concerning Descartes’ notion of the mind’s relationship to the body from a medical point of view is found in his late correspondence with Elizabeth of Bohemia. Elizabeth, at the beginning of the discussion of her state of health with Descartes, suffered from “low-grade fever” and “dry cough” (*Letter to Elizabeth, 18 May 1645*, AT IV 201). Descartes immediately attributes these to psychological causes, namely her “sadness” due to the misfortunes that affected her family.<sup>6</sup> Things that distress us, says Descartes, are “domestic enemies with whom we are forced to keep company, and we have

---

5 Cf. Des Chene 2001, 150: “If the body is a whole only in the union, then its ceasing to be a whole is only with respect to the union. The body considered apart from the union cannot die, no more than the soul.”

6 Elizabeth of Bohemia (1618–1680) was the daughter of Frederick V of Bohemia and Elizabeth Stuart (daughter of James I of England). Her family lived in exile (in Germany, later in The Hague) after her father lost the Bohemian throne. Her uncle was Charles I of England, beheaded in the English Civil War.

to be perpetually on guard lest they injure us” (*Letter to Elizabeth, May or June 1645*, AT IV 218 / CSMK 249). What can we do against them?

[...] consider on the one hand a person who had every reason to be happy but saw continually enacted before him tragedies full of disastrous events, and who spent all his time in the consideration of sad and pitiful objects. Let us suppose that he knew they were imaginary fables, so that though they drew tears from his eyes and moved his imagination, they did not touch his intellect at all. I think that this by itself would be enough gradually to constrict his heart and make him sigh in such a way that the circulation of his blood would be delayed and slowed down. The grosser parts of his blood, sticking together, could easily block the spleen, by getting caught and stopping its pores; while the more rarified parts, being continually agitated, could affect his lungs and cause a cough which in time could be very dangerous. On the other hand, there might be a person who had countless genuine reasons for distress but who took such pains to direct his imagination that he never thought of them except when compelled by some practical necessity, and who spent the rest of his time in the consideration of objects which could furnish contentment and joy.[...] I do not doubt that this by itself would be capable of restoring him to health, even if his spleen and lungs were already in a poor condition because of the bad condition of the blood caused by sadness (*Letter to Elizabeth, May or June 1645*, AT IV 219 / CSMK 249–250).

It is persistently having the wrong kind of representation that harms its subject, whatever the source of that representation – the imagination or the intellect – might be. The imagination, which does not present the affliction as applying to the subject, can still trigger the same harmful bodily processes. Therefore, the representation itself should be removed. Even though he does not see himself as a Stoic, Descartes in effect prescribes Stoic remedies to the passions (focusing on one’s internal resources and strength in the face of difficulties) (Shapiro 1999, 509–510, 516). Descartes cites his own case in confirmation of the recommendation to Elizabeth:

I was born of a mother who died, a few days after my birth, from a disease of the lungs, caused by distress. From her I inherited a dry cough and a pale colour which stayed with me until I was more than twenty, so that all the doctors who saw me up to that time gave it as their verdict that I would die young. But I have always had an inclination to look at things from the most favourable angle and to make my principal happiness depend upon myself alone, and I believe that this inclination caused the indisposition, which was part of my nature, gradually to disappear completely (*Letter to Elizabeth, May or June 1645*, AT IV 220–221 / CSMK 250–251).

The mind, via thoughts, influences health. The most health-preserving thought is the conviction that we are strong, and cannot easily fall into illness, and if we do, we are capable

of restoring health (with the help of such thoughts, for instance) (*Letter to Elizabeth, 8 July 1644*, AT V 65 / CSMK 237). At the same time, the self-curing power of the mind cannot be unlimited, if for no other reason, because “there are diseases which take away the power of reasoning and with it the power of enjoying the satisfaction proper to a rational mind” (*Letter to Elizabeth, 1 September 1645* AT IV 282 / CSMK 262). The mind can even destroy the body, via beliefs that are the opposite of the health-restoring ones above: some people are “convinced by an astrologer or doctor that they must die at a certain time, and for this reason alone fall ill, and frequently even die. *I have seen that happen to several people*” (*Letter to Elizabeth, 8 July 1644*, AT V 66 / CSMK 237. e.m.).

Besides the above examples, there are innumerable references to the influence of mental states on (physical) health in Descartes’ correspondence with Elizabeth. Descartes assures Elizabeth that she can open his letters without expecting to read any news such that “they will impede the digestion of the waters you are taking” (*Letter to Elizabeth, 21 July 1645*, AT IV 252 / CSMK 256). He also claims that the most frequent cause of prolonged fever is sadness (*Letter to Elizabeth, 18 May 1645*, AT IV 201).

What about “mental health”? For Descartes, it would seem to mean, if anything, the absence of excessive and destructive passions. These, through their proximate causes, have a lot to do with bodily processes “And those with an unhealthy spleen are apt [...] to be sadder than others...” (AT XI 421 / CSM 1.372). But, besides the dubiousness of the “physical” and “mental” health dichotomy in Descartes, any clear-cut separation of mind-to-body and body-to-mind direction of impact seems arbitrary. In the curing of disturbances, bodily and mental causes do not appear separable:

I know indeed that it is almost impossible not to give in to the disturbances which new misfortunes initially arouse in us. [...] But the next day, *when sleep has calmed the turbulence that affects the blood in such cases*, I think one can begin to restore one’s mind to a state of tranquility. *This is done by striving to consider all the benefits that can derived from the thing which had been regarded as a great misfortune...* (*Letter to Elizabeth, June 1645*, AT IV 236–7 / CSMK 253. e. m.).

It is not accidental that Descartes treats (bodily) illnesses and emotional disturbances together. As we have seen, when trying to find criteria for health, that is, orderly operation, the body will not provide such for itself, and its union with the mind is essentially involved in it. Similarly, although this is less explicit in Descartes, the appropriate condition of the mind can only be measured by the presence or absence of certain inordinate passions, which are, in their causes, bodily phenomena. An emotional disturbance is not a purely mental one: just like physical illnesses, it involves changes in the temperature of the blood, abnormal motion of the animal spirits, and so on. Emotional and bodily problems often go together, with similar bodily mechanisms and processes leading from one kind of abnormal functioning to another. It can be the proximate causes or accompanying bodily movements of emotions themselves that make the person sick (constriction of the heart, slowing down the circulation of the blood etc., e.g.: AT XI 403 / CSM 1.363). This is why, when

advising Elizabeth, Descartes can use words like “indisposition” and “disturbance” often indiscriminately referring to mental and physical problems, switching to the discussion of how to regain emotional balance after introducing the topic of Elizabeth’s bodily illness:

[...] it is impossible for me to believe that a mind capable of such thought and reasoning should be lodged in a body which is weak and ill. [...] I know indeed that it is almost impossible not to give in to the disturbances which new misfortunes initially arouse in us. I know too that ordinarily the best minds are those in which the passions are most violent and act most strongly on their bodies (ibid).

### III. Is this all new?

Is Descartes’ position in the Elizabeth letter concerning the longer-term mutual effects of mind and body and his leaning towards psychosomatic rather than mechanistic medicine a fresh development? Descartes indeed discusses the reciprocal permanent influence of mind and body in his correspondence with Elizabeth more than he did before, but has his position changed or is it really more of the same?

It is well-recognized that in his letters to Elizabeth, Descartes offers advice concerning what both correspondents take to be psychosomatic diseases in line with a Stoic approach to emotion management. But Stoic influences had been manifest in Descartes’ philosophy from early on. In the *Rules*, he identifies the goal of wisdom as figuring out “how to increase the natural light of his reason... in order that his intellect should show his will what decision it ought to make in each of life’s contingencies” (AT X 361 / CSM I 10). The mental flourishing to be achieved he calls “contentment of mind” or “tranquility”. The mind’s illnesses are to be cured by philosophy, just as bodily illnesses are treated by medicine. It is one of Descartes’ very early recorded notes that “I use the term ‘vice’ to refer to the diseases of the mind, which are not so easy to recognize as diseases of the body. This is because we have frequently experienced sound bodily health, but have never known true health of the mind” (AT X 215 / CSM I 3).

That the management of mental life has significance for the regulation of bodily processes is also acknowledged rather early in Descartes’ career. He observes the following as early as 1620 or thereabouts: “I notice that if I am sad or in danger and preoccupied by some serious undertaking, I sleep deeply and eat voraciously. But if I am full of joy, I do not eat or sleep”.<sup>7</sup> Sleep and eating (nutrition) are purely bodily functions: Descartes has taken them out of the competence of the vegetative soul and placed them in the competence of the body. As such, they are shared with animals, and the processes they involve take place

---

7 *Observations*, AT X, 215 / CSM 1.3. The theme comes back in Descartes’ *Letter to Elizabeth, May 1646*, AT IV, 409 / CSMK 286.

according to physical laws. Thus mental states, especially emotions, may interfere with bodily mechanisms, just as in the Elizabeth letters.

In order to restore mental health, besides emotion management, Descartes also endorses taking physical measures (such as using spa water). However, Descartes seems to be going only with the (theoretical and practical) flow here in the contemporary understanding of melancholy, and there is no reason to assume that he had a robust, considered position of his own. Both Robert Burton in his *Anatomy of Melancholy* (1621) and Thomas Willis in *De Anima Brutorum* (1672) took melancholy to be a disease of both mind and body, to be treated by both mental effort on the part of the subject, and physical cures. That grief and sorrow specifically can have disastrous effects on health “was a well-established literary trope” (Schmidt 2007, 30). Not much novelty so far.

We should also consider the other direction: the influence of the body on the mind’s properties and capacities. Shapiro believes that Descartes became convinced by Elisabeth that embodiment influences the way in which reason operates (Shapiro 1999, 515). The belief that the operation of the faculty of reason is significantly influenced by organic factors does not seem to be recently acquired, however. He writes in the *Discourse*: “For even the mind depends so much on the temperament and disposition of the bodily organs that if it is possible to find some means of making men in general wiser and more skillful than they have been up till now, I believe we must look for it in medicine” (AT VI 62 / CSM 1.143). That the constitution of the brain determines the quality of thinking is witnessed by a letter from 1640 as well: “In the case of very good and subtle minds, I think the gland must be free from outside influence and easy to move, just as we observe that the gland is smaller in man than it is in animals, unlike the other parts of the brain” (*To Meyssonnier, 29 January 1640*, AT III 20 / CSMK 143).

It emerges from the *Passions* that, when explaining the disposition to have certain feelings, we may have to have recourse to bodily dispositions:

And those with an unhealthy spleen are apt not only to be sadder than others, but also at times to be more cheerful and more disposed to laughter, inasmuch as the spleen sends two kinds of blood to the heart, one very thick and coarse, which causes sadness, and the other fluid and thin, which causes joy (AT XI 421 / CSM 1.372).

In Galenic medicine, which was dominant up to Descartes’ time, the physiology of which having had a big impact on Descartes, melancholic temperament is responsible for what we would call the depressed phase in bipolar disorder (Galenus 1521/1881; cf. Jackson 1969). It is associated with black bile and the spleen as its producing organ. Sanguine, in turn, the temperament associated with blood, is responsible for what we would call the manic phase of bipolar disorder. So again there seems to be a casual, cursory association with the tradition, rather than Descartes’ working out of a new position. This also applies to the hint that bodily processes may not only modulate but also *create* purely mental inclinations or volitions: the strong agitation of the animal spirits excite the imagination so much that it “*makes [the person] inclined to compose poetry*” (*Letter to Elizabeth, 22 February 1649*, AT

V 281 / CSMK 367. Emphasis mine). The idea that physiological factors may determine mental activity such as writing poetry was again present in contemporary humoralism: overwork of the imagination was associated with choleric temperament.

All in all, it appears that in his correspondence on matters of health with Elizabeth, Descartes acts much more as an understanding and soothing partner than a theoretical reformer. The later Descartes' conception of how mind and body influence each other in terms of health and disturbance was not significantly different from his previous views, while at the rare points, where we might get the impression that Descartes has something in a different vein to add, he largely goes with received wisdom.

The second half of the seventeenth century saw the flourishing of the mind-body issue in the wake of Descartes' philosophy. Philosophical and medical treatises were produced in the Cartesian spirit for the clarification of the relationship between mind and body. In the last section, I will look at three authors who were interested in the medical implications of this relation, conceived in a Cartesian way. While pointing out continuities with Descartes, I will also identify shifts of emphasis in the adoption and re-interpretation of his ideas.

## **IV. Medical thinking and the passions in Johannes Clauberg, Louis de La Forge and Tobias Andreae**

Two of the characters I am going to introduce may be known for their connection to Descartes. Clauberg was the person who produced the surviving copy, or even helped edit the text, of Descartes' conversation with Franciscus Burman. He also had an essential role in the propagation of Cartesianism in Germany. La Forge, a physician, took up the task of illustrating Descartes' *Treatise on Man*, which led to his elaborate interpretation and supplementation of Descartes' thoughts on mind and body.

Clauberg and La Forge consider psychophysical interaction the foundation of the mind-body union. For Descartes, the way in which a given person affectively reacts to stimuli had depended on associations formed in his past. The significance of individual history for the union was retained by these authors, especially by Clauberg. Interaction is individualized for both. Its intensity and frequency, as well as the content of the states involved, show individual variation, and, in Clauberg's view, even determine the closeness of the union.

The third author, Tobias Andreae, was a medical philosopher, a contemporary of Clauberg and La Forge, who made further steps in the development of Cartesian psychosomatic medicine. He shared many of the basic assumptions of the other two concerning the relationship of mind and body, with his interest geared toward concrete diseases and treatments, as well as an in-depth analysis of the sensation of pain.

In Descartes, the default operation of mind and body is not their joint functioning: passions are *special* modes. Clauberg, in turn, dispenses with the special-event-like character of psychophysical operations, which become the rule rather than the exception. So much so that human *life* depends on these functions: “[...] life consists in action and operation. The whole human lives as long as the soul and the body are in conjunction: dies when those two are separated” (Clauberg 1663, X/4). The joint functioning of mind and body is individualized and dynamic (Clauberg 1663, X; XXX). Each mind-body unit is a special one, according to the degree, kind, etc. of psychophysical interaction. It represents a higher level of perfection to exhibit more variety in joint functioning, but less frequency in operations related to the senses at any given time. Newborns only have natural psychophysical connections, are more given to the senses and desires (*Conjunctio*, XLIII/3), and are less capable of governing their bodies by their will than adults, further developing tamer connections by habit.

Not only do types of physical states “depend on” types of mental states and vice versa, but global characteristics of individual minds also depend on those of the bodies they are connected with and vice versa. Like in Descartes, mind and body do not simply act on each other on certain occasions, but permanently incline and influence each other: mental disposition follows the temperament of the body (Clauberg 1663, I/5), while mental dispositions have bodily manifestations, which is the basis of physiognomy (Clauberg 1663, I/7). A whole chapter is devoted to the way in which intellectual capacities and virtues are functions of the mind’s connection to the body. Fast comprehension is due to the agility of the animal spirits (Clauberg 1663, XLIV/4). The same agility of the spirits renders thoughts wandering and frivolous, especially in adolescence (Clauberg 1663, XLIV/5). Mnemonic capacities, just as differences in mindset, depend on the disposition of the brain (Clauberg 1663, XLIV/8, 11).

Clauberg, like Descartes, considered placing medicine on proper foundations a primary goal. Initially, he was more concerned with adequate knowledge of *physics* as the condition of significant advances in medicine. His more enduring ambition regarding the groundwork of medicine was to establish general philosophical, metaphysical foundations for it. While physiology is subject to the laws of nature, of *health*, we cannot give a scientific account (Trevisani 1992, 112–113), for the latter bears the mark of the mutability and diversity involved in the union of mind and body (Clauberg 1663, LVIII/13, 14). Moral life, as well as God’s grace, have an influence on the health of the body. Since medicine is not universal and scientific, and scientific truth is inflexible, there are more effective and less effective medical treatments. Like in Descartes, lack of lawlikeness in the sphere of the mental excludes the possibility of a “third science”, and, similarly to what can be gleaned from Descartes’ later letters, makes (psychosomatic) medicine somewhat intuitive and open-ended.

Part of what La Forge has to say about the mind-body union is almost identical with Clauberg’s ideas. Mind and body must be united by something that “goes out of” its subject, that is, action and passion: “a body and mind are united when some movements of the first depend on thoughts of the second and, reciprocally, some thoughts of the second de-

pend on movements of the first” (La Forge 1997, 121). La Forge’s “articles of the union” are a rendering of Descartes’ views concerning permanent connections between mental and physical states. According to articles 2 and 4, the particular motions of animal spirits are “joined naturally” to “one thought which will always accompany it”. “Naturally joined” is contrasted with “thoughts which custom or the will of the soul joined” and which “the soul can undo”. Also, in the case of natural connections, thoughts are “united with certain movements *in all human beings*”. According to article 3, “that this thought will correspond to the state of the body, that is, it will be happy or sad, confused or clear” depends on “the disposition in which it found the body the first time”. As an explanation of articles 2 and 3, movements of body being “naturally joined” to mental states, and this depending on “dispositions in which [the soul] found the body for the first time”, La Forge offers the drosy and missing limb cases. The proximate causes of the feeling of thirst and hurting of the missing limb keep producing the same effects as “the first time”, despite the changes that have taken place in the meantime.

La Forge affirms a relationship of mutual reference and reinforcement between the mental and the bodily states associated with the passions: the mind’s state is in accordance with the action tendencies of the body and perpetuate the proximate causes of the passion:

[...] according to the laws of the union about which we spoke above, the motions of the soul which are joined with the sensations of the passions are those which are most appropriate and suitable for getting the soul to consent to those things to *which the body is disposed*. Thus for example, when our body is disposed to love, the mind immediately comes to have thoughts *which could excite this passion in it*. On the other hand, when the soul is distraught with grief, the body assumes a disposition which serves to maintain it (La Forge 1992, 194).

Why this should be so seems to be a pointless question to La Forge: it is among the “articles of the union” that lay the foundations of mind-body metaphysics. It is fairly clear, however, that the harmony between mental and physical and the sustained cycles of interaction between mind and body are a high-priority, unquestioned metaphysical value. This is all the more remarkable because there is no *normative* element in the whole schema. In Descartes, we recall, the passions were meant to focus the mind on what “nature deems useful” for us. Here, whatever processes the one substance has are to be reinforced by the other, regardless of their import. By this, the conception of the mind-body union as a unit of cooperation and of self-perpetuating psycho-physical cycles seems to be taken to a new level.

The work of Tobias Andreae<sup>8</sup> belongs to the second decade of the new university at Duisburg, where he was appointed professor of medicine in 1662. There was intense

---

8 Tobias Andreae (1633–1685), professor of medicine in Duisburg, is not to be confused with his uncle, who, by the same name, was a professor of Greek and history in Groningen and a friend of Clauberg. The elder Tobias Andreae (1604–1676) corresponded with Descartes about the latter’s dispute with Martin Schook, rector in Groningen, in 1645. Clauberg became a Cartesian under the influence of the elder Andreae: Sprecht 1966, 108, n. 38; Bouillier 1868, 294.

interaction between the medical and the philosophical faculties; sometimes the same person was affiliated with both and worked on topics of disciplinary overlap.<sup>9</sup> In the medical faculty, dissertations were produced with such titles as *Dissertatio philosophico-medica de homine microcosmo*; *Brevis explicatio corporis humani prout anima vegetativa pollet* (Trevisani 1992, 131–132).

In his *Exercitationes de conjugio mentis et corporis*, Tobias Andreae addresses a topic that cropped up in his medical work on epilepsy. Human health, claims Andreae, is not merely a physiological-mechanistic issue – if it were, it would not differ from veterinary science. Not only because the mind can be ill just as much as the body, but because the treatment of a body which is part of a psychophysical union has to be, at least in part, different from that of one which is not.

The *Exercitationes* start out with the usual early Cartesian introduction to the mind-body issue. We experience that our minds are affected by our bodies in different ways, e.g. in pain, and that our bodies are in turn affected by our minds, e.g. in the case of trembling hands. The principle or reason for this conjunction of mind and body cannot be found either in the mind or in the body itself, for physical conjunction has to do with surfaces in contact and bodies moving each other, and mental conjunction means the agreement of wills. Nor is it a simple combination of the two kinds. The psycho-physical unity experiences the world in *qualitative* terms (odors, light, pain, etc.). In internal and external perception, we do not experience things as mere motions and figures, even though perceptions are the result of the physical motions of the brain (Andreae 1679, I, 16). The “three grades” of perception, which include physical motion and qualitative feel, are however only conceptually separable. The physical motion in the brain (first grade) gives the mind a qualitative sensation (second grade). The mind does not judge this sensation “indifferently”, but simultaneously with, or rather as part of, the perception it adjudicates on its goodness or badness, whether its subject should seek or shun it (third grade) (Andreae 1679, I, 17).

States of the body do not just cause, or are registered in, whatever states of the mind God establishes correspondence with. A single qualitative state has both physical and mental input. Pain does not merely consist in the mind registering the harm of the body; it is not a physical fact (the body being harmed) signaled to the mind by way of a sensation of a certain (unpleasant) quality. Rather, the feeling of pain is informed both by physical processes in the body and the interpretation of what is happening in the body by the mind. Mucius Scaevola felt less pain when he put his hand into the fire than a coward would have felt (Andreae 1679, II, 6). Toothaches differ in degree according to what we take their causes to be: if we think the pain is caused by a piece of food being stuck in our tooth, the ache is less acute than it is when we put it down to the decay of the tooth (Andreae 1679, I, 25). Qualitative states are irreducibly states *of the union*.

For Andreae, the union has strong medical implications, especially with regard to the understanding and treatment of emotional-mental problems. Epilepsy, for instance, has

---

9 E.g. Wilhelm Fuchs, who received a doctorate in both philosophy and in medicine, in his *medical* dissertation denied two Cartesian theses: that the essence of the soul is thinking and that its seat is the pineal gland (Trevisani 1992, 131–132).

a neuro-physiological and a psychological understanding; the two do not exclude each other.<sup>10</sup> Relatedly, thoughts can be as good or bad for a disease with physical manifestations as physical influences. Thinking that I am extremely ill can make the problem worse; anticipation or remembrance of epileptic fits can induce another outbreak.

Mental problems typically have to do with the mismanagement of passions. Passions are bodily *and* mental phenomena, but not in the sense of being complexes consisting of physical changes that trigger mental ones or the other way round. Passions do not *cause* physiological changes, but *are* physiological changes *and* mental states at the same time (Trevisani 1992, 118). Being sad, for instance, *is* the constriction of the heart, rather than standing in a causal relationship with it.

In Descartes, there was a strong mechanistic-automatic element to emotional behavior. When threatened, the body has the tendency to run away, which the feeling of fear reinforces. The running would take place without the mind as well, without the involvement of any emotion. For Andreae, in contrast, there is nothing automatic in running away or staying and taking the beating (Trevisani 1992, 147). Rather, in the decision as to which course to take, there is a balancing of physical integrity against moral integrity, a meshing of the two perspectives, as it were. The mind and the body are each other's *instrument* and *goal* (ibid).

As mental and physical processes are inseparable in the case of emotional-mental problems, mental treatment can work on the body and bodily treatment can work on the mind (Trevisani 1992, 151); for any treatment really works on the *connection* between mind and body, and the symptoms, mental and physical, are those of a *misconnection*. The doctor may try to restore the right connection by either physical or mental means. In the normal case, there is a transparent, harmonious relationship between mind and body. The ways in which the harmony of this relationship gets disrupted are not predictable (Trevisani 1992, 152). Because of this, and because the right treatment depends on innumerable circumstances, medicine is not a science but an art (Trevisani 1992, 117).

## V. Conclusion

I have argued that Descartes did not substantially revise his position concerning the relationship between mind and body from a medical point of view under Elizabeth's influence. I have also tried to show ways in which Cartesian psychosomatic thinking survived and was pushed in an interactionist, individualized direction in early Cartesianism. The work of Tobias Andreae may claim added interest due to its remarkable closeness in some core features to our contemporary understanding of mental illness, such as attributing a significant role to the patient's self-understanding, and the validity of both physical and psychological explanations, also manifested in treatment.

---

<sup>10</sup> On Andreae's *Disputatio Inauguralis* on epilepsy, see Trevisani 1992, 117ff.

## Literature

- Andreae, Tobias. 1679. *Exercitationum Prima. De conjugio mentis et corporis*. Frankfurt an der Oder: Eichhorn.
- Bouillier, Francisque. (ed.) 1868. *Histoire de la Philosophie Cartésienne*. 2 vols. 3rd ed. Paris: Delagrave.
- Carter, Richard B. 1983. *Descartes' Medical Philosophy: The Organic Solution to the Mind-Body Problem*. Baltimore, London: The Johns Hopkins University Press.
- Clauberg, Johannes. 1663. *Corporis et animae in homine conjunction*. In: Clauberg, Johannes (1691/1968): *Opera omnia philosophica*. Ed. J. Th. Schalbruch, reprint. Hildesheim: Olms.
- Des Chene, Dennis. 2001. *Spirits and Clocks. Machine and Organism in Descartes*. Ithaca, London: Cornell University Press.
- Descartes, René. 1985. *The Philosophical Writings of Descartes*. Vols. 1–2. Tr. by John Cottingham, Robert Stoothoff and Dugald Murdoch. Cambridge: Cambridge University Press. (CSM)
- Descartes, René. 1991. *The Philosophical Writings of Descartes*. Vol. 3. Tr. by John Cottingham, Robert Stoothoff, Dugald Murdoch, and Anthony Kenny. Cambridge: Cambridge University Press. (CSMK)
- Descartes, René. 1996. *Oeuvres de Descartes*. Publiées par Charles Adam et Paul Tannery. Paris: Vrin. (AT)
- Galen. 1521/1881. *Galenus Pergamensis De temperamentis, et de inaequali intemperie libri tres*. (facsimile with translation). Tr. by Thomas Linacre. Cambridge: Sipherch.
- Gaukroger, Stephen. 1995. *Descartes. An Intellectual Biography*. Oxford: Clarendon Press.
- Jackson, Stanley W. 1969. Galen – On Mental Disorders. *Journal of the History of the Behavioral Sciences* 5 (4): 365–384.
- La Forge, Louis. 1997. *Treatise on the Human Mind*. Tr. by Desmond Clarke, Dordrecht: Kluwer.
- Shapiro, Lisa. 1999. Princess Elizabeth and Descartes. *British Journal for the History of Philosophy* 7 (3): 503–520.
- Schmidt, Jeremy. 2007. *Melancholy and the Care of the Soul: Religion, Moral Philosophy and Madness in Early Modern England*. The History of Medicine in Context. New York: Routledge
- Sprecht, Rainer. 1966. *Commercium Mentis et Corporis. Über Kausalvorstellungen in Cartesianismus*. Stuttgart-Bad Cannstatt: Friedrich Frommann Verlag.
- Trevisani, Francesco. 1992. *Descartes in Germania. La ricezione del cartesianesimo nella Facolta filosofica e medica di Duisburg (1652–1703)*. Milan: Francoangeli.

# VI

*Davide Monaco*

## A New Account of the Objective-Formal Distinction in Spinoza's Parallelism Theory

### Introduction

References to the objective-formal distinction play a fundamental role in the parallelism theory in the *Ethics*. For example, they are explicit in EIIP7C.<sup>1</sup> If this is true, then the correct understanding of the terms 'objective' and 'formal' must be a preliminary step for any valid interpretation of parallelism. Without shedding light on them and acknowledging their importance, Spinoza's theory would remain obscure in the worst case and would look none too original in the best.

At first, it must be noted that Spinoza deploys these terms in the *Ethics* with parsimony,<sup>2</sup> without defining them and taking for granted that his contemporary readers will understand their meaning. This implicit and unproblematic use might be explained by the fact that he wanted them to refer simply to a shared conceptual framework: if this is true, this shared framework would be perhaps Neo-scholastic and Cartesian philosophy (see van Bunge et al. 2011, 208–209), with which Spinoza partly agrees in his early works.

- 
- 1 I refer to the *Ethics* (E) in the following way: the first Roman numeral identifies the part of the *Ethics*, while the following letter indicates a definition (D), an axiom (A) or a proposition (P), which is followed by an Arabic numeral. Demonstrations (D), or scholia (S) and corollaries (C) when present, are indicated after this numeral. Scholia and corollaries might be also followed by a numeral: for example, EIP33S2 is the second scholium of proposition 33 of Part I of the *Ethics*. This notation is also followed when referring to the *Principles of Philosophy* (PP), while for other works I refer to parts (indicated with Roman numerals), chapters or paragraphs (Arabic numerals): *Treatise on the Emendation of the Intellect* (TdIE), *Short Treatise* (KV), *Metaphysical Thoughts* (CM). All translations are from Curley's edition.
  - 2 *Esse formale* is the most frequent term with eight occurrences (EIIP5, EIIP5D, EIIP6C, EIIP7S, EIIP15, EIIP15D), while *formaliter* has just one occurrence in EIIP7C. *Formalis essentia* has three occurrences (EIP17S, EIIP8, EIIP40S2). *Objective* has three occurrences (EIP17S, EIP30D, EIIP7C), while *esse obiectivum* has only two occurrences (EIIP8C, EIIP48S), one of which is from the *Nagelate Schriften* (NS).

However, there are many reasons to doubt that Spinoza was passively conforming to a traditional philosophical vocabulary. Accordingly, I will claim that he takes a rather original approach to this philosophical distinction, and that this renders Spinoza's position unique in respect to his predecessors.

In the first section of this paper, I will look at the main source of Spinoza's use of these notions: Descartes' *Meditations*, where this distinction plays a crucial role. Since Suárez's account influenced Descartes' theory of ideas, this will also be presented and then deployed in the next section as a yardstick to evaluate Spinoza's originality regarding the objective-formal distinction. In this section, I will show that Spinoza rejects the notion of the objective reality of ideas, while retaining only that of their formal reality; and I will explain why he does so. At last in the third section, I will demonstrate that Spinoza's parallelism demands a new variant of the objective-formal distinction, namely: one *in which the notion of the objective being of ideas is replaced by that of the formal being of ideas*. As a consequence, his distinction is more properly established between the latter and the formal being of things.

## I. From *Timaeus* to *Meditation III*

Reconstructing the history of the objective-formal distinction in detail is beyond the scope of this paper. However, I will briefly sketch some salient points, in order to highlight its importance for Spinoza's philosophy. This distinction originated in the Thomistic and Scotistic tradition, and by Suárez's time it had become widely popular. In his *Metaphysical Disputations* (DM II.I.1),<sup>3</sup> Suárez explains that the formal concept (*conceptus formalis*) is an act and product of our mind that *represents* the known thing, while the objective concept (*conceptus objectivus*) is the *represented* thing, to which the formal concept refers. For example, a human as an object of our thought is the objective concept, while our act of conceiving toward this object is the formal concept. It should be noted that this terminology can be also applied to the divine intellect, especially in connection with the issue of the 'demiurge scheme', discussed by several Scholastic philosophers. For Plato, the demiurge's acts of creation in the *Timaeus* are carried out according to exemplary ideas, which function as archetypes of created things. This dialogue was widely circulated in Calcidius' partial translation throughout the Middle Ages (see Klibansky 1981, 28), thus the myth was inherited and adjusted to Christian metaphysics, and the demiurge's ideas were moved to God's mind. However, this raised several theological issues, one of which was the threat to the theory of *ex nihilo* creation. In order to preserve this doctrine, things would have to possess no real being in God's mind before being created, but only a diminished type of reality, which Suárez qualifies as objective being (*esse objectivum*). The earliest technical notion of 'idea' appeared in this context, i.e. it was originally deployed to indicate God's ideas

3 When referring to the *Metaphysical Disputations* (DM), I indicate the disputation and the section with a Roman numeral and the number with an Arabic numeral, while I cite from Doyle's translation for DM LIV and from Wells' translation for DM XXXI.

(see Nuchelmans 1983, 23–24; Ayers 1998, 1062), and Descartes himself is well aware of this when replying to Hobbes' objections in his *Meditations* (AT VII 181; CSM II 127).<sup>4</sup>

Given that the objective-formal distinction was still popular in the first half of the seventeenth century (see Nuchelmans 1983, 30–31), it is not surprising to find some variants of it, both in Descartes<sup>5</sup> and Spinoza. In other words, such technical notions were part of a common philosophical vocabulary and early modern authors were clearly familiar with them. Indeed, even those who intended to break with Scholastic tradition did not discard them entirely, but continued to deploy them in specific contexts and for their own purposes (see Nuchelmans 1983, 35). I think that this is clearly displayed in Spinoza's parallelism, which can be said to be grounded at once in the reaction to and in the reevaluation of the 'demiurge scheme'. As I will show in the second section, one can catch a glimpse of this Scholastic debate in EIP17S and EIIP8, where Spinoza discusses the notion of God's intellect or idea<sup>6</sup>, which, as is clear from the beginning of Part II of the *Ethics*, plays a crucial role in the parallelism theory.

It is now time to consider the importance of the objective-formal distinction in Descartes' *Meditation III*, which Spinoza certainly read. In this meditation, the concept of the objective reality (*realitas objectiva*)<sup>7</sup> of ideas appears as a cornerstone of the argument demonstrating God's existence (Cronin 1966, 1). It should be noted that this is the first work in which Descartes uses this concept. The fact that this terminology was deemed very important by him can be best acknowledged by the note added in the Latin translation of the *Discourse on the Method* in 1644, where the notion of idea is explicitly connected with that of *esse objectivum* (AT VI 559). In the *Preface to the Reader*, Descartes presents the distinction, which will be implicitly held in *Meditation III*, between an idea as materially (*materialiter*) taken, i.e. "as an operation of the intellect", and an idea as objectively (*objective*) taken, i.e. "as the thing represented by that operation" (AT VII 8; CSM II 7). In the first case the idea bears no reference to something else, but must be considered solely in itself (and thus it is always true, whether it is the idea of a goat or that of a chimera (AT VII 37; CSM II 26), while objectively taken, the idea indeed does have a reference to something else. The first type of consideration is useless in order to prove God's existence (or the reality of external things), while it is only the second type which can be deployed in this regard. This is because the objective reality of ideas accounts for differences in representation e.g. ideas representing modes have less objective reality than ideas representing substances, and the idea representing God therefore has the highest objective reality (AT VII 165–166; CSM II 117). Therefore, beside their objective reality, ideas also have a formal reality, i.e. they are representational acts of our mind in the Suárezian sense.

---

4 I refer to the Adam and Tannery edition (AT) while I cite from Cottingham, Stoothoff and Murdoch edition (CSM): in both cases the Roman numeral indicates the volume, while the Arabic numeral indicates the page.

5 According to Gilson it is Suárez's definition of *conceptus objectivus* that Descartes deploys in his *Meditation III* (Gilson 1979, 49).

6 I use the notions of God's idea, God's intellect and God's mind interchangeably throughout my paper.

7 Descartes conflates this notion with that of *esse objective* (AT VII 102; CSM II 74). See also "l'estre objectif de l'idée" in a letter to Mersenne (AT III 545; CSM III 211).

While developing his argument, Descartes introduces his principle of efficient causality, according to which “there must be at least as much reality in the efficient and total cause as in the effect of that cause” (AT VII 40; CSM II 28). This principle does not apply only to physical objects but also to objects of our thought. Therefore, both the formal and the objective reality of ideas must have a cause: while the formal reality of ideas can be causally explained simply by referring it to the thinking subject (the *cogito*), i.e.: to ourselves; those ideas having an objective reality must have accordingly an extra-mental cause, i.e. a cause containing formally what the idea contains objectively. What must be underlined in the Cartesian argument is the presence of this causal relationship between the “objective mode of being” (*modus essendi obiectivus*) of ideas and the “formal mode of being” (*modus essendi formalis*) of an extra-mental reality (AT VII 42; CSM II 29). Obviously, not only ideas but also external objects have a formal reality, e.g. the cause of a stone must possess at least as much reality (formally) as the actual stone (conceived as an effect); likewise the cause of the idea of a stone must also contain at least as much reality (formally) as the one contained in the objective reality of this idea (also conceived as an effect) (AT VII 41; CSM II 28). Descartes readily points out that the cause of an idea (such as the cause of the idea of the stone) does not transfer its formal reality to the idea, because the idea derives it from the thought, of which it is a mode. The formal reality of the cause here does not bring about the formal reality of the idea, while it is the objective reality of this very same idea which must be derived from a cause possessing a formal reality. Since the thinking subject cannot certainly be the cause of the idea of God, because the objective reality of this idea cannot have been produced by it, this means that something else must be its cause, in which this reality is contained formally or eminently (i.e. in a more perfect way), and this is God itself (AT VII 42–45; CSM II 29–31). In *Meditation VI*, this argument will be completed. There Descartes will say that God does not transmit ideas of sensible objects to the thinking subject eminently, but that these ideas are derived from the objects themselves (AT VII 79–80; CSM II 55). In the fifth axiom contained in his geometric arguments in the second set of *Replies* (AT VII 165; CSM II 117), Descartes clearly says that the idea of the sky is caused by the sky itself, because “every idea must have a really existing cause of its objective reality”. The actual sky here is also the cause of the idea of the sky, because the former contains formally what is given objectively in the latter. Given this account, it is clear that Descartes is more interested in the concept of the objective reality of ideas, than in that of their formal reality, which he dismisses quite hastily. I claim that, for Spinoza, the opposite is true.

## II. Idea sive essentia obiectiva

Spinoza summarises Descartes’ notion of the objective reality of ideas in the *Principles of Philosophy* (PP ID3), however this definition does not shed any light on his own views, because this work is only meant to provide a reliable exposition of Cartesian philosophy. Therefore, it is necessary to look at other works. First, I will analyze the objective-formal

distinction in some other early texts, such as: the *Treatise on the Emendation of the Intellect* and the *Short Treatise*, where the Scholastic-Cartesian terminology is mostly applied to finite human intellects. Afterwards, I will move to the *Ethics*, where this distinction is established from the point of view of God's intellect.

Even if Spinoza is rather close to Descartes' views in his early works, his terminology is already quite original. In these works, Spinoza uses the concept of objective essence (*essentia objectiva/voorwerpelyk wezen*) simply as a synonym for idea. In the first work, the objective essence of something is simply the true idea of the thing, i.e. the certainty of the truth. For example: "a true idea of Peter is an objective essence of Peter" (TdIE 34). To have certainty means to have the objective essence or the adequate idea of something (TdIE 35). It is interesting to note that while ideas still have a formal essence here, they do not have an objective essence, because they are objective *essences* themselves. *An objective essence can be predicated not of an idea, but a thing*. This original use finds further confirmation when Spinoza talks about the objective essence of one being causing all of our ideas. This objective essence here is clearly a synonym for God's idea (TdIE 99). Similarly, in the *Short Treatise*, when Spinoza talks about the human mind, he renames it as 'idea' or 'objective essence' (KV II, appendix II). It can be therefore assumed that objective essence is only a synonym for the word idea (see Yakira 2015, 143). In other words, it can be said: idea is objective essence.<sup>8</sup> This vocabulary will be slightly modified in the *Ethics*: here Spinoza no longer deploys the notion of objective essence but only that of objective being, the latter only referring to things. Here also, Spinoza equates the concept of *objective being* with that of idea, thus rejecting the concept of the objective being *of ideas* altogether. Thus, in EI-IP8C we are asked by Spinoza to conceive of ideas as the objective being of singular things. Moreover, in EIIP48S we read:

[...] By ideas I understand, not the images that are formed at the back of the eye (and, if you like, in the middle of the brain), but concepts of Thought [NS: or the objective Being of a thing insofar as it consists only in Thought].<sup>9</sup>

However, one immediately notices that in the *Ethics* the references to the objective side of the distinction become rarer, and I believe there are at least two reasons why this is the case: one is that the concept of the objective reality of ideas is connected in Descartes with

---

8 One problem with this interpretation is that if all ideas are objective essences, and the objective essences are necessarily true, then there would be no more inadequate or false ideas. All ideas would be true. I do not think this is a serious problem if we consider Spinoza's notions of inadequacy and falsity in the *Ethics*. Inadequate ideas are not false ideas, because "all ideas, insofar as they are related to God, are true" (EIIP32) and "there is nothing positive in ideas on account of which they are called false" (EIIP33), but they are partial or 'mutilated' ideas (EIIP35), i.e. they entail privation of knowledge. In other words, inadequacy exists only for finite minds. Given that the objective-formal distinction relates with God's mind (both in EIIP17S and in the parallelism propositions of Part II), there is no doubt that Spinoza is considering only true ideas there.

9 Akkerman thinks that this unnecessary (albeit intelligent) addition from the NS should be removed from the text (Akkerman 1980, 149).

an extra-mental causal principle, i.e.: it is clearly caused by something outside the nature of thought (see Lord forthcoming). In spite of Spinoza's equation between idea and objective essence, this is also true for his early works, in which he held different causal views. Indeed, in the *Short Treatise* he still thought that ideas were causally connected with their objects (KV I 1; KV II 22), and he heavily relied on the Cartesian doctrine of 'animal spirits' to explain the mind-body interaction (KV II 19). Therefore, his objective-formal distinction was *de facto* heavily influenced by the Cartesian model of causality. However, Spinoza's ban of an inter-attribute causality in the *Ethics* makes it impossible to accept what was clearly possible in the Cartesian account of the *Meditations* and in his own early works. Spinoza might have thought that, if he did not want to confuse his reader, he should have refrained from referring to the objective essence to a large extent in his late philosophy. This is because any Cartesian reader would have immediately interpreted this concept in the 'traditional' way.

Another reason could be that some ideas having objective reality might not correspond to any actual thing in some of Descartes' Scholastic sources, for example in Suárez where the objective concept can refer to beings of reasons and universals (DM II.I.1.). By this Suárez means that nothing would correspond to them in reality, e.g. beings of reason would have just *esse objective* in our mind.<sup>10</sup> Spinoza does not grant universals and beings of reason any ontological status, i.e. the former are said to be nothing in the *Short Treatise* (KV I 6),<sup>11</sup> while the latter are also nothing and cannot be classified as ideas in the *Metaphysical Thoughts* (CM I 1). This is because every idea must be matched with an *ideatum*, and beings of reasons and universals clearly have none. It is true that according to EIIP38C, there are some notions (or ideas) that are common to all men and therefore can be said to be universal (EIIP40S2). However, what is common to all things can be rightly considered to be the object of an idea which is necessarily adequate (EIIP38D). That we are talking about a real idea possessing a real *ideatum* is confirmed by Spinoza's reference to EIIP7C in the demonstration of EIIP38. In that corollary Spinoza refers to God's power of thinking and clearly God's knowledge is of particulars and not of universals (CM II 7). For Suárez there is no intrinsic universality in the things (i.e. there are no 'universals *in re*') (see Ross 1962), and also for him, beings of reason are not real beings (in the sense of ultrarealism); however in DM LIV.I.6. he makes clear that they can have objective being in the mind:

[...] That which is [...] objectively in the mind sometimes has in itself, or can have, true real being, in line with which it is an object for reason. Absolutely and without qualification, this is not a genuine being of reason but rather a real being, for this [true and real] being is what simply and essentially belongs to it; whereas,

10 See also Goclenius' *Lexicon*, where the being of reason "*habet esse obiectivum (non reale) in intellectu*" (Goclenius 1613, 270). Also *objective*, when used in this context, has this particular meaning: to exist objectively would mean to exist only in the intellect and not in reality.

11 In one case, Spinoza refers to those collections of images termed 'transcendental' as confused ideas in EIIP40S1. However, he is perhaps not using the word 'idea' in its technical meaning there (this also happens with the non-technical use of the word 'attribute' in the same passage).

to be an object for reason is extrinsic and accidental to it. But at times something is an object for reason, or considered by reason, which does not have in itself any other real and positive being besides being an object for the intellect or for the reason thinking of it. And this is most properly called a being of reason. For it is in reason somehow, that is, objectively, and it does not have another more noble or more real way of being from which it could be called being. Therefore, what is normally and rightly defined as a being of reason is *that which has being only objectively in the intellect* or is *that which is thought by reason as being, even though it has no entity in itself* (Suárez 1995, 62).

Moreover, objective being pertains to beings of reason also when they are contained in the divine intellect (see Novotný 2013, 78–79).<sup>12</sup> Suárez’s second sense of the objective concept could not have been accommodated in the system of the *Ethics* and it is reasonable to suppose that Spinoza was rather suspicious of this terminology on the whole or not entirely ready to borrow it wholeheartedly. As it is clear from EIIP7C, Spinoza would only opt for the first sense of ‘being objectively in the mind’, and his use of *objective* there confirms this. Whatever is in God’s mind, while existing objectively, is *always* matched with a real being, i.e. *objective* is always paralleled by *formaliter*.

Spinoza’s theoretical contribution can be further appreciated if we consider his demonstrations for God’s existence in the *Ethics*. Descartes’ use of this terminology is subject to his initial assumption that the external reality must be proved. As a consequence, the meaning of the categories considered so far is highly dependent on this practical need. After all, the notion of the objective reality of ideas must serve to prove God’s existence. The Cartesian argument of the *a posteriori* proof for God’s existence and the corresponding use of the objective-formal distinction are well received also by early Spinoza, e.g. in the *Short Treatise*, where he deploys exactly the same strategy. However, there he also states that the *a priori* demonstration is to be preferred over the *a posteriori* one (KV I 1). He indeed abandons this proof in the *Ethics* while replacing it with another type of *a posteriori* proof (EIP11aliter2). For Spinoza there is no thinking subject to start from and there is no outer reality that must be called into question and then proved real. Indeed, God’s existence can be better proved *a priori*: in the concept of God its essence necessarily involves existence. At first sight, EIP30 with its demonstration, where the word *objective* comes up, could seem to be recalling the Cartesian *a posteriori* argument, but this proposition is not designed to prove God’s existence in Spinoza’s intentions. Here Spinoza demonstrates only that an actual intellect, whether finite (e.g. human) or infinite (e.g. divine), must comprehend everything that is in reality. Here, Spinoza is considering both a finite and an infinite intellect,

12 In CM II 7 Spinoza says that God knows beings of reason only insofar as he “preserves and produces the human mind”, i.e. insofar as he constitutes the essence of the human mind. But he does not know them “outside human minds”. This point is strikingly similar to Suárez’s position, who maintains that God does not need beings of reason in order to conceive things, and that he knows them only insofar as they are formed by our minds. However, Suárez also adds that they receive *some* being in virtue of divine cognition (see Novotný 2013, 78–79).

and he is therefore interested in their common properties. It is interesting to find some of the building blocks of the Cartesian *a posteriori* proof deployed now by Spinoza as a plain description of the content of all intellects.

Therefore, Spinoza in the *Ethics* does not deploy the concepts of objective and formal being in connection with the demonstration of God's existence (see Yakira 2015, 139). On the contrary, it must be observed that he establishes his objective-formal distinction from the perspective of God's intellect, and does not deploy the Cartesian twofold consideration of ideas anymore. The objective-formal distinction appears for the first time in the *Ethics* in EIP17S, in which Spinoza uses a *reductio ad absurdum* argument (Lærke forthcoming):

If intellect pertains to the divine nature, it will not be able to be (like our intellect) by nature either posterior to (as most would have it), or simultaneous with, the things understood, since God is prior in causality to all things (by [EI]P16C1). On the contrary, the truth and formal essence of things is what it is because it exists objectively in that way in God's intellect.

Spinoza intends to show that such an intellect would be anterior to the things understood, i.e. it would have to exist before these things exist. This would also mean that the truth and the formal essence of things would be what it is *because* it exists objectively in that way in God's intellect. This option would resemble very closely the 'demiurge scheme'. If the intellect pertained to God's essence, then it would follow that all things would have an *esse objectivum* in God's intellect before their 'creation'. However, Spinoza does not hold this opinion, thus the absurd option he displays in this scholium could be meant as a polemical reference to those Scholastic philosophers who held this opinion.<sup>13</sup> For Spinoza, God's intellect belongs to *Natura naturata* and not to *Natura naturans*, as Koyré (1950) rightly pointed out. Spinoza thinks that God's intellect is an infinite immediate mode of the substance under the attribute of thought (EIP21). Being created alongside the other immediate modes, it can also parallel them, while if it was equated with God's essence, the ontological symmetry of these modes would be violated, and God's cognition would have to precede God's production. Therefore, God is not supposed to have ideas of the things 'before' they follow from him. The same problematic and absurd outcome would be indeed present if we imagine that God's intellect is posterior (like the human intellect) to the things understood. In that case, he would know things only after they have followed from him and the sequence of ideas would follow that of things. Instead, Spinoza asks his reader to reflect on the word 'because' (*quia*),<sup>14</sup> connecting the two sequences, i.e. on the causal dependence that would exist between the formal essence of things and their objective counterpart in God's intellect, if that sentence had to be read literally. If we were to rephrase Spinoza's words and to reverse the logic of his argument, he would clearly be saying that "the truth and formal essence of things is what it is, *not* because it exists objectively

---

<sup>13</sup> In Spinoza's *Ethics*, scholia have often a polemical function.

<sup>14</sup> *Quia* has a clear causal meaning in this passage.

in that way in God's intellect". In other words, there cannot be a cause-effect relationship between ideas and things. If this view is correct, then God's intellect must be *simultaneous* with the things it knows. What does this mean exactly? In EIIP8 and EIIP8C, Spinoza suggests that ideas (the objective being of things) exist in God's mind and their *ideata* exist in God's attributes, even though neither has any durational dimension. Non-durational ideas are paralleled with non-durational things. Likewise, durational ideas are matched with durational things:

*The ideas of singular things, or of modes, that do not exist must be comprehended in God's infinite idea in the same way as the formal essences of the singular things, or modes, are contained in God's attributes.*

[...] So long as singular things do not exist, except insofar as they are comprehended in God's attributes, their objective being, or ideas, do not exist except insofar as God's infinite idea exists. And when singular things are said to exist, not only insofar as they are comprehended in God's attributes, but insofar also as they are said to have duration, their ideas also involve the existence through which they are said to have duration.

Since God's idea is eternal (EIP21) and God's essence is also eternal (EID6), whatever is included in them must be conceived *sub specie aeternitatis*, i.e. the essences of things are eternal (CM I 2). When ideas and things are instead conceived *sub specie durationis*, they are also matched with one another. Thus, it follows that things have an *eternal objective being*, insofar as they are contained in God's idea, and a *durational objective being*, insofar as they are said to exist durationally. At this point, one could ask what distinguishes Spinoza's position from that of the 'demiurge scheme'. Does not Spinoza also hold that things have an objective being in God's mind before their creation, like some Scholastic philosophers thought?

A comparison with Descartes on this point might not be very telling. Even though *Meditation V* is a privileged *locus* for the discussion of Descartes' Platonism concerning eternal truths and their relation to God (see Schmaltz 1991), he does not mention the objective-formal distinction there. It is only Caterus who brings this problem to Descartes' attention in his objections. The theologian implies that all ideas are like that of the triangle discussed in *Meditation V*: they could have an objective reality, which is not derived from any external cause (AT VII 93; CSM II 67). Descartes rebuts Caterus' objection by implicitly hinting at the fact that the objective reality of ideas of geometrical beings also have eternal essences that must be causally produced (AT VII 104–106; CSM II 76–77). It has been rightly claimed that Descartes' account of ideas in *Meditation III* must be viewed in connection with his doctrine of eternal truths and that the objective being of eternal truths is efficiently caused by God (see Wells 1990, 61); however, I think it is preferable to neglect Descartes' perspective here and rather compare Spinoza's position to Suárez's account. There, one can find an explicit reference to the objective being used in connection with God's cognition. This comparison also holds better given Spinoza's theocentric point of view in the *Ethics*.

The Spanish Jesuit suggests that things have a potential being before being created, i.e. as ‘creatable’. This being is said to be real “not by a true reality of its own which it has actually in itself, but because it can be made real” (DM XXXI.II.2). Moreover, he maps the potential being onto the objective being (*esse potentiale objectivum*).

However, it is certain in faith that God did not make created essences from eternity, neither from necessity [...]; since, it is a matter of faith that God does nothing out of absolute necessity; nor from free will. For it is likewise a matter of faith that He began to operate in time. Furthermore, it is evident that, if the essences of things had been made by God from eternity, they also would have been existing from that time on, because every effecting is terminated at existence [...]. This is confirmed, for otherwise God could not return something into nothing since something of the thing would always remain, namely, the essence. Also God would not have created all things from nothing but would have transferred them from one being to another being (DM XXXI.II.3; Suárez 1983, 59).

For Spinoza things certainly have an objective being in God’s mind before existing in duration, but this is not a potential being as in Suárez’s case, because for him there is no creation *ex nihilo* (CM II 10). He refers to *esse Potentiae* only with respect to God’s power of creation, and not to the being of things themselves before being actualised (CM I 2). On the other hand, these ideas are not archetypes that can be participated by things, but *singular ideas of singular things* as stated in EIIP8. For these reasons, the ‘demiurge scheme’ is rejected from the outset. In EIIP8 Spinoza clarifies that parallelism is valid both *sub specie aeternitatis* and *sub specie durationis*, i.e. ideas in God’s mind still correspond to real objects, even if they cannot be found in duration. This clearly shows that Spinoza’s notion of objective being is entirely different from Suárez’s. On the other hand, Spinoza was certainly aware that his readers would draw comparisons between his account and the Scholastic ‘demiurge scheme’ when confronted with this terminology. For this reason, he probably wanted this theory to peek out in his EIP17S and EIIP8C, so that the originality of his position might be immediately noticeable to his readers.

### III. Parallelism and the ‘formal-formal distinction’

For Spinoza, the objective-formal distinction is no more operational within the field of ideas like the Cartesian account, but it is only applicable to the relationship between reality and thought. While *formaliter* refers to God’s nature (consisting of infinite attributes), *objective* always refers to the way things are contained in God’s intellect. In fact, as showed in the last section, the context in which Spinoza’s objective-formal distinction appears for

the first time is unrelated to his *a posteriori* proof of God's existence. It certainly cannot be a coincidence that Spinoza refrains from deploying these concepts in his *a posteriori* proof, as he did in his early works. Meanwhile he introduces them when he clarifies the nature of God's intellect in EIIP17S. Therefore, it is also plausible that when these terms appear once again in the parallelism propositions in Part II, they stand on completely new ground. In other words, when Spinoza deploys these terms there, they are free from the function that Descartes applied to them.

It is time to examine separately both the concept of the formal being of things (*esse formale rerum* of EIIP6C) and that of the formal being of ideas (*esse formale idearum* of EIIP5), in order to show that Spinoza wished to establish his parallelism between the two. However, given that the terminology deployed by Spinoza varies greatly in the *Ethics* (e.g. Spinoza uses both 'formal essence' and 'formal being'), one must preliminarily enquire whether there is any difference between the concepts of essence and being in this regard. While in the *Metaphysical Thoughts*, for Spinoza there seems to be a distinction between *essentia* and *esse* (CM I 2); in the *Ethics* this issue is not addressed, and he seems to use the concepts of the formal being and the formal essence as equivalents. For example, in EIIP7S and EIIP15 Spinoza refers the formal being of ideas to finite ideas: in the first case Spinoza talks about the "formal being of the idea of the circle", while in the second he talks about the idea that constitutes the "formal being of the human Mind". Formal being can be also referred to things as in EIIP6C. On the other hand, formal essence is never referred to ideas (as it was in the *Emendation*) but always to things, and it is used both in the singular and the plural. The term objective essence is never used but is replaced by the term objective being, which becomes a *hapax legomenon*, if we accept Akkerman's advice. Therefore, I think it can be assumed that in the *Ethics* there is no special distinction between formal essence and formal being on one side, and objective essence and objective being on the other, and that for Spinoza they would be interchangeable.

In EIIP6C Spinoza says that the *formal being of things* does not follow from God's nature because the latter has first (*prius*) known the things. These words sound very familiar if we keep EIIP17S in mind: in fact, they look like the paraphrase of that earlier passage. Here Spinoza does not refer God's cognition to his intellect but to his nature, and he wants to show that absurd consequences follow from this. There cannot be any nature that knows things before they are causally derived from this very same nature. This is a clear sign that he expects his readers to go back to EIIP17S, where he denied that God's intellect pertains to God's essence or nature. Once again, Spinoza warns them against the ontological asymmetry that would exist between God's intellect and the sequence of things if God's intellect were anterior to them. Instead, the formal being of things does not depend on and does not follow from any objective being in God's mind. Therefore, Spinoza establishes simultaneity between the two patterns also in EIIP6C. In this corollary, he also provides an implicit definition for the formal being of things, when he tries to explain how these things actually follow from divine nature. The ideated things, i.e. the objects of the ideas, follow from their attributes in the same way as the ideas follow from the attribute of thought, i.e. things do not follow 'after' God has formed ideas of them, but at the same time. We can thus establish an equation between the formal being of things, which are not modes of thinking, and

ideated things. In other words, here the formal being of things includes all things, which are not modes of thinking. It is worth noting that Spinoza's formal being of things is not reducible to physical things like it was for Descartes, i.e. to the attribute of extension, because it also includes the things in the unknown attributes.

In EIIP5, Spinoza states that the formal being of ideas is caused by God as conceived under the attribute of thought, i.e. as *res cogitans*. In the demonstration, Spinoza clarifies this even further: here he implicitly equates the formal being of ideas with God's idea, which refers to his essence and to all things that follow from it, according to EIIP3. Spinoza deploys three different expressions to highlight the causal relationship between the formal being of ideas and God, conceived under the attribute of thought. (I) The formal being of ideas "admits God as a cause only insofar as he is considered as a thinking thing"; (II) the ideas of God's attributes and of singular things *do not* admit the objects themselves as their efficient cause, *but* God himself, insofar as he is a thinking thing; and (III) "God can form the idea of his essence, and of all the things that follow necessarily from it, solely from the fact that God is a thinking thing". This means that the formal being of ideas, of which Spinoza is talking about, is a product of God's causality. The literal meaning of the adjective 'formal' is an explicit reference to God's productive activity, conceived as *formare*: God forms his own idea, i.e. his own intellect.

It is reasonable to think that these two concepts (the formal being of ideas and the formal being of things) are those which will be paralleled in EIIP7, where Spinoza expresses his parallelism theory: "*the order and connection of ideas is the same as the order and connection of things*". In the following corollary, Spinoza comes back to the *formaliter-objective* distinction, but it must not be forgotten that there is no Cartesian objective reality of ideas anymore. Rather, we can only find here the objective being of things in God's intellect, which must be regarded only as a synonym for the formal being of ideas following from God *qua res cogitans*:

[...] Whatever follows formally from God's infinite nature follows objectively in God from his idea in the same order and with the same connection (EIIP7C).

If God's intellect is simultaneous with the things it knows, as EIP17S and EIIP6C both affirm, and if it is a *created* intellect rather than a *creative* one (an infinite immediate mode of thought which belongs to *Natura naturata*), then it cannot contain any pre-existing ideas that will be later matched with created things. The act of 'creation' happens simultaneously in God's intellect and in God's nature. For Descartes, God's idea has a heuristic role, clearly connected with the crucial importance of the objective reality that it contains: God's idea in us is a proof that God exists. This role is also evident in Spinoza's early works, while in the *Ethics*, God's idea possesses new ontological features. Rather than being evidence of God's existence, which can be best acknowledged simply by considering the definition of God, it is the product of God's efficient causality within the attribute of thought: for this reason, the concept of formal being is preferred in this context over the Cartesian concept of the objective reality of ideas.

Spinoza is indeed much more interested in equating the formal being of ideas with that of things. While for Descartes, the concept of the formal reality of ideas does not play any

decisive role in *Meditation III*, Spinoza revives it in the *Ethics*, and this interest is clearly visible in the parallelism propositions. There Spinoza is trying to walk down a path that Descartes had just outlined and abandoned. Descartes suggested that “the nature of an idea is such that of itself it requires no formal reality except what it derives from my thought, of which it is a mode” (AT VII 41; CSM II 28), but immediately after, he added that this was not enough and one had to move further to consider its objective reality. For Spinoza, instead, this Cartesian statement is sufficient, once it is adjusted to his new ontology. Ideas still receive their formal being from thought, of which they are modes, but now thought is one of God’s infinite attributes and constitutes his essence. This is also evident if we compare these two rather opposite remarks, the first by Descartes and the second by Spinoza:

And although one idea may *perhaps* originate from another, *there cannot be an infinite regress* here; eventually one must reach a primary idea, the cause of which will be like an archetype which contains formally “and in fact” all the reality “or perfection” which is present only objectively “or representatively” in the idea (AT VII 42, CSM II 29. Emphasis added).

And:

[...] The formal being of the idea of the circle can be perceived *only* through another mode of thinking, as its proximate cause, and that mode again through another, *and so on, to infinity* (EIIP7S. Emphasis added).

Indeed, as confirmed by EIIP9, for Spinoza there is no problem with such an infinite regression in the sequence of ideas. One must not posit a first idea, let alone an archetype. In the *Ethics* ideas can only originate from other ideas and their connection must be carried out to infinity. For Descartes God causes the “primary idea” *sic et simpliciter*, while for Spinoza the cause of God’s idea could be only God *qua res cogitans*. Only in this way, God can be the cause of his intellect and of other ideas. In Spinoza the formal sequence of ideas must parallel and correspond to the formal sequence of things, but God causes each sequence while being constituted by different attributes. If the parallelism demands a non-causal correspondence between the formal being of ideas and the formal being of things, then the causal interaction between ideas and things still present in Descartes is relinquished, because it yields to a ‘formal’ correspondence between the two sequences. The objective-formal distinction is thus replaced, as it were, by a formal-formal distinction in the parallelism. It appears now clear that Spinoza is using Descartes’ principle of efficient causality in each attribute separately.<sup>15</sup> The cause must contain as much formal reality as one can find in the effect, insofar as both the cause and its effect are situated in the same

---

15 The application of the principle of efficient causality within the sole attribute of thought is neither contradictory nor it yields to any materialist reading of Spinoza. This is because Spinoza’s principle of efficient causality is rather different from Aristotle’s one, which evidently includes in its definition the reference to change and rest and therefore to motion.

attribute. In other words, this causal principle cannot have any cross-attribute validity and the formal being in an attribute is the product of God's causality insofar as he is conceived under that particular attribute.

## Conclusion

Spinoza's attitude toward the objective-formal distinction is rather original and the latter is reducible neither to previous Scholastic models nor to the Cartesian formulations contained in the *Meditations*. Since this distinction plays a crucial and new role in the architecture of the *Ethics*, which is clearly visible in the parallelism theory of Part II, Spinoza must be credited with a good degree of originality regarding this doctrine. The dismissal of the Cartesian concept of the objective reality of ideas is the direct consequence of the ban of inter-attribute causality, fully developed by Spinoza only in this work. Here, the newly enhanced position of the concept of the formal being (of ideas and things) is the outcome of three different steps: first, in EIP17S Spinoza places the formal essence of things and their objective being side by side, denying that any causal relationship exists between them. Secondly, in EIIP5 and EIIP6C he clearly states that the formal being of ideas and that of things are the products of God's efficient causality insofar as the latter is exerted in the attribute of thought and in all other attributes respectively. Thirdly, given that God's causality is identical in all attributes, a complete parallelism or identity is established between the formal being of ideas and that of things.

## Bibliography

- Akkerman, Fokke. 1980. *Studies in the Posthumous Works of Spinoza: On Style, Earliest Translation and Reception, Earliest and Modern Edition of Some Texts*. Groningen: Rijksuniversiteit Groningen.
- Ayers, Michael. 1998. Ideas and Objective Being. In: Garber, Daniel – Ayers, Michael (eds.): *The Cambridge History of Seventeenth-Century Philosophy*. Vol. 2. Cambridge: Cambridge University Press. 1062–1107.
- Cronin, Timothy J. 1966. *Objective Being in Descartes and in Suarez*. Roma: Gregorian University Press.
- Descartes, René. 1897–1913. *Œuvres de Descartes*. 12 vols. Edited by Charles Adam and Paul Tannery. Paris: Léopold Cerf.
- Descartes, René. 1984. *The Philosophical Writings of Descartes*. Volume II. Translated by John Cottingham, Robert Stoothoff, Dugald Murdoch. Cambridge: Cambridge University Press.
- Gilson, Étienne. 1979. *Index Scolastico-Cartésien*. 2nd ed. Paris: Librairie Philosophique J. Vrin.
- Goclenius, Rodolphus. 1613. *Lexicon Philosophicum, Quo Tanquam Clave Philosophiae Fores Aperiantur*. Frankfurt am Main: M. Becker.

- Klibansky, Raymond. (1939) 1981. *The Continuity of the Platonic Tradition During the Middle Ages*. London: Warburg Institute. Reprint, München: Klaus.
- Koyré, Alexandre. 1950. Le Chien Constellation Céleste, et le Chien Animal Aboyant. *Revue de Métaphysique et de Morale* 55 (1): 50–59.
- Lærke, Mogens. (forthcoming). Aspects of Spinoza's Theory of Essence. Formal Essence, Non-Existence and Two Types of Actuality. In: Sinclair, Mark (ed.): *The Actual and the Possible: Modality and Metaphysics in Modern Philosophy*. Oxford: Oxford University Press.
- Lord, Beth. (forthcoming). Spinoza on Thinking Substance and the Non-Substantial Mind. In: Copenhaver, Rebecca – Shields, Christopher (eds.): *The History of the Philosophy of Mind*. Vol. 4. London: Routledge.
- Novotný, Daniel D. 2013. *Ens rationis from Suárez to Caramuel: A Study in Scholasticism of the Baroque Era*. New York: Fordham University Press.
- Nuchelmans, Gabriel. 1983. *Judgement and Proposition: From Descartes to Kant*. Amsterdam: North-Holland Publishing Company.
- Ross, James F. 1962. Suarez on Universals. *The Journal of Philosophy* 59 (23): 736–748.
- Schmaltz, Tad M. 1991. Platonism and Descartes' View of Immutable Essences. *Archiv für Geschichte der Philosophie* 73 (2): 129–170.
- Spinoza, Baruch. 1925. *Spinoza Opera*. 4 vols. Edited by Carl Gebhardt. Heidelberg: Carl Winter.
- Spinoza, Baruch. 1985–2016. *The Collected Works of Spinoza*. 2 vols. Edited and translated by Edwin Curley. Princeton NJ: Princeton University Press.
- Suárez, Francisco. 1861. *Disputationes Metaphysicae*. Vols. 25–26 of *Opera Omnia*. Edited by Charles Berton. Paris: Vivès.
- Suárez, Francisco. 1983. *On the Essence of Finite Being as Such, On the Existence of that Essence and Their Distinction. (De Essentia Entis Finiti Ut Tale Est, Et De Illius Esse, Eorumque Distinctione)*. Translated by Norman J. Wells. Milwaukee WI: Marquette University Press.
- Suárez, Francisco. 1995. *On Beings of Reason. (De Entibus Rationis): Metaphysical Disputation LIV*. Translated by John P. Doyle. Milwaukee WI: Marquette University Press.
- van Bunge, Wiep – Krop, Henri – Steenbakkers, Piet – van de Ven, Jeroen (eds.) 2011. *The Bloomsbury Companion to Spinoza*. London: Bloomsbury.
- Wells, Norman J. 1990. Objective Reality of Ideas in Descartes, Caterus, and Suárez. *Journal of the History of Philosophy* 28 (1): 33–61.
- Yakira, Elhanan. 2015. *Spinoza and the Case for Philosophy*. Cambridge: Cambridge University Press.

# VII

*Filip Buyse*

## A New Reading of Spinoza's *Letter 32* to Oldenburg: Spinoza and the Agreement between Bodies in the Universe

Indeed, he [Spinoza] would undoubtedly have been horrified by any suggestion that he and his philosophy are remote from modern science, not just because he spent much time experimenting, studying experiments, and discussing experimental results with scientists, as well as assembling microscopes and telescopes, but still more, because it was basic to his conception of his philosophy that his thought should be firmly anchored in the rules and procedures of mathematics and science (Israel, J.I.: *Radical Enlightenment*, 242).

### 1. Introduction

The starting point of this paper is an apparent paradox in Spinoza's reply to a question – concerning the agreement (or the coherence) between bodies in the universe – that Robert Boyle had addressed in *Letter 31* to the Dutch philosopher via Henry Oldenburg. In the first chapter (2) of this paper, I will indicate what the problem is and put it in context. In the next chapter (3), this paper tries to resolve the paradox by suggesting that Spinoza had applied the mechanical analogy of the synchronization of pendulum clocks. My claim is that although it seems that bodies can determine themselves, they are externally determined by synchronization. Subsequently, this paper gives some arguments (4) in favor of the plausibility of this hypothesis. In the next chapter (5), the difference between the synchronization hypothesis and

Gueroult's pendulum hypothesis will be addressed and opposed to Deleuze's interpretation. Finally (6), this paper gives a reason why Spinoza does not mention the "pendulum clock" in his explanation of *Letter 32*, even though he might have been inspired by the motions of pendulums (clocks) in his conception of the *ratio*<sup>1</sup> of motion and rest of bodies.

## 2. The paradox

In the third definition of the third part of his masterpiece, *The Ethics* (1677), Spinoza defines affects (or emotions)<sup>2</sup> as: "affections of the Body by which the Body's power of acting is increased or diminished, aided or restrained, and at the same time, the ideas of these affections". Affects are thus always affections, however affections are not always affects. New interpretations of Spinoza's work concerning affections are consequently always relevant for the study of the affects, and this is the theme of the bundle wherein we find this article.

According to postulates III and VI of E2 numerous other bodies affect the human body in many ways. Furthermore, for each body there *is* a corresponding idea in God and, simultaneously [*ad simul*] with the affections of the body, the mind (the idea of the body) *has* ideas of these affections. However, regarding the concept of "affections" – and consequently also related to the concept of "affects" or "emotions" – there seems to be a twofold problem in Spinoza's *Letter 32* to Oldenburg. The aim of this paper is to suggest a possible way to resolve this problem.

Robert Boyle (1627–1691) had asked in *Letter 31* – via his good friend Henry Oldenburg – how "knowledge of how each part of Nature agrees with its whole and how it coheres with the others". In his definition of *Mechanical Philosophy*, which he had introduced in the preface of the second part of a book (*Physiological Essays*)<sup>3</sup> that he had sent<sup>4</sup> to Spinoza, Robert Boyle states that a body has only a limited set of intrinsic properties and that these properties had to be explained in terms of the mechanical properties of the corpuscles that compose the body. However, the question remained: how do these parts cohere with each other? In *Letter 32*, Spinoza gives the following definition of the coherence of the parts:

By coherence of parts I mean simply this, that the laws or nature of one part adapts itself to the laws or nature of another part in such wise that there is the least possible opposition between them. On the question of whole and parts, I consider things as parts of a whole to the extent that their natures adapt themselves to one

---

1 The link between the concept of *conatus* will not be discussed in this article although there is obviously a link between Spinoza's *concept* of *conatus* and his concept of *ratio* of motion and rest. Spinoza, however, does not use the term '*conatus*' in the context of his *Letter 32*.

2 For a discussion of the terminology concerning affects and emotions, see: Jaquet 2004, 67–71.

3 *Physiological Essays* should be understood in the same sense as 'physical' or simply as that which concerns nature.

4 In fact, Boyle sent the Latin translation of the *Physiological Essays* to Spinoza. The title of this work is: *Tentamina quaedam physiologica diversis temporibus & occasionibus conscripta*.

another so that they are in the closest possible agreement. Insofar as they are different from one another, to that extent each one forms in our mind a separate idea and is therefore considered as a whole, not a part.

Obviously, what Spinoza writes in the passage above seems to be in contradiction with his radical, metaphysical determinism which is an essential and characteristic part of his philosophy. Indeed, according to E1p28<sup>5</sup>, a body which is determinate or finite (according to the first definition<sup>6</sup> of E2) is necessarily determined by another finite thing. Moreover, according to E1p10 and E2p6 this finite<sup>7</sup> thing should be of the same attribution of the unique and eternal substance. Consequently, a body (a mode of the attribute Extension) is always determined by another body to act in a determinate way. Spinoza expressed this idea in Lemma 3 of the *Physical Interlude* and he referred to it several times in the course of his *Ethics*, highlighting the importance of this principle. In his *Letter 58* (1674) to G.H. Schüller, Spinoza explains his metaphysical determinism through the example of a moving stone, emphasizing that what is true for a stone is true for each individual thing.<sup>8</sup>

By contrast, in the passage quoted above from *Letter 32*, Spinoza does not write that the parts (or natures) of a body are externally caused by other bodies. On the contrary, he writes that they adapt themselves: “the laws or nature of one part adapts itself to the laws or nature of another part”. Moreover, he does not write this once, as though it were a mere exception, but it appears several times: “their natures adapt themselves to one another”, suggesting here a spontaneous, internal cause.

It is important to notice that these phrases are not a slip of the pen, nor are they the result of misleading translations. It is well-known that Spinoza wrote very rigorously and in the original text of the *Ethics*, written in Latin, we read:

---

5 The abbreviations applied for Spinoza's works: E – *Ethics (Ethica)*; CM – *Metaphysical Thoughts (Cogitata metaphysica)*; PPC – *Descartes's Principles of Philosophy (Principia Philosophiae Cartisiana)*, *Letters (Epistolae)*; KV – *Short Treatise on God, Man and his Well-being (Korte Verhandeling van God de Mensch en deszelfs Welstand)*; TP – *Political Treatise (Tractatus-Politicus)* and TTP – *Theological Political Treatise (Tractatus Theologico-Politicus)*.

Passages in Spinoza's *Ethics* will be referred to by means of the following abbreviations: a (axiom), ap (appendix), c (corollary), d (demonstration), def (definition), p (proposition), le (lemma) and s (scholium). For instance: E2p16c2 = Part 2 of the *Ethics*, proposition 16, corollary 2.

Other abbreviations: OCH [Huygens, Christiaan. 1888–1950. *Œuvres Complètes de Christiaan Huygens* (Publ. par la Société hollandaise des sciences). La Haye: M. Nijhoff].

All citations in English from Spinoza's work are translations by Edwin Curley. All translations from Christiaan Huygens' correspondence are from Alex Boxel. These translations can be found on the following site: <http://idolsofthecave.com>

6 In the Latin text, the definition of the body is: “*Per corpus intelligo modum, qui Dei essentiam, quatenus, ut re extensa, consideratur, certo, et determinatio modo exprimit; vid. Coroll. Prop. 25.p.1.*”.

7 Spinoza defines “a finite thing” in E1d2 as: “That thing is said to be finite in its own kind that can be limited by another of the same nature. For example, a body is called finite because we always conceive another that is greater. Thus a thought is limited by another thought. But a body is not limited by a thought nor a thought by a body”.

8 What Spinoza writes here in this letter on freedom and necessity might be inspired by chapter 21 of Hobbes' *Leviathan*, a book translated in 1667 into Dutch by Spinoza's friend Abraham Van Berkel.

Per partium igitur cohaerentiam nihil aliud intelligo, quàm quòd leges, sive natura unius partis ità sese accommodat legibus, sive naturae alterius, ut quàm minimè sibi contrarientur. Circa totum, et partes considero res eatenus, ut partes alicujus totius, quatenus earum natura invicem se accommodat, ut, quoad fieri potest, inter se consentiant, quatenus verò inter se discrepant, eatenus unaquaeque ideam ab aliis distinctam in nostrâ Mente format, ac proinde, ut totum, non ut pars, consideratur. Ex. gr. cum motûs particularum lymphae, chyli, etc. invicem pro ratione magnitudinis, et figurae ità se accommodant, ut planè inter se consentiant, unumque fluidum simul omnes constituent, eatenus tantùm chylus, lympa, etc. ut partes sanguinis considerantur: quatenus verò concipimus particulas lymphaticas ratione figurae, et motûs, à particulis chyli discrepare, eatenus eas, ut totum, non ut partem, consideramus.

It is clear in this original Latin text that bodies are not caused by other bodies to adapt. Obviously, Spinoza writes here that the parts adapt themselves, as he applies the expression *se accommodate* several times.<sup>9</sup> Moreover, this point is also confirmed by what we read in the Dutch translation by his friends and posthumously published under the title *Nagelate Schriften* (1677). In the *Vyftiende Brief*<sup>10</sup> of this work, we read the translation of the passage above.<sup>11</sup>

Obviously, this Dutch translation confirms our interpretation of the definition of coherence in *Letter 32*. This translation also clearly expresses the idea that the laws [*de wetten*] or the nature [*de natuur*] of parts [*de delen*] adjust themselves [*zich voegen naar; zich schikken en voegen naar*] in such a way that there is the least possible opposition between them [*dat zy op het minste tegen malkander strijden*].

The passage from *Letter 32* seems to be very intriguing for yet another reason since – as Albert Rivaud has argued already in his *La physique de Spinoza*<sup>12</sup> – Spinoza gives very

9 The only places Spinoza uses the verb *accommodare* or any variation of it in the *Ethics* are: the corollary of proposition 4 of E4 and, consequently, caput VI and VII of the appendix of E4, and proposition 7 of E5. Interestingly, in each of these cases, Spinoza uses this term in the sense of adapting to the external causes or to the laws of nature to which things are forced to accommodate themselves.

10 This letter corresponds with *Letter 32* in the modern numeration.

11 Cf. *Vyftiende Brief*. In: *Nagelate Schriften* 1677, 495: “By zamenvangend der delen dan versta ik niets anders, dan dat de wetten, of de natuur van een deel zich in dier voegen naar de wetten, of natuur van ‘t ander schikt en voegt, dat zy op het minste tegen malkander strijden. Omtrent het geheel en de delen aanmerk ik de dingen als delen van enig geheel, voor zo veel hun natuur zich onderling in dier voegen schikt, dat zy, zo veel als mogelijk is, met malkander overënkomen: maar voor zo veel zy van elkander verschillen, vormt yder een denkbeelt, van d’anderen verscheiden, in onze ziel, en word dieshalven als een geheel, niet als een deel, aangemerkt. Tot een voorbeelt, gelijk de bewegingen der bijzondere delen van ‘t water, van de gijl, enz. zich naar hun grootheit en gestalte onderling in dier voegen schikken, dat zy gantschelijk met malkander overënkomen, en alle een vloedig lighaam maken; en dus verre worden het gijl, het water, enz. alleenlijk als delen van ‘t bloet aangemerkt. Maar voor zo veel wy de waterdeeltjes naar reden van hun gestalte en beweging bevatten van de deeltjes van ‘t gijl te verschillen, voor zo veel aanmerken wy hen als een geheel, en niet als een deel.”

12 Cf. Rivaud 1924–1926, 24–57.

different explanations on a similar subject (i.e.: the cohesion among parts of a body) in other parts of his work and correspondence. For instance, in Axiom 3 of the *Physical Interlude*, just after his definition of the body, Spinoza writes:

As the parts of an Individual, or composite body, lie upon one another over a larger or smaller surface, so they can be forced to change their position with more or less difficulty; and consequently the more or less will be the difficulty of bringing it about that the Individual changes its shape. And therefore the bodies whose parts lie upon one another over a large surface, I shall call *hard*; those whose parts lie upon one another over a small surface, I shall call *soft*; and finally those whose parts are in motion, I shall call *fluid*.

Interestingly, Spinoza had already developed views similar to what he would express in the *Physical Interlude* as early as 1661/62. In letters known as the Boyle/Spinoza correspondence,<sup>13</sup> transacted after Oldenburg's visit to Spinoza during the summer of 1661, Spinoza writes in *Letter 6* (December 1661) the following explanation:

To understand the first, it must be noted that bodies in motion never meet other bodies with their largest surfaces, whereas bodies at rest lie on others on their largest surfaces.

Thus, in both the *Physical Interlude* (1677) and *Letter 6* to Boyle (1661), Spinoza conceives of the agreement of coherence between the constituent parts of bodies in a purely mechanistic way, in contrast to the explanation of *Letter 32* to Boyle (1665). In other words: coherence is conceived of in terms of relative position, contact, motion and rest of the parts that constitute the whole.<sup>14</sup> Nowhere in this context does Spinoza write that the natures of bodies adapt themselves spontaneously to other bodies in order to form a single unity. On the contrary, in the definition of a body that he gives in the *Physical Interlude*, Spinoza states clearly that a new physical individuality is a whole of bodies "constrained by other bodies" [*a reliquis ita coërcetur*].

In conclusion, the question arises: why does Spinoza, in *Letter 32*, give explanations that seem to violate an important element of his metaphysics, and how does this relate to the other, very divergent explanations that he gives of the same subject on other occasions? How can this puzzle possibly be resolved?

---

13 The correspondence between Baruch Spinoza and Henry Oldenburg is composed of 17 letters from Oldenburg to Spinoza and 10 from Spinoza to Oldenburg. This correspondence was between 1661 and 1676 with hiatuses between 1663 and 1665 and between 1665 and 1675. What is known as the 'Spinoza-Boyle' correspondence forms a part of this larger whole and consists of the letters 6, 11, 13 and 16 written between 1661 and 1663.

14 Cf. Rivaud 1924–1926, 24–57.

### 3. Resolving the paradox

I suggest that the answer to these questions lays in the fact that, when Spinoza wrote his *Letter 32*, he was in contact with Christiaan Huygens (1629–1695), and that he was inspired by a spectacular new discovery that Holland’s most famous physicist had made only a few months before.

In the spring of 1663, Spinoza moved from Rijnsburg to Voorburg where the Huygens family had a residence, named “Hofwijck”. Based on what Spinoza writes in his *Letter 26* (dated May 1665), there is historical evidence that Spinoza visited Hofwijck in that period and that he had (especially around 1665) discussions with the most famous Dutch physicist of that time concerning new scientific publications such as Hooke’s *Micrographia* (1665) and numerous scientific topics. Moreover, several Spinoza scholars such as J. Israel, W. Klever and A. Rivaud, argued that it is not incorrect to assume that Spinoza was informed of all Huygens’ activities during that period.<sup>15</sup>

But what kind of discovery did Huygens make in that period that might be relevant for the interpretation of *Letter 32*? Around the 22 February 1665, while Christiaan Huygens was sick and lying in his bed, he observed that two pendulum clocks, which were hanging in front of him, started to beat in synchronicity. He couldn’t believe his eyes. Initially, he was unable to explain this phenomenon and referred to it as “some sort of sympathy” [*une espèce de sympathie*]. He struggled to find a causal explanation for this effect. Why did the clocks mysteriously synchronize with each other? How could mechanical objects transmit an influence when they were not touching? What is the cause of this “odd kind of sympathy”?

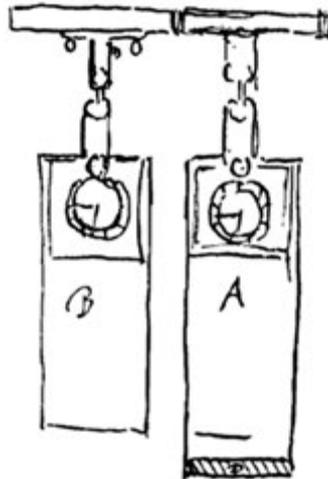


Figure 1<sup>16</sup>

15 Cf. Israel 2001, 246–252; Keesing 1985; Klever 1997, 14–31; Nadler 1999, 221–222 and van Bunge 2011, 16–24.

16 Original drawing by Christiaan Huygens (Cf. OCH XVII 183).

Initially, Huygens was convinced that there could not be any other cause of the agreement of the clocks “*than an imperceptible agitation of the air which is produced by the movement of the pendulums*”. However, a few days later he wrote with a pencil in the margin of his notes: “*causam hujus rei postea inveni ex communi fulcro*”. More precisely, *on the first of March*, he conducted some additional experiments and determined that not the air, but rather the mechanical connection between the two clocks was essential for their synchronization.

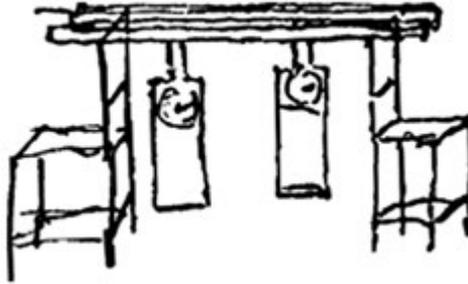


Figure 2<sup>17</sup>

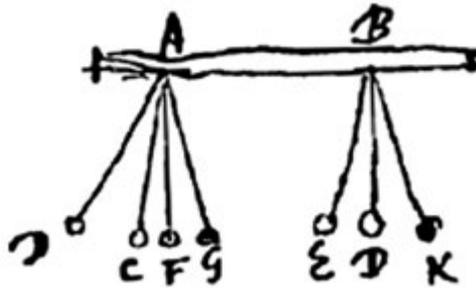


Figure 3<sup>18</sup>

Huygens was the first physicist to observe and analyze this phenomenon which is known today as synchronization. But what, precisely, is synchronization and how does it differ from other phenomena such as resonance? A. Pikovsky defines synchronization as “an adjustment of the rhythms of oscillating objects due to a small interaction”<sup>19</sup>. It is important to note that synchronization is about two or more objects which are already in motion. Consequently, this effect is different from resonance where one of the objects may be at rest as illustrated by the resonance of the Tacoma Bridge which collapsed on 9 November 1940<sup>20</sup> as a result of elementary forced resonance with the wind providing an external periodic frequency that matched the natural structural frequency.

17 Drawing from Huygens (OCH XVII 185, Fig.76).

18 Drawing from Huygens (OCH XVII 185, Fig.77).

19 Cf. Pikovsky *Synchronization*, 8–14.

20 Cf.: <https://www.youtube.com/watch?v=nkXl8JJBH7E> or <https://www.youtube.com/watch?v=3mclp9QmCGs>

Synchronization is not a state. On the contrary, it is a dynamic, equilibrium process.<sup>21</sup> Today we can demonstrate this phenomenon perfectly well by placing metronomes<sup>22</sup> upon a thin board that is isolated from a table using, for example, empty soda cans. The synchronization takes place with two or more metronomes which were initially in motion but which need not be in phase. Additional metronomes can be added which, after a brief time, will also start to synchronize. Even if the synchronized system is disturbed (for example, by touching one of the metronomes to block its motion), the synchronization will spontaneously restore itself after a suitable interval. In sum, the pendulums have the capacity to adapt themselves so that they form a unity. As Kurt Wiesenfeld puts it: “The phenomenon of spontaneous mutual synchronization offers perhaps the most primitive example of emergent behavior”<sup>23</sup>. Needless to say, this behavior perfectly suits Spinoza’s description of the behavior of parts that adapt themselves in his definition of coherence in *Letter 32*.<sup>24</sup>

As an interesting aside we should note that Bennett *et al.*<sup>25</sup> from the Georgia Institute of Technology have reconstructed Huygens’ clocks and re-examined the 350-year-old synchronization in modern experiments. This research has confirmed that, in principle, Huygens’ analysis was correct. As Huygens had suggested, the antiphase synchronization he observed was the result of the weak communication between the two similar, heavy, oscillating pendulum clocks via the board. As Huygens’ drawings illustrate, the clocks were attached to a common supporting beam which itself was supported on the backs of two chairs.

## 4. The plausibility of the synchronicity hypothesis

The hypotheses that Spinoza was inspired by Huygens’s discovery of the pendulum clocks and that he applied it as a mechanical analogy is plausible for at least four reasons, which we will present here.

First of all, there is the location. As already mentioned, Spinoza was in contact with Huygens during that period and he was informed directly of Huygens’ work. Consequently, he must have been informed of Huygens’ discovery of synchronization because the famous mathematician did not hide his discovery. On the contrary, he made it public. Directly after his discovery, he informed several leading people in Holland and abroad. And almost directly afterwards, on the 16<sup>th</sup> and 23<sup>rd</sup> of March, he published his discovery

---

21 Bennett et al. 2002, 563–579.

22 Pantaleone 2002, 992–1000. For a simulation of this experiment, please visit the site of the Harvard Natural Sciences Lecture Demonstrations at: <https://www.youtube.com/watch?v=Aaxw4zbULMs>

23 Wiesenfeld – Borrero-Echeverry 2011, 047515.1; Bennett et al. 2002, 563–579.

24 Cf.: <https://www.youtube.com/watch?v=nkXl8JJBH7E> or <https://www.youtube.com/watch?v=3mclp9QmCGs>

25 Cf. Bennett et al. 2002, 563–579; Czolczynski et al. 2011, 2047–2056.

in Europe's earliest scientific journal, the *Journal des savants*.<sup>26</sup> Huygens' discovery was also discussed in a meeting of the Royal Society of which he had been made a member in 1663. He had informed the other members via his contact person R. Moray. As the *Journal Book of the Royal Society* shows, they did discuss the question during their meeting on the first of March (which corresponds to March 10<sup>th</sup> in the Gregorian calendar). The minutes of the Royal Society make clear that members decided during this meeting that the president (mathematician William Brouncker) and Robert Hooke (1635–1703) would devise and conduct some experiments to determine whether the cause of the odd sympathy was the “agitation of the air” or a “magnetical cause”. Henry Oldenburg was present<sup>27</sup> at this meeting so that we may assume that he knew of the “sympathie des horloges” at the moment that he and Boyle addressed their question concerning the coherence between the bodies to Spinoza. Consequently, the members of the *Republic of Letters* knew about it. Hence, it is plausible to assume that Spinoza must have known about it and might have been inspired by it given the fact that he was interested in the physics of the pendulum clock illustrated by the fact that the *Horologium oscillatorium*<sup>28</sup> (1673) was found in his personal library<sup>29</sup> after his death.

Spinoza visited Hofwijck and discussed several scientific topics with Huygens. At that moment, the Dutch philosopher already had a strong background in physics because he had taught *mathesis* to university students<sup>30</sup> while he lived in Rijnsburg and had written a book (the only book he would publish under his name during his lifetime) which had started as a dictate of a private course on the second part of Descartes' *Principia* (1644). As *Letters* 26 and 30 show, Holland's most famous physicist informed him of brand new scientific publications such as Hooke's *Micrographia* and Boyle's *Experiments and Considerations Touching Colours*.

Spinoza and Huygens must have also discussed the latter's experiments with the pendulum clocks. In *Letter* 26 to Oldenburg, his comments on Kircher's book, which he had seen in Huygens' residence, clearly illustrate that he knew about the competition between several scientists to find the right applications for several types of clock. Spinoza even gives his own interpretation of Huygens' lack of enthusiasm for the publication of the author of the *Mundus subterraneus*. According to Spinoza, it must have had something to do with the fact that Athanasius Kircher (1602–1680) did not believe that the pendulum clock could be used for the determination of longitude at sea:

26 See the *journal des savants* of the 16<sup>th</sup> of March 1665 and the 23<sup>rd</sup> of March.

27 Cf. Birch 1756, 18.

28 Christiani Hugenii Zulichemii, Const. f., *Horologium oscillatorium, siue, De motu pendulorum ad horologia aptato demonstrationes geometricae* / Christiaan Huygens. – Parisiis: apud F. Muguet, 1673. The clock was first described in Huygens' small book of 1658 *Horlogium*. His major work, *Horologium Oscillatorium* (The Pendulum Clock or Geometrical Demonstrations Concerning the Motion of Pendulums as Applied to Clocks) was published 15 years later in 1673.

29 Cf. number 62, in: *Catalogus van de bibliotheek der vereniging Het Spinozahuis* (Leiden: E.J. Brill, 1965), 24.

30 Cf. Bontekoe, C.: *Brief Aan Johan Frederik Swetser, Gesegt Dr. Helvetius, Geschreven en uytgeeven tot een Korte Apologie voor den Grote Philosoph Renatus Descartes* [...], ('s Gravenhage: 1680).

I have seen Kircher's Subterranean World at Mr. Huygens'. He praises Kircher's piety, but not his ability. I don't know whether this is because Kircher treats pendulums, and concludes that they will not help at all to discover longitudes (which is completely opposed to Huygens' opinion).<sup>31</sup>

Oldenburg had repeatedly asked Spinoza questions about Huygens' new pendulums, instead of asking Huygens directly. The co-secretary of the Royal Society had probably noticed Spinoza's interest in the new science during his visit in the summer of 1661. However, he could not tell Oldenburg anything definite, though he knew that the clocks were not a commercial success at that time, probably because of the high cost:

You want to know what Our People think of Huygens' new Pendulums. As yet I can't tell you anything definite about this. Still, I know this: the craftsman who has the exclusive right to make them is completely giving up the work, because he can't sell them. I don't know whether this is because commerce has been interrupted [by the war] or because he's trying to sell them at too high a price. He's asking 300 Caroline florins each.<sup>32</sup>

This illustrates that the Dutch philosopher was well aware of the importance of pendulum clocks in 1665. Pendulum clocks and their applications were a hot topic for Spinoza and his circle.

Of course, there is also the date. Spinoza wrote his *Letter 32* in November 1665, just a few months after his neighbor, Christiaan Huygens, had discovered "the sympathy of the clocks". Moreover, according to his diary,<sup>33</sup> Huygens was still busy doing research on this subject when Spinoza visited him in May of 1665 and wrote his *Letter 32* in November of the same year.

Secondly, there is the fact that "clocks" became a paradigmatic example for understanding the supposed agreement or harmony between bodies in the universe, at least for philosophers in the region where Spinoza lived around 1665. There was already a longstanding tradition<sup>34</sup> among natural philosophers going back to innovators such as Leonardo da Vinci, Giovanni de' Dondi (ca. 1330–1388), Richard of Wallingford (1292–1336),<sup>35</sup> and Oresme (ca. 1323–1382), to use the mechanical clock as a metaphor for understanding the universe and natural phenomena. And, it is well documented that later also Kepler, Descartes, Malebranche and other early modern philosophers applied this metaphor.

However, this tradition was completely renewed after 1658 when Christiaan Huygens, with the help of his technician, Salomon Coster (ca. 1622–1659), invented and constructed his pendulum clocks, which were of a much better quality than former versions of the

---

31 *Letter 30* from Spinoza to Henry Oldenburg, Voorburg, 1 October 1665.

32 *Ibidem*.

33 See OCH XVII, 187, note 3.

34 See Wootton 2016, 437–441, 484–486.

35 See North 2004.

mechanical clock. The clock was in the air so to speak! People were enthusiastic about this discovery and they hoped that this sophisticated thing would solve the long existing problem of determining longitude at sea.

Clocks were not only used by astronomers and physicists. They became a fruitful metaphor for many natural philosophers. The distinction between the artificial and the natural had been essential for Aristotelians just as the distinction between the sublunary world of change and the perfectly constant world of the celestial bodies making perfect circular motions had been essential.<sup>36</sup> However, the destruction of these distinctions after Galileo's observations opened the door for thinkers to make analogies between man-made machines (such as pendulum clocks) and mechanical analogies for understanding the homogenized universe. Kepler, who Spinoza mentions in a note that he adds to his *Letter 32*, wrote already in 1605:

My aim is this, to show that the celestial machine is not like a divine creature, but like a clock (he who believes the clock to be animate assigns the glory of the artificer to the work), insofar as nearly all the diversity of motions are caused by a simple, magnetic and corporeal force, just as all the motions of a clock are caused by a most simple weight. I will also show how this physical account is to be brought under mathematics and geometry.<sup>37</sup>

More importantly, in light of the theme of this article, thinkers in Holland around 1665 used the clock to explain the agreement between things. A good example to illustrate this is Comenius [Jan Amos Komenský (1592–1670)]. The Czech philosopher moved to Amsterdam in 1656 and only a few years later, just after Huygens had invented his new pendulum clock, he began to apply the metaphor of this clock in his philosophical and pedagogical work. More precisely, in his *Opera didactica omnia* (1657), he not only applied this mechanical analogy to explain the agreement between parts in the universe but also to clarify the harmony between the parts of a body. Moreover, he applies the clock to the human spirit and soul in detail.<sup>38</sup> In his analogy, the escapement is reason, the most important wheel is the will, the weights are the desires and the affections and the less important wheels are the other movements of the soul. Reason balances the desires and affection so that the agreement of virtues follows. Obviously, this is a good example of how around 1660 the pendulum clock became a paradigmatic example for understanding the agreement and harmony between parts.

Just as the great world itself is like an immense piece of clockwork put together with many wheels and bells, and arranged with such art that throughout the whole structure one part depends on the other, and the movements are perpetuated and

---

36 See Bensaude-Vincent – Newman 2007.

37 Letter dated in 1605 to Herwart von Hohenburg, translation from Snobelen.

38 Comenius. 1657. *Opera didactica omnia* (Oeuvres didactiques complètes) D 42; Comenius 1657/1910, 47–48.

harmonised; thus it is with man [...]. The weight, the efficient cause of motion, is the brain, which by help of the nerves, as of ropes, attracts and repels the other wheels or limbs, while the variety of operations within and without depends on the commensurate proportion of movements.

In the movements of the soul the most important wheel is the will; while the weights are the desires and affections which incline the will this way or that. The escapement is the reason, which measures and determines what, where, and how far anything should be sought after or avoided. The other movements of the soul resemble the less important wheels which depend on the principal one. Wherefore, if too much weight be not given to the desires and affections, and if the escapement, reason, select and exclude properly, it is impossible that the harmony and agreement of virtues should not follow, and this evidently consists of proper blending of the active and the passive elements.

Man, then, is in himself nothing but a harmony [...] as in the case of a clock or of a musical instrument which a skilled artificer has constructed.<sup>39</sup>

Another, more well-known example from the same period comes from Arnold Geulincx (1624–1669). In the annotations to the *Ethica*, the Flemish philosopher (who moved to Leiden in 1658) compares the mind and the body with two clocks:

My will does not move the mover to move my limbs; but the one who has imparted motion to matter and has laid down the laws to it, the same one has formed my will and thus has closely united these very dissimilar things (the motion of matter and the determination of my will), so that, when my will wishes, motion of the desired kind is present, and, on the contrary, when the motion is present, the will has willed it, without any causality or influx from one to another. Just as with two clocks that agree with each other and with the daily course of the sun: when one sounds and indicates the hours to us, the other sounds in the same way and indicates the same number of hours to us; not because there is any causality from one to the other, but because of the mere dependence, in which both are constructed by the same art, and by similar activity.<sup>40</sup>

In this passage, Geulincx clarifies the agreement between the determination of the will (the mental) and the motion of matter (the physical) by means of the mechanical analogy of two clocks that agree with each other “without any causality or influx from one to another”. The agreement or the mere dependence is due to, he argues, a similarity in art and construction. Obviously, this explanation echoes Huygens’ observation and analysis. The application of the analogy to the understanding of the mind/body problem is, however, much better known from the explanation that Leibniz gave to Basagne de Beauval in

39 Comenius 1657/1910, 47–48; McReynolds 1980, 99 and Matthews 2000, 217.

40 Annotata ad Ethicam, Sämtlichen Schriften [10.24], 3: 211–212; cf. Annotata ad Metaphysicam, Sämtlichen Schriften [10.24], 2: 307.

1696; except Geulincx had already articulated this explanation more than 25 years earlier. In Leibniz's version, we find a much more explicit reference to "the way with which Mr. Huygens experimented", with results that surprised him:

Imagine two clocks or watches which are in perfect agreement. Now this can happen in *three ways*. The *first* is that of a natural influence. This is the way with which Mr. Huygens experimented, with results that greatly surprised him. He suspended two pendulums from the same piece of wood. The continued strokes of the pendulums transmitted similar vibrations to the particles of the wood, but these vibrations could not continue in their own frequency without interfering with each other, at least when the two pendulums did not beat together. The result, by a kind of miracle, was that even when their strokes had been intentionally disturbed, they came to beat together again, somewhat like two strings tuned to each other. The *second* way of making two clocks, even poor ones, agree always is to assign a skilled craftsman to them who adjusts them and constantly sets them in agreement. The *third* way is to construct these two timepieces at the beginning with such skill and accuracy that one can be assured of their subsequent agreement. Now put the soul and the body in the place of these two timepieces. Then their agreement or sympathy will also come about in one of these three ways.<sup>41</sup>

Third, there is the argument that in his definition of *Letter 32* Spinoza refers to laws which he identifies with the nature of parts: "the laws *or* nature of one part adapts itself to the laws or nature of another part". Why does Spinoza mention "law" in this context? And, what is the link with "the nature of a part"? The reason might be that he is referring here to an early version of the pendulum law. This law expresses the relation between the natural frequency of a pendulum and the elements that determine this frequency: the length of the pendulum and the acceleration of gravity. Galileo had formulated in his *Dialogo* (1632)<sup>42</sup> the concept of isochronism, arguing that it is not the mass of the pendulum but the length that is decisive. But, it was Christiaan Huygens who discovered the pendulum law in 1673. This law was – according to Vincent Icke<sup>43</sup> – the first modern physical law expressed in the form of a mathematical formula. It is therefore possible that Spinoza had an anticipation of this physical law in mind when he discussed the agreement between the parts of a whole in his *Letter 32*.

---

41 G IV, Leibniz to Basagne, 1696, 498–500.

42 Galileo 1632, 267.

43 Cf. Icke, Vincent. 2005. *Christiaan Huygens in de onvoltooid verleden toekomstige tijd*. Groningen: Historische Uitgeverij.

## 5. Gueroult's hypothesis of the pendulum

The hypothesis that the synchronization of the pendulum clock inspired Spinoza corresponds with the hypothesis that Martial Gueroult formulated in the 1970s.<sup>44</sup> In his influential book *Spinoza II – l'Âme* (1974) the French structuralist claimed that Spinoza conceived of the simplest bodies [*corpora simplicissima*] as single pendulums, arguing that this hypothesis is compatible with all the axioms and lemmas<sup>45</sup> of the *Physical interlude* that treat the simplest bodies. Moreover, Gueroult claimed that Spinoza conceived of complex bodies as complex pendulums.

However, for Gueroult, it was not the pendulum as such that is important here but rather the kind of motion that the pendulum typically represents, namely, what is known today as harmonic oscillation. Indeed, just like single pendulums, the simplest bodies are either in motion or at rest (*axiom 1*); each can move at varying speeds (*axiom 2*); and they are distinguished from one another in respect to motion-and-rest, quickness and slowness, and not in respect to substance (*Lemma 1*). According to Huygens' formula describing the period of a pendulum, the oscillating time depends only upon the local gravitational acceleration and the length of the pendulum and not on the mass or the material of the oscillating bodies. Furthermore, the simplest bodies agree in certain aspects (*Lemma 2*) in the same way that a pendulum's motion or rest must be determined by another body. In other words, just like pendulums, the simplest bodies cannot oscillate on their own (*Lemma 3*); rather, an external cause must set them in motion or bring them to a halt. Hence M. Gueroult concludes that we should conceive of *corpora simplicissima* as corpuscles that are in perpetual vibration and that the nature of their vibration constitutes the fundamental essence of their individuality.<sup>46</sup>

It is not just the conception of the simplest body that leads us to the pendulum, but the origin of Spinoza's conception of the complex body also owes something to the pendulum. Just after the lemmas' treating of the simplest bodies, Spinoza gives his definition of a body. According to this definition a body is a whole of modally distinct parts participating in the same mutual relation of a physical nature. Among others, François Zourabichvili<sup>47</sup> (referring to Gueroult) has argued that the mutual relation, or the *ratio* of motion and rest, is neither to be understood as an exact number nor as an exact mathematical proportion. According to Zourabichvili, it should rather be understood as a norm.

But, where does his idea of the *ratio* of motion and rest come from? Again, we can learn something here from the proto-*Ethica* where the young Spinoza was still under a greater Cartesian influence. In the preface of the second part of this book, he writes:

---

44 Cf. Gueroult 1974, 159–165.

45 Cf. axioma 1 and 2, and lemma 1, 2 after Ep13s.

46 Cf. Gueroult 1974, 161.

47 See not only M. Gueroult but also and more recently Zourabichvili 2002, 53: "C'est bien ce que Spinoza a en vue lorsqu'il donne son exemple: le corps peut être soumis « à un changement constant mais non à un si grand qu'il dépasse la limite de 1 à 3". Thus the relation does not concern the equilibrium of two quantities but the amplitude of an oscillation: the relation between a maximum of rest and a maximum of movement, and "si d'autres corps agissent sur le nôtre si puisement que la proportion de 1 à 3 de son mouvement ne puisse subsister, alors c'est la mort".

So if such a body has and preserves its proportion – say of 1 to 3 – the soul and the body will be like ours now are; they will, of course, be constantly subject to change, but not to such a great change that it goes beyond the limits of from 1 to 3, and as much as it changes, so also the soul changes each time.<sup>48</sup>

As the earlier Dutch version of the work makes especially clear, the *ratio* of a body rather indicates the borders of an oscillating or vibrating motion around an equilibrium state, which may not be surpassed by the parts belonging to the same physical individuality. Hence, this ratio indicates a norm rather than a number.

Gilles Deleuze severely criticized Gueroult's thesis. Deleuze argued that it is absurd to think that Spinoza conceived of the simplest bodies as pendulums. The author of *Spinoza et le problème de l'expression* (1968) argued that these simplest bodies are infinitely small and have therefore no size and no form. Hence, it is absurd to think that they are a kind of pendulum. In Deleuze's interpretation, while the simplest bodies are the final terms of an analysis, they are not "atoms" in the sense of finite existing bodies either. Rather, Deleuze argues that Spinoza's simplest bodies are indeed the end of an analysis (or the beginning of a synthesis), but that they are infinitely small. They are the actual existence of the infinite ("*L'infini en acte*") and are therefore different than the infinite in the Cartesian sense of being infinitely divisible.

In Deleuze's interpretation the simplest bodies have no essence and no real existence outside complex bodies, which are wholes composed of an infinite number of infinitely small simple bodies. Deleuze argues that these simplest bodies should be conceived of as differentials. The concept of a "differential" was already a fundamental concept in the first half of the 17<sup>th</sup> century, although it became a well-known mathematical concept only later via the work of Leibniz and Newton. In this view, it only makes sense to speak of a differential 'dx' on condition that there is an x. Interestingly, this idea is also present in Spinoza's *Letter 32*. Here the Dutch philosopher emphasizes that it makes no sense to speak about the parts of chyle without the concept of chyle; just as it makes no sense to speak about the parts of blood without the whole of blood.

It is possible, however, to reconcile Gueroult's thesis with Deleuze's criticism. As the French structuralist writes in the conclusion of his chapter on the pendulum hypothesis, Spinoza's simplest bodies should be conceived of as oscillating bodies vibrating around an equilibrium position. Of course, the simplest bodies are not pendulum clocks. However, in Gueroult's interpretation, Spinoza applied the metaphor of the pendulum clock to illustrate the fact that the simplest bodies make harmonic oscillations. Adding Deleuze's views to this, the simplest bodies are transformed into vibrating, infinitely small bodies or "points".

---

48 In the Dutch version of the text we read: "12. Zoodanig een lichaam dan, dese zijne proportie, als e.g. van 1. tot 3, hebbende en behoudende, zo zal de ziel en 't lichaam zijn gelijk het onze nu is zijnde wel gestadig verandering onderworpen, maar niet zo groot dat ze buiten de palen van 1. tot 3. gaat; dog zo veel het verandert, zo veel verandert ook telkens de ziel".

## 6. Why does Spinoza not mention the pendulum clock?

At first glance, Spinoza seems to exclude the pendulum clock from his thinking and writing. Moreover, he seems to mock Huygens' clocks in *Letter 30* (1665), telling Oldenburg that Huygens cannot sell them because they are too expensive. Moreover, there seems to be a kind of diminishment in Spinoza's application of the pendulum analogy. The Dutch philosopher applies the pendulum clock analogy solely in his proto-*Ethics*, the *Short Treatise* (ca.1660), but the clock disappears in the corresponding passage of the *Ethics* (1677).

More precisely, in the second chapter of the first part of the *Short Treatise*, Spinoza writes the following:

From all that we have said so far it is clear that we maintain that extension is an attribute of God. Nevertheless, this does not seem possible at all in a perfect being. For since extension is divisible, the perfect being would consist of parts. But this cannot be attributed to God, because he is a simple being. Moreover, when extension is divided, it is acted on; and that too cannot in any way be the case in God (who is not susceptible of being acted on, and cannot be acted on by any other being, since he is the first efficient cause of everything).

To this we reply:

1. That part and whole are not true or actual beings, but only beings of reason; consequently, in Nature there are neither whole nor parts.
2. A thing composed of different parts must be such that each singular part can be conceived and understood without the others. For example, in a clock that is composed of many different wheels, cords, etc., I say that each wheel, cord, etc., can be conceived and understood separately, without needing [the understanding of] the whole as a whole. Similarly, with water, which consists of straight, oblong particles. Each part of it can be conceived and understood, and can exist, without the whole. But since extension is a substance, one cannot say of it that it has parts, since it cannot become smaller or larger, and no parts of it could be understood separately. For in its nature it must be infinite.

In this chapter, Spinoza discusses some of the most essential concepts of his metaphysics. One of these subjects is the part/whole relation, which is also the main theme of *Letter 32*. In the citation above, the Dutch philosopher argues that – modally speaking – parts of a whole can be conceived of separately and can exist on their own. Importantly, he gives two examples to illustrate his views: the parts of water that can exist separately from the whole “water” and the parts (wheels, cords, etc.) of a pendulum clock which can be conceived of and exist separately. To this idea, Spinoza contrasts the concept of “corporeal

substance". In contrast to the parts of a whole body of water and a pendulum clock, the corporeal substance as substance is a continuum. There are no parts that can be conceived or exist separately.

But of course in Spinoza's metaphysics of the *Ethics*, all physical objects are modes of the attribute extension of the unique substance. In the corresponding part of the *Ethics*, in E1p15, Spinoza no longer sets up a contrast between water and the clock to the corporeal substance. On the contrary, he treats "water" as a modus of the corporeal substance and argues that there are two ways to look at this parts/whole relation:

This will be sufficiently plain to everyone who knows how to distinguish between the intellect and the imagination – particularly if it is also noted that matter is everywhere the same, and that parts are distinguished in it only insofar as we conceive matter to be affected in different ways, so that its parts are distinguished only modally, but not really. For example, we conceive that water is divided and its parts separated from one another – insofar as it is water, but not insofar as it is corporeal substance. For insofar as it is substance, it is neither separated nor divided. Again, water, insofar as it is water, is generated and corrupted, but insofar as it is substance, it is neither generated nor corrupted. And with this I think I have replied to the second argument also, since it is based on the supposition that matter, insofar as it is substance, is divisible, and composed of parts.

The first way is to conceive of the whole (water) as water (modally), which implies that the parts can be thought of individually and can exist separately. The second way is to conceive the same body of water as substance, which entails that it cannot exist on its own.

Amazingly, the example of the pendulum clock disappeared from Spinoza's explanation in the corresponding passage from the *Ethics*. Why? In Spinoza's metaphysics, generally speaking, there is a rejection of the Aristotelian distinction between the artificial and the natural. Natural things such as water are not more or less natural than artifacts produced by men since men are just parts of nature like all other things. As a consequence, modes such as machines designed and produced by men are no less natural than other modes like water, trees or flowers.<sup>49</sup> However, Spinoza is very cautious about applying artifacts as models for natural phenomena. For example, he does not use the pendulum clock in the *Ethics* although he could have. Moreover, it is very hard to find texts where he actually makes use of this kind of analogy with man-made machines. Spinoza differs from his contemporaries Huygens, Descartes, Leibniz and Boyle, who all evidently applied the machine metaphor.

Why does Spinoza not mention the pendulum clock? It is possible to figure this out by drawing from the appendix of *De Deo*. The main problem that Spinoza has with artifacts is that they are designed and constructed by a creator who is distinct from his creation and who has a certain purpose (*telos*). For instance, Huygens and his technician constructed a pendulum clock in order to measure time and longitude at sea. In this appendix, Spinoza

---

<sup>49</sup> See E3p2s.

radically rejects the idea that a human body or nature as a whole could be conceived as a man-made machine, although he had applied the analogy of the automaton (probably thinking of the pendulum clock) to explain the functioning of the human mind in chapter 85 of his *The Emendation of the Intellect*. He wrote already in *Letter 4* (1661) that man is not created but generated: “I ask you, my friend, to consider that men are not created, but only generated, and that their bodies already existed before, though formed differently”. In summary, Spinoza’s anti-finalism might be the most important reason why he does not explicitly mention the machine analogy in general and the pendulum clock in particular.

## 7. Conclusion

To summarize, the spectacular observations made by Christiaan Huygens in February 1665 concerning self-synchronizing pendulum clocks may provide a context for understanding Spinoza’s otherwise paradoxical *Letter 32* to Henry Oldenburg, written in November of 1665. In this letter, Spinoza appears to imply that bodies can adapt themselves to other bodies in a non-mechanistic way, and absent the agency of an external cause – a claim that is completely contradictory with the metaphysical determinism that is an important and characteristic element of Spinoza’s philosophy.

Christiaan Huygens was the first to observe the phenomenon of synchronization in two pendulum clocks. Synchronization is a phenomenon whereby oscillators that are appropriately coupled together will adjust their oscillations so as to exhibit a synchronous motion that is regulated by weak impulses communicated through their mutual coupling. Thus the synchronized oscillators appear to behave as if they had spontaneously adapted themselves [*se accommodat*] to each other without any corporeal contact. They appear to act as if they “feel each other” or “communicate with each other” at a distance. Consequently, despite his strongly mechanistic worldview, Huygens initially referred to this phenomenon as “the sympathy of clocks”, which seems to suggest a kind of action-at-a-distance as a result of a hidden, or ‘occult’, quality inherent in the clocks.

Huygens did not hide his discovery. On the contrary, he directly notified his father and the fellows of the Royal Society. Moreover, his discovery was published only a few weeks later in the first scientific journal, the *Journal des savants*. Consequently, within the space of a month, the entire *République des lettres* was made aware of this odd phenomenon. Therefore, it is not unreasonable to assume that Spinoza was likewise aware of Huygens’ observations, all the more so since, according to *Letter 26* (1665), the Dutch philosopher visited Huygens in Voorburg during that time.

This may explain the context within which Spinoza wrote “I consider things as parts of a whole to the extent that their natures adapt themselves to one another so that they are in the closest possible agreement”. This assertion seems not only to represent a discontinuity in his views on the nature of the relations between bodies but also – and more importantly – it seems to be inconsistent with his metaphysical determinism. However,

the paradigm of the synchronization of bodies explains that bodies can “adapt themselves” to other bodies. Moreover, the effect is completely explainable in terms of the mechanistic model of the collision of bodies so that the phenomenon is entirely compatible with Spinoza’s mechanistic views as well as his concept of the causality of bodies. Moreover, this hypothesis allows us to explain why Spinoza speaks of “the laws” which adapt themselves. It is likely that he had Huygens’ law of the pendulum in mind. By this line of reasoning, each pendulum has its own natural frequency, its own law, so to speak, which adapts itself slightly during the process of synchronization to the law of other pendulums in order to form one law, that is, one whole united by a mutual relation of motion and rest.

## Bibliography and abbreviations

- Baker, Gregory L. 2011. *Seven Tales of the Pendulum*. Oxford: OUP.
- Bennett, Matthew – Schatz, Michael F. – Rockwood, Heidi – Wiesenfeld, Kurt. 2002. Huygens’s clocks. *Proceedings: Mathematical, Physical and Engineering Sciences* 458 (2019): 563–579.
- Bensaude-Vincent, Bernadette – Newman, William R. (eds.) 2007. *The Artificial and the Natural: An Evolving Polarity*. Cambridge: The MIT Press.
- Birch, Thomas. 1756. *The History of the Royal Society of London for Improving of Natural Knowledge from Its First Rise, in which the Most Considerable of Those Papers Communicated to the Society, which Have Hitherto Not Been Published, are Inserted as a Supplement to the Philosophical Transactions*. Vol. 2. London: A. Millar in the Strand. [*The History of the RS*]
- Boyle, Robert. 1999. *The Works of Robert Boyle*. Edited by Michael Hunter and Edward B. Davis. London: Pickering & Chatto.
- Boyle, Robert. 2001. *The correspondence of Robert Boyle*. Edited by Michael Hunter, A. Clericuzio and L.M. Principe. London: Pickering & Chatto.
- Chareix, Fabien. 2014. Le bal des pendules: Spinoza et Leibniz face à la mécanique théorique de Huygens. In: Moreau, P.F. et al. (ed.): *Spinoza et Leibniz*. Paris: PUPS. ch. XII.
- Comenius, Jan Amos. 1657/1910. *The Great Didactic*. Translated from the Latin, and edited by M.W. Keatinge, London: Adam and Charles Black.
- Czolczynski, Krzysztof. et al. 2011. Huygens’ odd sympathy experiment revisited. *International Journal of Bifurcation and Chaos* 21 (07): 2047–2056.
- Deleuze, Gilles. 1981. *Cours Vincennes: the actual infinite-eternal, the logic of relations*. (10/03/1981). Confrontation with Gueroult’s commentary. <https://www.webdeleuze.com/textes/42> (accessed May 06, 2017).
- Gullemeau, Evelyne. 2007. El paradigma pendular en la teoría spinozista de los afectos. In: García, Eugenio Fernández – Cámara García, María Luisa de la. (ed.): *El gobierno de los afectos en Baruj Spinoza*. Madrid: Trotta. 93–106.
- Galileo, Galilei. 1632. Dialogo di Galileo Galilei sopra i due massimi sistemi del mondo tolomaico e (Florence: Landini). English translation by S. Drake. In: *Dialogue Concerning the World Systems, Ptolemaic and Copernican*. 2nd edition, 1967 with a foreword by Albert Einstein. Berkeley: University of California Press.

- Gueroult, Martial. 1974. *Spinoza II – L'âme*. Paris: Editions Aubier-Montaigne.
- Huygens, Christian. 1888–1950. *Œuvres Complètes de Christiaan Huygens*. (Publ. par la Société hollandaise des sciences.) La Haye: M. Nijhoff. (OCH)
- Israel, Jonathan. 2001. *Radical Enlightenment*. Oxford: Oxford University Press.
- Jaquet, Chantal. 2004. *L'Unité du corps et l'esprit (Affects, actions et passions chez Spinoza)*. Paris: Presses Universitaires de France. 67–71.
- Keesing, Elisabeth. 1985. Les frères Huygens et Spinoza. (Paris: Éditions République.) *Cahiers Spinoza* 5: 109–128.
- Klarreich, Erica G. 2002. Huygens's clocks revisited. *American Scientist* 90 (4): 322–323.
- Klever, Wim. 1997. Spinoza en Huygens. Een geschakeerde relatie tussen twee fysici. *Gewina* 20 (1): 14–31.
- Matthews, Michael R. 2000. *Time for Science Education. How Teaching the History and Philosophy of Pendulum Motion Can contribute to Science Literacy*. Dordrecht: Springer Netherlands.
- McReynolds, Paul. 1980. The Clock Metaphor in Psychology. In: Nickles, T. (ed.): *Scientific Discoveries: Case Studies*. Dordrecht: Reidel Publishing Company. 97–112.
- Nadler, Steven. 1999. *Spinoza: A Life*. Cambridge: Cambridge University Press.
- North, John. 2004. *God's Clockmaker: Richard of Wallingford and the Invention of Time*. Oxford: Oxbow Books.
- Pantaleone, Jim. 2002. Synchronization of Metronomes. *American Journal of Physics* 70 (10): 992–1000.
- Parkinson, George H.R. (ed.) 1993. *Routledge History of Philosophy Volume IV: The Renaissance and 17th Century Rationalism*. London – New York: Routledge.
- Parrochia, Daniel. 1984–1985. Physique pendulaire et modèles de l'ordre dans l'Éthique de Spinoza. *Cahiers Spinoza* 5: 71–92.
- Pikovsky, Arkady – Rosenblum, Michael – Kurths, J. Jürgen. 2001. *Synchronization: A universal concept in nonlinear sciences*. Cambridge: Cambridge university press. (*Synchronization*)
- Rivaud, Albert. 1924–1926. La physique de Spinoza. *Chronicon Spinozanum* 4: 24–57.
- Spinoza, B. 1677. *De Nagelate schriften van B.D.S.* Amsterdam: J. Riewertsz. (NS)
- Spinoza, Benedictus. 1985–2016. *The Collected Works*. Vol. I–II. Edited and translated by Edwin Curley. Princeton: Princeton University Press.
- van Bunge, Wiep. et al. (eds.) 2011. *The Continuum Companion to Spinoza*. New York: Continuum.
- Vermij, Rienk. 2013. Instruments and the Making of a Philosopher. Spinoza's Career in Optics. *Intellectual History Review* 23 (1): 65–81.
- Whitstone, Sebastian. 2012. Christian Huygens' Lost and Forgotten Pamphlet of his Pendulum Invention. *Annals of Science* 69 (1): 91–104.
- Wiesenfeld, Kurt – Borrero-Echeverry, Daniel. 2011. Huygens (and others) revisited. *Chaos* 21 (4): doi:10.1063/1.3665201 (Accessed May 5, 2017).
- Wootton, D. 2016. *The Invention of Science*. Milton Keynes: Penguin Books.
- Zourabichvili, François. 2002. *Spinoza – Une physique de la pensée*. Paris: Presses Universitaires de France.

# VIII

*Keith Green*

## Spinoza on Reflexive Affects and the Imitation of Affects

In *Ethics* (*E*) part 3, proposition 30, Spinoza gives an account of reflexive affects such as self-esteem (*gloria*), shame (*pudor*), penitence (*pœnetentia*), humility, and despondency (*abjectio*) in terms of imitating the affects of others. The affects of others that are imitated are forms of joy and sorrow with an idea of the cause as ‘external’ (to oneself – another person), which makes them forms of love or hatred (as Spinoza defines love and hatred in *E3p13* and in defs.6 and 7 at the end of *E3*). So we might suppose that these reflexive affects are also forms of love or hatred. Indeed, this comports with patterns of use in natural language; we readily speak of self-love or self-hatred. About these affects, Spinoza comments (in the *demonstration* of *E3p30*):<sup>1</sup>

*Dem:* He who imagines that he affects others with Joy or Sadness will thereby (by p27) be affected with Joy or Sadness. But since man (by 2p19 and p23) is conscious of himself through the affection by which he is determined to act, then he who has done something which he imagines affects others with Joy will be affected with Joy, together with a consciousness of himself as the cause (*cum conscientia sui, tanquam causâ*), or, he will regard himself with Joy, and the converse, *q.e.d.*

Even though Spinoza immediately goes on to say that the joy and sadness that constitute the imitated affects are “species of love and hatred”, he argues that because love and hatred

---

<sup>1</sup> All quoted English translations of Spinoza’s texts come from Curley, Edwin. (ed. and trans.) 1985. *The Collected Works of Spinoza*. 2 Vols. Princeton: Princeton University Press, unless otherwise noted. Citations for quoted texts will subsequently be bracketed in text, along with page and vol. location in Gebhardt, Karl. (ed.) 1925. *Spinoza Opera*. 4 vols. Heidelberg: Carl Winter. I will use the following standard, short reference to Spinoza’s writings; *Ethics*: E2p47s, Appendix = app, Axiom = ax, Corollary = c, Definition = def, Definition of emotion = def.em., Proof = pr, Explanation = exp, Lemma = lemma, Proposition = p, Postulate = post, Preface = pref. Scholium = s.

“have reference to external objects” (*Sed quia Amor, & Odium, ad externa objecta referentur...*), different names must be assigned to the reflexive affects that imitate the love and hatred of others for oneself. The reflexive affects that Spinoza is about to define are, as he describes them in the exposition of *E3def.em.24*, affects “together with an idea of an internal thing as the cause” (*quos idea rei internæ comitatur, tanquam causa*). And it seems that this sets them apart from love and hatred as affects: joy and sadness “accompanied by an external cause”. It seems to follow that the reflexive affects must somehow both “have reference to an external cause” yet are also constituted as a group by the fact that they affects “together with a consciousness of (one)self as the cause”. It is clear enough that ‘*internæ*’, as Spinoza uses the term in the *scholium*, should be taken to mean “with a consciousness of oneself as the cause”. So the ‘set up’ for definitions of reflexive affects, both in *E3p30* and in the exp. of *E3def24* gives rise to both unresolved textual and conceptual issues.

*Schol.*: Since Love (by *p13s*) is Joy, accompanied by an *external* cause (*concomitante etiam ideâ causæ externæ*), and Hate is Sadness, accompanied also by the idea of an *external* cause, this joy and sadness are species of Love and Hate. But because Love and Hatred are related to *external* objects (*ad objecta externa referentur*), we shall signify these affects by other names. Joy accompanied by an *internal* cause (*concomitante ideâ causæ internæ*), we shall call *love of esteem* (*gloriam*), and the *Sadness contrary to it*, *Shame* (*puorem*) – I mean *when the joy or sadness arise from the fact that the man believes he is Praised or Blamed*. Otherwise, I shall call *joy accompanied by the idea of an internal cause* (*concomitante ideâ causæ internæ*), *Self-esteem* (*acquiescentiam in se ipso*), and *the Sadness contrary to it* *Repentance* (*pœnitentia*) (Latin and italics inserted).

Spinoza’s remark about the names of the affects clearly bears some weight, since he takes pains at a later point in *E3*, in the explication of *def. 24*, to note for his reader that he is moving on, at that point, to define other affects which are “accompanied by the idea of an internal cause” (*Hinc ad alios transeo, quod idea rei internæ comitatur, tanquam causa*). Then follow definitions for self-esteem (*acquiescentia in se ipso*) (XXV), humility (*humilitas*) (XXVI), penitence (*pœnitentia*) (XXVII), pride (*superbia*) (XXVIII), despondency (*abjectio*) (XXIX), love of esteem (*gloria*) (XXX), and finally, shame (*puor*) (XXXI) – the very affects that were his subject matter in *E3p30s* and following.

The issue about the names of affects tracks Spinoza’s deployment of this distinction between affects ‘accompanied by’ an *idea* of the cause as something ‘external’, and those accompanied by the idea of the cause as something ‘internal’.<sup>2</sup> The trouble, both with the text and conceptually, seems to arise in his deployment of this distinction between ‘*internæ*’ and ‘*externæ*’, as well as in questions about whether ‘have reference to external objects’ (*ad objecta externa referentur* – *E3p30s*), and ‘accompanied by the idea of an external cause’

2 I note here that earlier readers of Spinoza have noted the conceptual issue raised by Spinoza’s language. See Boros 2007, 15. Boros points out the ambiguity of “accompanying” (trans. for *concomitante*) in In.18, *E3p30*, as translated by Curley.

(*concomitante etiam ideâ causæ externæ* – E3p30s, E3p13s, E3def.em.6 & 7) should be taken as equivalent in meaning, and opposite ‘together with consciousness of oneself as cause’ (*concomitante ideâ sui, tanquam causâ afficietur* – E3p30) and ‘accompanied by the idea of an internal cause’ (*quod idea rei internæ comitatur, tanquam causa*). So it behooves Spinoza’s readers to consider what work this ‘*internae/externae*’ distinction is doing more generally, and to question any intuition that this distinction is exclusive (in the sense that an affect can only actually have an ‘external’ or ‘internal’ cause).

E3p30 is among the many of Spinoza’s texts that are haunted by an unclear relationship between an affect *having* an ‘external’ or ‘internal’ cause, in fact (*de re*), and being ‘accompanied with an idea (*de dicto*) of’ an ‘external’ or ‘internal’ cause. Spinoza’s language in E3p30dem is that the love and hatred “have reference to” external objects. We might, after all, infer that according to Spinoza’s metaphysics, since a cause is represented in its effects, even when its effects are images in the mind of a subject (as we see in E2p17, for example), being caused by an ‘external’ cause implies that the idea in the mind of the subject must be an idea of an ‘external’ thing. But we should resist this inference, however intuitive it might seem. The causal account that Spinoza actually goes on to give of the reflexive affects after he turns to “the mind’s consideration of itself and its power of acting” in E3p53, and which he goes on to define after E3def.em. 24 (down to E3def.em.31) is an account in terms of external causes. We should also note that in the exposition of E3def.em.28 (pride), he also emphasizes that if we want to understand how it is possible for someone to think ill of oneself, we must attend *solely* to the opinions of *others*. Yet, as the explication of E3def.em.24 makes clear, subjects of the reflexive affects that Spinoza goes on to define take something about themselves or their actions as the cause of their pain or joy.

It follows that it not clear that when Spinoza claims, in E3p30s, that love and hatred “have reference to” external objects, he should be interpreted only as making a straightforward causal claim. And when (if these are the words he actually used), he goes on a few lines later in the *scholium* to describe love of esteem (*gloria*), shame (*pudor*), self-esteem (*acquiescentia in se ipso*), and penitence (*poenitentia*) as “accompanied by an idea of an internal cause”, we likely should not interpret him as making a claim about the actual cause, or its ‘location’ with respect to the subject. Indeed, Spinoza’s language suggests that we should treat his deployment of the ‘internal/external’ distinction in E3p30 as claims about subject’s *idea* (*de dicto*) of the cause, and not the actual (*de re*) cause, of the affect. What sets apart the reflexive affects that Spinoza defines after E3p30 (for which that proposition is a preface), and after def. 24, are that their cause is taken by subjects to be ‘internal’ though the proximate cause is actually and necessarily ‘external’ to them. Indirect support may be inferred from Spinoza’s claim (in E2p17s) that the nature of a subject’s body is a ‘cause’ – albeit an ‘inadequate cause’ of the ideas that the subject forms of ‘external’ bodies, as well as the causal effects of another body upon a subject’s own body.

If there is any doubt that Spinoza takes the actual (proximate) causes of the reflexive affects that he sets out to define after E3p53 and E3def.em.24 as ‘external’, consider the following: Spinoza accounts for subjects’ vulnerability to forming some such ideas in terms of the ‘weakness’ of their own minds or ‘spirits’ (in E3p54ff). Note that this outcome fol-

lows from the fact that they are forms of sorrow (*tristitia*), which are necessarily cases of subjects' 'cognizance' of their *conatus* being weakened by a cause that is necessarily 'external' (*E3p11*). Two critical outcomes follow. (i) Though subjects form an idea of the cause of reflexive affects as (something) *internal*, to the degree that these affects are 'the affection by which he is determined to act', the actions they cause will be 'seen' under the illusion of 'freedom' – where the proximate cause is actually *external* and *internal* (the condition of the subject's own body/mind), but it is only taken to be *internal*. (One cannot 'weaken' oneself according to Spinoza's *conatus* doctrine: *E3p4*, *E3p10*, among other places.) (ii) It actually matters that in the *scholium* of *E3p30*, Spinoza speaks of (and Curley aptly translates) the reflexive affects *accompanied by the idea of an internal cause*. His claim is not about the actual cause, but the subject's idea of the cause.

## The textual issue: Spinoza's use of 'internae' and 'externae' in *E3p30*<sup>3</sup>

It should first be noted, however, as Edwin Curley does in his English translation of the *Ethics*, that Spinoza's use of the terms "internal" and "external" to differentiate the 'objects' of love and hatred from the objects of reflexive affects in *E3p30s* also poses a vexing textual problem. And it remains a matter of disagreement among scholars – perhaps since it appeared difficult to construe which uses in *E3p30s* were most consistent with Spinoza's uses of the '*internae/externae*' distinction in other contexts. The *Opera Posthuma* (*OP*) and *Nagelata Schriften* (*NS*) do not agree, where "*internae*" and "*externae*" are employed in *E3p30s*. And editors and translators have disagreed in their reconstruction of what they took Spinoza's choice of words to be. In the former (*OP*), an idea of an "external" cause is imputed to both 'love of esteem' and 'self-esteem' in *E3p30s*. In the later (*NS*), which Curley follows, an "internal" cause is imputed to these affects. Gebhardt, it should be noted, follows the *OP* in the first case (G II/163/24) and *NS* in the second (GII/163/27, 28).

Curley justifies following the *NS*, however, by alluding to the fact that both love and hatred (in both *E3p13*, and in the *definitions of affects* at the end of part 3, explication of def. 24) are defined, "in their simplest forms" as affects with an idea of the cause as something *external*. And he adds that in the exposition of def. 24, at the end of part 3, Spinoza gathers many of the same reflexive affects that are countenanced in *E3p30s* and for which he accounts after *E3p53*, under the description that they are affects *accompanied by the idea*

3 This discussion of textual issues in *Ethics* 3p30 is entirely indebted to the generous and extensive help of Piet Steenbakkers, by way of correspondence. For a more complete discussion of these issues, extended remarks from correspondence are appended, with his permission, at the end. Steenbakkers refers to the following works in his appended remark from correspondence: Akkerman 1977; Camerer 1877; Spruit–Totaro 2011; van Vloten – Land 1914; *Opera Posthuma* 1677.

of an internal cause. He offers what I would argue is a correct interpretation, overall (Curley I, p. 511, ftnt. 27).

It seems clear that the *OP* compositor, influenced by the appearance of ‘external’ three times earlier in this scholium, misread the manuscript, and failed to see that Spinoza intended to draw a contrast between love and hatred in their simplest forms, and the four more complex forms enumerated between lines 24–29. What lends plausibility to the *OP* reading is the fact that the joy and sadness whose origin is described in P30 are explicitly said to be species of love and hate (at lines 20–22), from which it follows that they must involve an idea of an ‘external’ cause.

The argument that he proceeds to make, however, is not based upon purely textual considerations. It is, in part, a conceptual argument that the subject of reflexive affects can have an internal idea of the cause, even though the idea of sadness of another that one takes oneself to have caused, and which is ‘external’ to oneself, is ‘indispensable’.

But this does not exclude the possibility of the idea of an internal cause (as Akkerman seems to assume, 189). I take it that this is Spinoza’s point: the four more complex forms of love and hatred defined in lines 24–29 arise from the fact that I believe myself (perhaps mistakenly) to have caused joy or sadness to another, *thereby* causing myself joy or sadness. In either case, I am the indirect cause of joy or sadness to myself. But the idea of the other as a subject of the joy (or sadness) I (take myself to have) caused is indispensable to this particular form of joy (or sadness). So my affect does involve the idea of something external as its partial and immediate cause, viz. the other’s affect.

Ongoing debates about the original or final language of the text continue to reflect certain conceptual assumptions by later generations of textual editors, aiming to figure out which version of the text was actually *consistent*. Curley considers two factors. He notes the language of the exposition of *E3def.em.24*, where Spinoza informs his reader that he is proceeding from that point forward to define affects where the subject has an idea of the cause as something ‘internal’. He also attempts a brief argument that a subject of the affects in question can also have an idea of the cause that is both ‘internal’ and ‘external’.

Resolving the textual issue also has implications for resolving conceptual questions in a way that avoids imputing inconsistency to Spinoza. The account that Spinoza actually gives of painful reflexive affects, specifically after *E3p53*, is in terms of ‘external’ ‘causes–others’ painful ideas of oneself, which subjects passively affirm, and thereby imitate others’ affects towards them. And Spinoza must account for painful affects in terms of external causes, if he holds to the implication of the conatus argument that one cannot weaken or destroy, and thus cause pain or sorrow to oneself. Yet these are affects where oneself is the object. So the conceptual issue that motivates the different positions that interpreters have

taken about the right language of the text has been that the cause could not be ‘external’ in fact, if subjects of reflexive affects have an idea of the cause as something ‘internal’.

## The conceptual issue

The conceptual problem raised in *E3p30* is that if one loves or hates something only insofar as one’s joy or sorrow is accompanied by an idea of a cause that one takes to be ‘external’ (something not oneself, and so, not taken to be an attribute of oneself) [by *E3p13* and *E3def.em.6 & 7*], it appears to follow that one can only literally be loved or hated by another. Yet, as we saw, Spinoza flatly claims [in the *scholium* – lines 20–22] that reflexive forms of joy and sadness are “species of love and hatred”. What, then, are we to make of this claim?

Spinoza has claimed, in *E2ax3*: “There are no modes of thinking, such as love, desire, or whatever is designated by the word ‘affects’ of the mind, unless there is, in the same individual the idea of the thing loved, desired, etc.” And in *E2p27dem*, where he introduces the notion of imitation of affects, he claims: “The images of things are affections of the human body, whose ideas represent *external* bodies as present to us (by *E2p17s*), i.e. (by *E2p16*), whose ideas involve the nature of our Body and at the same time the present nature of the external body” (*italics inserted*). So in *E3p27*, Spinoza claims: “*If we imagine a thing like us, toward which we have had no affect, to be affected with some affect, we are thereby affected with a like affect.*” Michael Della Rocca gives the following schematization of imitation of affects (from a close reading of *E3p27* and *3p27d*):<sup>4</sup>

x imitates y with respect to affect A iff y is similar to x, and x imagines (has an image of) y as having affect A.

The first corollary (of *E3p27*) provides the critical link between the imitation hypothesis and the reflexive affects:

If we imagine that someone toward whom we have had no affect affects a thing like us with joy, we shall be affected with Love toward him. On the other hand, if we imagine him to affect it with Sadness, we shall be affected with Hate toward him.

If one imagines oneself affecting another, who is ‘similar’ in the relevant sense to oneself, with joy or sorrow, then is one not, by implication, ‘affected with’ love or hatred toward oneself? [If, as Spinoza insists, x’s having an image of ‘y having an image of x’ amounts to

---

<sup>4</sup> Della Rocca 1996, 249. Note that Della Rocca’s schematization is not quoted precisely as he schematizes it. Since, as Della Rocca himself notes, Spinoza specifies that “affect imitation involves x’s *imagining* y to have a certain affect”, I have substituted “imagines” for “regards” in Della Rocca’s schematization.

x having an image of x – and then “x’s being affected with the same affect” toward x with which y is affected toward x] – *then does one not have to be able to have an image of one’s own body as something ‘external’?* This, however, at least looks incompatible with his definitions of love and hatred in *E3p13s*, and in *defs. 6 and 7* at the end of part 3 as affects *accompanied by an idea of the cause as external*.

I take it as clear and non-controversial that Spinoza accounts for reflexive affects like love of esteem, shame, and penitence in terms of imitating others’ love and hatred [both in the propositions following *E3p53* and especially in the exposition of *def. 28* (of pride), at the end of part 3]. We should also note that in the exposition of *E3def.em28*, Spinoza insists that if we want to understand how it is possible for a person to think too ill of oneself, we must consider specifically how others view him. Why, we might ask, would reflexive affects require anything more, as a necessary condition, than simply being able to have or form any idea of oneself that others can have – or more specifically, form ‘third-person’ beliefs about oneself? ‘Third-person’ beliefs about oneself would be beliefs where the same norms of justification apply to one’s own beliefs about oneself that would apply to others’ beliefs about oneself. And if one’s belief about oneself can be expressed with the same proposition (*x is a.*) that another person’s belief about one could be expressed: [(*x* believes ‘*x is a*’) and (*y* believes ‘*x is a*’)], then it seems obviously to follow that the truth conditions that apply to *x*’s belief would also apply to *y*’s belief.

This, however, is precisely the issue. It is not so obviously the case, given Spinoza’s distinction between ‘internal’ and ‘external’ causes and ideas of causes, that just because one’s beliefs about oneself answer to the same epistemic standards as other’s beliefs about one, and that either hold justified or unjustified beliefs about one, that one can ‘see oneself just as other see one’ or ‘form an image of oneself from the same perspective from which others form an idea of oneself’. Spinoza’s language in *E3p30s* seems to imply that one somehow cannot have a ‘third-person’ image or ideas of oneself, and so one cannot love or hatred oneself.

## “Internal” and “external”

To see why the ‘internal/external’ distinction in *E3p30* is an issue, it is important to consider what stock Spinoza has already invested in it (as it were) up to this point, and elsewhere, in the *Ethics*. The down payment is made in *E2p17* [*If the human Body is affected with a mode that involves the nature of an external body, the human mind will regard the same external body as actually existing, or as present to it, until the Body is affected by an affect that excludes the existence or presence of that body.*]. In the *scholium* of *E2p17*, Spinoza makes it clear that he regards himself as having shown, in the two preceding demonstrations of this proposition, how “it can happen that we regard things as present that do not exist”. The first demonstration makes the point that insofar as a human body is affected by another

body, the human mind will ‘have the idea of a mode that actually exists’, the idea that ‘posits the existence or presence’ of the nature of the external body’.

I would argue that in making this claim, Spinoza is implicitly contrasting a *de re* idea – ‘an idea of a mode that actually exists’ – with a *de dicto* idea – what that idea is *taken by the subject* of the idea to be an idea of or about (Steinberg 2013, 283–407). So space is already open for the possibility of nonveridical ideas, or for failing to form an idea (*de dicto*) of necessarily external causes that may nonetheless be represented in their effects upon one’s image- and belief-forming, and ‘dispose’ (*disponunt*) one’s imagination, as in the case of evidently uncompelled suicide (as Spinoza adduces its causes in *E4p20s*). Here, he redeploys an example he introduced in *E2p17s*: the contrast between the idea of Peter that ‘constitutes the essence of Peter’s mind’ and the idea of Peter which is ‘in’ the mind of Paul (another man). Peter’s idea of Peter is said to “directly explain the essence of Peter’s body” and it ‘involves existence’ only so long as Peter actually exists. Paul’s idea of Peter, on the other hand, is said to “indicate the condition of Paul’s body more than Peter’s nature”; and while the condition of Paul’s body lasts (being causally affected by the presence of Peter’s body), Paul’s mind will still regard Peter as present. But Paul (unlike Peter) can be caused to ‘have an idea of Peter as present’ even though Peter’s body is no longer there, by the laws of association that determine which images and ideas will be ‘present to mind’ for any agent/subject at any given point in time. Even if Peter can, say, reach across his lap with his right hand and caress the back of his other hand, presumably encountering it as an object that can produce effects or perceptions, and even if light refracting from one’s body can be reflected back upon one’s retina by a mirror (causally producing an image, or perception, of one’s own body in one’s own body), it evidently matters more, on Spinoza’s view, that Peter can only be conscious of his body so long as his body actually exists. Paul’s body may, however, outlast Peter’s. And images produced by Peter’s body upon Paul’s, and so yielding Paul’s ideas of Peter, may outlast Peter’s body.

In *E2p17s*, the ‘*internae*’/‘*externae*’ distinction distinguishes images ‘in’ or ‘present to’ one’s own mind as opposed to images, even of one’s own body, in another mind; but implicitly it is (already) a distinction between ‘species’ of affects. The distinction’s significance and implications become clearer in a vexing passage at the end of the *scholium* of *E3p11*.<sup>5</sup> The *scholium* of *E3p11* begins with Spinoza emphasizing that the mind can “undergo great changes”, passing to a lesser or greater perfection, and that this fact “explains the affects of Joy and Sadness”. It is here that he introduces definitions of Joy (*laetitia*): “*passion by which the mind passes to greater perfection*” and Sadness (*tristitia*) “*passion by which (the Mind) passes to lesser perfection*”. He *then* states that he would like to “explain” *E3p10* [*An idea that*

---

5 *E3p10dem* must also be seen in light of *E3p7* [*The mind’s conatus or power is the very essence of mind.*] And *E3p54* [*The mind endeavors to think only of the things that affirm its power of activity.*] and: [*But the mind (from its own nature, or in and of itself) affirms only what the mind is and can do, and not what the mind is not and cannot do.*]. So as Garrett translates *E3p55dem*: “it is of the nature of the mind to think only of those things that affirm its power of activity” *E3p55*; Garrett 2002, 127. Garrett’s translation makes more clearly the critical distinction between something’s ‘nature’ as an individual mode vs. its ‘conatus’, or ‘striving’ *quantum in se est*. Garrett 2002, 141.

*excludes the existence of our Body, cannot be in our mind but is contrary to it.*] – and this, he claims, amounts to explaining how “an idea can be contrary to an idea”, specifically (by implication) contrary to the idea that constitutes one’s mind.

Referring back to *E2p17s*, Spinoza reveals what he takes himself to have proven there: Three things always depend upon (or “involve”) the *present* existence of one’s body: (i) “the idea which constitutes the essence of the Mind”, (ii) “the present existence of the mind” and (iii) “the power of the mind by which it imagines things and recollects them”. The implication is that when the mind ‘affirms’ that it will not exist at a future time, did not exist at a past time, or existed in the past at a time which it cannot remember, or that it cannot do something, it affirms an idea that can only exist (as an image) in another mind. So the mind is necessarily caused to affirm such ideas (of itself) by causes ‘external’ to itself. And Spinoza makes it clear enough later in *E3p11s* that in the mind’s affirming such ideas, it is doing so passively or ‘with indifference’, and so also being “weakened”.<sup>6</sup>

But the cause of the mind’s ceasing to affirm the existence of the body cannot be the mind itself (by *E3p4*), nor also that the body ceases to exist. For (by *E2p6*) the cause of the mind’s affirming the body’s existence is not that the body has begun to exist. So by the same reasoning, it does not cease to affirm the body’s existence because the body ceases to exist, but (by *E2p8*) this [*sc.* ceasing to affirm the body’s existence] arises from another idea which excludes the present existence of our body, and consequently of the mind, and which is thus contrary to the idea that constitutes our mind’s existence (Curley, Vol. I, p. 501).

‘Ceasing to affirm’ amounts to ‘ceasing to believe’. And it follows, on Spinoza’s reasoning, that since one cannot literally ‘have an image of’ one’s body not existing (i.e. one literally cannot perceive or experience oneself not existing), affirming an idea of oneself not existing (i.e. unthoughtfully believing) in the past, or having an idea of oneself not existing at a future time, can only have a that is of an (following cause) ‘external’ to one’s body or mind, and its activity. Spinoza seems to infer, in other words, that only another mind (from the perspective of a body ‘external’ to one’s own) can ‘have an image of’ one’s body not existing. Only in the idea of that other body’s affections (i.e. ‘in’ its mind) can one’s not existing literally be experienced, and so is necessarily its ‘idea’. But Spinoza clearly recognizes, in *E3p30* (among other places) that one can passively affirm such an idea about oneself, and that is why it is possible, for example, to fear, or desire, or grieve the idea of one’s not

---

<sup>6</sup> In Letter 21, to Blyenbergh, Spinoza claims, “(Our) freedom is placed neither in contingency nor in a certain indifference, but in a manner of affirming and denying, so that the less indifferently we affirm or deny a thing, the more free we are” (Curley I, p. 378, G IV/130). At work here is the distinction he has made between “passive” and “active” affirmation in *E2p18*, among other places. Even if one cannot but “imitate” others’ love or hatred of oneself, one can decline to affirm hateful ideas of oneself, or resist wrongfully imputing such ideas to others, only from ‘*strength of mind*’. And as the comment from Letter 21 makes clear, only in being able to do this is one able to act from the power of one’s own mind, and so “from one’s own power” as an agent.

existing in the future. And this is how it is possible to have (i.e. passively or unthoughtfully affirm) painful ideas of oneself, that *qua* painful, cannot be the effect of one's own natural striving, or the activity of one's own mind. For Spinoza, in other words, one's having such an idea of oneself has to have an 'external' cause – like Paul's image or idea of Peter's body in the example of *E2p17s*.

We must point out, however, that just as Curley claims, there is, nonetheless, an 'internal' cause of one's imitating another's affects toward oneself. And that is whatever it is about the current strength or weakness of one's mind or 'nature', and the various 'aptitudes' (*aptum*) of one's body, that explains why one is liable to passively affirm another's ideas of oneself. In *E5p20s*, Spinoza implies that the power of contrary emotions or affects over one's mind is 'defined by knowledge alone' where 'lack' of power of mind (passion) is judged by 'lack of knowledge'.

It appears to follow, on Spinoza's terms, that when one's mind 'affirms' other 'third-person' beliefs about itself – about one's body – such as the idea that it will not exist at a future time, or that it cannot do something, it is being 'weakened' by a cause necessarily 'external' to it; so its cognizance of this weakening is 'sorrow', by its very nature. And this must encompass all the reflexive affects that are forms of sorrow, and thus, of hatred, as well as affects that are 'joyful' just to the degree that they 'offset' other painful affects. And this is just the point that we hear Spinoza making at the outset of the *scholium* of *E3p11*.

To put the matter in somewhat more contemporary philosophical terms, we can 'triangulate' the images or ideas ourselves that can only exist as an idea of something 'external' in the mind of another, in virtue of which others love or hate us; and in doing so (but only in doing so) come to love or hate ourselves. But we can never just 'see ourselves as others see us'. We cannot, 'redouble' or replicate (from a 'third person' perspective that is genuinely 'ours' and not someone else's) another's image or idea of ourselves, as if we are seeing our own image in a mirror, or watching ourselves upon a stage as one member of the audience amongst others. We are, thus, always dependent upon images and ideas that can only form in the minds and experiences of other – or 'in' or from the perspective of other bodies – to have a sufficiently complex and unified idea of ourselves in order to function as an agent.

Two points must be made here. First, I am not defending Spinoza's rather counterintuitive argument that one cannot form at least a wide range of 'third-person' beliefs 'from one's own 'experience' or the perspective of one's own body without triangulating someone else's beliefs; nor am I taking a position about just what reflexive ideas fall within this range. Second, we must see that this is not a feature of affect formation that Spinoza thinks we either can, or should aim to, "outgrow" or "get beyond". That is to say, it cannot be regarded an initial stage of self-consciousness that is somehow "transcended" in the development of self-knowledge, and as a condition for the emergence of a 'social' self or agent. It is simply (but more radically) an account of the formation of the mind (one's complex idea of one's own body, constituted quite literally by all other ideas of one's affections) and its contents.

Though I do not think the following comment is a complete argument, in itself, for the second point made here, We can see why Spinoza could never embrace a developmental account of reflexive affects, where it is assumed that one can, and must aim to 'transcend'

dependence upon others' images or ideas of oneself, or their role in moral or social agency. We perhaps glimpse this most clearly if we reread closely some of Spinoza's comments about the example of the Spanish poet in *E4p39s* – the example which Spinoza deploys to argue that a body can “die without becoming a corpse”. The Spanish poet is “incredible” because he got sick, and upon recovering, could not remember having written any of the poems or tragedies that he had written, though he had not reverted to being an infant only because he could still speak his native language. It actually remains unclear whether Spinoza thinks that the poet has actually ‘died without becoming a corpse’ or whether it is retaining an ability to speak his native language that ‘saves’ him from dying.

In the business of trying to persuade his reader that the case of the Spanish poet is not so exceptional as to be uninformative, Spinoza points out that people in general are actually not so different from the Spanish poet in the relevant way. We are only able to have a more (rather than less) adequate, ‘unified’ idea of ourselves as having been the infants we once were, not because we can remember having been one. We can have this idea only because we affirm others' ideas of our having been one. Though we cannot remember having been an infant, we see other infants, and then, reasoning “by analogy” that our body developed and grew as those of the infants we experience, we *actively* affirm *others'* ideas of our having once been an infant. We are never able to see or ‘experience’ ourselves as an infant, but since we can speak a language, others can communicate their idea of having been an infant in the past, and, drawing upon our experience of other infants, we can affirm others ideas of our having been one. So it is not just in not remembering having written the works that he wrote, that the Spanish poet “dies without becoming a corpse” but in not being able to speak his ‘native language’, and, through the activity of his mind, affirm *any* idea of himself having written them, that he dies. One's mind, after all, just is one's idea of one's body. If the poet can have that idea (i.e. if he can actively affirm the more adequate idea that he wrote the works that he actually wrote, then to that extent, he continues to be able to have an idea of his body (i.e. he continues, quite literally, to ‘have a mind’). But he is dependent upon the ideas of others in just the way we all are dependent upon others ideas of us, when we believe that we were once the infant we cannot remember having been.

What matters (what constitutes his individual *nature* continuing to “live”) is not only “internal” ideas that, as Spinoza has it in *E2p17s* “explicate” one's body, but the capacity to actively affirm “external” ideas of oneself (ideas which can only form of one as an ‘object’ in the experience of another). Others' images of the Spanish poet, and others' images of ourselves, make it possible for him, and for us, to conceive of ourselves as the individual that we are over time. And this is a critical necessary condition that we do not “outgrow”. There is never a point in time when we, or the poet, do not depend upon others ideas of us (as something ‘external’ to their bodies) in order to have a more adequate idea of ourselves – i.e. have a mind, and thus a nature, as an individual. And it is a necessary condition of our being accountable agents and emotional participants in a wide range of broadly narrative social relationships over time.

If one *cannot but* affirm others' ideas of oneself – ‘external’ ideas that, inferring from the Peter/Paul example of *E2p17*, can only ‘image’ in the minds of others – then one can be either

“strengthened” or “weakened” by affirming those ideas; and one’s affirming them can be either passive or active. But the fact is that, at least implicitly, they remain a necessary ‘part’ of our complex ideas of our bodies, the ‘idea’ of our body that Spinoza says ‘just is’ our mind.

## The upshot of the conceptual issue raised in *E3p30*

We can see that Spinoza’s notion of the imitation of affects as the causal source of reflexive affects, in *E3p30*, clearly anticipates critical aspects of notions of imitation (of affects) and sympathy that circulate through the works of a later generation of early modern thinkers. As early as Shaftsbury and Mandeville, we already see such ideas being incorporated into developmental narratives of the emergence of sociality and agency, where imitation is an initial (though perhaps universal and necessary) ‘stage’. Or in the case of Rousseau, by contrast, where ‘socialisation’ corrupts one’s unalloyed, unselfconscious (and, therefore, authentic) responses, it renders one vulnerable to being moved by vanity. For Hume and Adam Smith, it also accounts ‘developmentally’ for the emergence of a capacity for moral sentiments, and for the motivational salience of ‘artificial’ virtues, such as justice, and, thus, for something like moral autonomy.

In these developmental narratives or accounts, however, “full” autonomy – an ability to morally or socially self-regulate without the ongoing ‘input’ of others’ actual praise or blame – only emerges when one acquires the capacity to ‘transcend’ dependence upon others responses. Only then can one respond ‘impartially’ to a ‘view’ of oneself, and determine whether one’s actions and sentiments are ‘praiseworthy’ or ‘blameworthy’ as opposed to merely eliciting the praise or blame of a specific ‘Other’. And one presumably acquires this capacity by being able to ‘see oneself just as others see one’ – from a ‘third-person point of view’ – as if one is seeing one’s own image as a separate ‘thing’ in a mirror, or as one member of an audience among others, viewing oneself upon a stage. And so, fully autonomous agency emerges only when subjects ‘transcend’ *internalizing* or *redoubling* others’ ‘views’ and thus, ‘imitating’ their emotional responses towards oneself.

Perhaps we see this version of imitation of affects as ‘internalization’ on the way to autonomy, most starkly in a famous example of Adam Smith’s, from *The Theory of the Moral Sentiments* (Smith 1982, TMS forthwith):

To a man who from birth is a stranger to society, the objects of his passions, the external bodies which either pleased or hurt him, would occupy his whole attention. The passions themselves, the desires or aversions, the joys or sorrows, which those objects excited, though of all things the most immediately present to him, could scarce ever be the objects of his thoughts. The idea of them could never interest him so much as to call upon his attentive consideration. The consideration

of his joy could excite in him no new joy, nor that of his sorrow any new sorrow, though the consideration of the causes of those passions might often excite both. Bring him into society, and all his own passions immediately become the causes of new passions. He will observe that mankind approve of some of them, and are disgusted by others. He will be elevated in the one case, and cast down in the other; his desires and aversions, his joys and sorrows, will now often become the causes of new desires and new aversions, new joys and new sorrows; they will now, therefore, interest him deeply, and often call upon his most attentive consideration (Adam Smith, *TMS*, p. 111).

Here, the developmental trajectory is clear. One moves from only being able to respond to ‘external bodies’ that are ‘immediately present’ to one, to ‘coming into society’, where one becomes, as it were, acquainted with oneself, as a body among others, and subject to a whole ‘new’ range of reflexive affects, which simply are affects with which others respond to one from the same ‘perspective’, but without a dependence upon (or literal imitation) of what one imagines others’ affective responses are. That these affects are constitutive of agency and socially, and something like moral autonomy, becomes clearer in passages that almost immediately follows. In the first edition of *TMS*, Smith continues:

To judge of ourselves what we judge of others, to approve or condemn in ourselves what we approve or condemn in others, is the greatest exertion of candour and impartiality. In order to do this, we must look at ourselves with the same eyes with which we look at others; we must imagine ourselves not the actors, but the spectators of our own action and conduct, and consider how these would affect us when viewed from his new station, in which their excellencies and imperfections alone can be discovered. We must enter, in short, into what, if the whole circumstances of our conduct were known, we imagine would be the sentiments of others, before we can either applaud or condemn it.

In the 2<sup>nd</sup>, and subsequent editions of *TMS*, Smith reformulates these claims: (p. 113) in ways that accentuate both the possibility and the moral psychological significance of taking up a ‘third-person’ perspective with regard to one’s conduct and affects. One can be, to use Spinoza’s language, simultaneously ‘internal’ and ‘external’ to myself. And it is not insignificant that the third-person perspective from which one can, with sufficient development, view and self-regulate one’s own motives and character, is specifically figured in juridical language.

When I endeavor to examine my own conduct, when I endeavor to pass sentence upon it, and either to approve or condemn it, it is evident that in all such cases, I divide myself, as it were, into two persons; and that I, the examiner and judge, represent a different character from that other I, the person whose conduct is examined into and judged of. The first is the spectator, whose sentiments with regard

to my own conduct I endeavor to enter into, by placing myself in his situation, and by considering how it would appear to me, when seen from that particular point of view.

What is singular about Smith's reformulation of his account of 'taking up' the spectator's perspective, is the implication that one 'gets beyond' merely imitating, or 'internalizing' others' ideas of oneself, by becoming able to take up a 'third person' – the "spectator's" – perspective. One can 'view' one's actions, intentions, and 'first-order' affective responses from the same perspective as others, and in doing so, describe them in terms subject to the same norms or conditions of warrant that govern others' beliefs and judgments (others' ideas) about one's actions. In the former 'condition, where one 'internalizes' others' views of oneself, one imitates their passional responses toward oneself. Where one can 'view oneself' as 'another', one's affective responses are not imitations of others, but authentically 'one's own.' And from this 'third person' perspective, it is possible to function autonomously as one's own 'judge', without one's sorrow, self-hatred, or approval being caused by causes that would figure as 'external' for Spinoza. We might call such a view an '*internalization*' or '*dédoublement*' view (of the subjective emergence of 'autonomous' agency).<sup>7</sup>

What do we learn, by way of contrasts, about the conceptual issue raised by Spinoza's account of reflexive affects in terms of imitating others' affects in *E3p30*? Spinoza's conception of the 'internal/external' distinction implies that an account of reflexive affects in terms of imitation of others' affects can never be assimilated to a '*dédoublement*' account of the emergence of agency or moral autonomy. There are two reasons. (i) It is literally impossible for an individual to experience, or be conscious of, their own body 'as an object' or as 'external' in just the way that others experience it, as a *different*, or 'external' body. Yet, (ii) any 'idea of oneself' adequate to sustain accountable agency in any meaningful sense, cannot even attain minimal 'unity' of consciousness without affirming others' ideas of oneself (as an 'object' or different body) – ideas that can only 'image' in the experience, and from the perspective of another body. One can affirm such ideas either actively or passively (thoughtfully or unthoughtfully) through the activity of one's mind; but one can never 'replicate' them. One can never, therefore, extirpate or somehow 'rise above' others' ideas of oneself as a different body altogether. They are (after all) quite literally 'parts' of the complex idea of one's body that, on Spinoza's view, *is* one's mind. We noted above that it is for just this reason that Spinoza's imitation account of reflexive affects better exemplifies what Gerald Postema "*triangulation*".<sup>8</sup>

It is critically important to appreciate the far-reaching implications of Spinoza's nuanced deployment of the '*internaexternae*' distinction in *E2p17s*, *E3p11s*, and in his definitions

---

7 See Marshal 1986 for a study of *dédoublement* as a theme and idea across literary and philosophical texts.

8 Postema 2005, 265–269. Triangulation implies that reflexive ideas that are necessarily only from the perspective of another body, and that imitate the affects imputed to that other individual always necessarily continue to involve three causal factors: (i) X's idea of Y, (ii) Y's idea (and affect toward) X, and (iii) X's affirming Y's idea of X: So X's idea (and affect) toward X.

of love and hatred in *E3p13* and *E3def.em.6* & 7. Its deployment in *E3p30*, entails that even though ideas of one's body can only 'represent' it from as 'internal' (as oneself) or 'external' (different from oneself), one's mind (one's complex idea of one's body) necessarily includes (as 'parts' or constitutive ideas), if it is to be a more (rather than less) adequate idea, ideas from both perspectives. But since ideas of one's body as 'external' are necessarily an image of the body as a different body, and one cannot literally 'see' or experience one's body as a different body, one's mind (one's complex idea of one's body) is at least tacitly or passively affirming, and thus dependent upon, others' (actually different bodies') ideas of one's body. Painful reflexive affects can never involve 'duplicating' others' experience of one's body as a different body (as external), but only passively affirming necessarily different (or other, or 'external') bodies' ideas of one's own. And one is necessarily continuing to passively affirm such ideas so long as one is subject to painful (or hateful) reflexive affects.

What the conceptual issue raised by Spinoza's deployment of the 'internae/externae' distinction in *E3p30* reveals – I would argue – is that if his construal of this distinction is plausible, it throws into relief the ideological slight of hand at work in deployments of the pervasive *dédoublement* model (exemplified by Smith) to account for 'autonomous' social or moral agency. As we saw above, with Smith, these accounts hold that we 'come into our own' as autonomous agents via *dedoublement* – being able, by ourselves, to see ourselves – our own body – as an object, just as others see it – as if we are part of an audience constituted by everyone else who observes us. It perpetrates (essentially begging the question) the notion that self-hateful reflexive ideas and affects can become 'fully our own', and are not a diminishing or 'causing' of one's agency 'from the outside'. Such 'fictions' perpetrate the impression of a notion of individual autonomy as a certain sort of independence or authenticity by erasing or denying the ongoing passive causal dependence that yields us phenomenal agency.

Finally, it helps to see what is right, but also incomplete, about Curley's conceptual argument for resolving the textual questions raised about Spinoza's deployment of the 'internal/external' distinction in *E3p30s*. Curley seeks to resolve the tension inherent in the question about Spinoza's deployment of the 'internae/externae' distinction to describe the location of the cause of painful reflexive affects by introducing a distinction between 'direct' and 'indirect' causes. He takes Spinoza's point to be that the more 'complex' forms of reflexive love and hatred defined in lines 24–29 "arise from the fact that I believe myself (perhaps mistakenly) to have caused joy or sadness to another *thereby* causing myself joy or sadness" (quoted above). I am, then, the indirect cause of joy or sadness to myself. The other's affect (that I take him to have) is an 'indispensible' "partial and immediate" cause. Curley clearly sees that the reflexive affects necessarily involve both external and internal causes. It is worth repeating the point that if one's essential natural striving (or *conatus*) can never register in self-consciousness as painful, or a painful idea of oneself, then painful or self-hateful reflexive affects such as shame, humility, penitent regret, or despondency can only have an 'external' cause even though the subject's idea of the cause is only of something 'internal'. Curley's resolution, however, foregrounds triangulation; so long as painful reflexive affects are 'present to mind' and have motivational salience for me, I am necessarily believing, or

unthoughtfully assuming, that another person has a particular belief about me. I am never, and, indeed, can never ‘step outside of myself’ or ‘split myself into two’ and form the kinds of beliefs about myself that only others can form, and that are necessary constituents of the reflexive affects that Spinoza accounts for in *E3p30*.

## Appendix: Piet Steenbakkers, comment from correspondence on the text of *Ethics 3p30*

This is the text of *E3p30s* in V (the Vatican MS), fol. 65v–66r:

Cum amor (per sch. pr. 13. h.) sit laetitiâ concomitante idea causae externae, et odium tristitia concomitante etiam ideâ causae externae, erit ergo haec laetitia et tristitia amoris et odij species. Sed quia amor et odium ad objecta externa referuntur, ideo hos affectus alijs nominibus significabimus; nempe laetitiam concomitante ideâ | causae’ *externae*’ Gloriam, et tristitiam huic contrariam Pudorem appellabimus. Intelligo quando laetitia vel tristitia ex eo oritur, quod homo se laudari vel vituperari credit, alias laetitiam concomitante ideâ causae; *externae*’ acquiescentiam in se ipso; tristitiam verò eidem contrariam Paenitentiam vocabo. Deinde quia (per coroll. pr. 17. p. 2.) fieri potest, ut laetitia, quâ aliquis reliquos afficere imaginatur, imaginaria tantum sit, et (per pr. 25. h.) unusquisque de se id omne conatur imaginari, quod laetitiâ ipsum afficere | imaginatur; facile ergo fieri potest, ut gloriosus superbus sit, et se omnibus gratum esse imaginetur, quando omnibus molestus est.

(I quote from the MS rather than from the edition of it by Spruit–Totaro, p. 192. Abbreviations in the text – but not those in the the cross-references – have been tacitly expanded.)

As you can see, there are some minor differences between V and OP, but they agree when it comes to the two disputed occurrences of ‘*externae*’ (italicized in the quotation). I agree with Camerer, Akkerman and Curley that we should read ‘*internae*’ in both cases (see below for the reasons why), but the agreement of V and OP shows that Spinoza at one stage did in fact write ‘*externae*’. It would be far too coincidental if both the compositor of the OP and the scribe of V (Pieter van Gent) had misread both occurrences. If we accept that Spinoza at one time did write ‘*externae*’, there are three possibilities:

- (1) Spinoza wrote and indeed intended ‘*externae*’ here, so OP, V and Vloten-Land are correct and NS, Camerer, Akkerman and Curley wrong;
- (2) Spinoza did write ‘*externae*’, but changed his mind and corrected both occurrences to ‘*internae*’ in a copy of the text that the Dutch translator (Glazemaker) had access to, but the correction did not reach the compositor of the OP nor the

scribe of V; in that case NS, Camerer, Akkerman and Curley are right and OP, V and Vloten-L are wrong;

(3) Spinoza did write ‘*externae*’, changed his mind but failed to correct the two erroneous occurrences of ‘*externae*’ in 3p30s; the mistake was noticed by Glazemakers and/or the editors of NS, and they corrected it.

That we should read ‘*internae*’ here is, I think, clear if we combine 3ad24exp (“I pass now to the [other affects of Joy and Sadness], which are accompanied by the idea of an ‘*internal*’ thing as cause”) with 3ad30–31, and compare that with 3p30s. In the definitions of the affects, Spinoza explicitly distinguishes between joy-and-sadness-cum-idea-of-external-cause (love and hate) and joy-and-sadness-cum-idea-of-internal-cause (among them love of esteem and shame). That disposes of possibility #1. (This is also Camerer’s argument; I attach the page where explains this.) Possibility #3 is out, too: it is inconceivable that Glazemaker or an editor would have carried out a correction without consulting the OP editors. That leaves only #2: Spinoza changed his mind. V was copied by Van Gent between the end of 1674 and May 1675, so perhaps Spinoza made the correction after that. Apparently, the compositor of the OP here worked from a copy in which the correction had either not been implemented, or had not been indicated clearly enough.

If my reconstruction is correct, that would confirm my suspicion that the definitions of the affects were added relatively late. This is what I think happened (but I cannot prove it): Spinoza had initially noticed structural affinities between love/hate on the one hand and love of esteem/shame on the other; that accounts for the wording of the initial sentence of *E3p30s*. The second sentence, however, indicates that he also was aware of the specificity of love/hate: their external objects. It was only in elaborating the systematic distinction external/internal in *E3def.em25–31* that Spinoza realized the formulation 3p30s was inaccurate, and he corrected it, but not (or not clearly enough) in all copies. Curley’s suggestion that there is some externality involved (because my affect is mediated by someone else’s affect) may account for Spinoza’s initial hunch that the two types of affects were structurally similar, but it cannot be an argument to retain the readings of OP and V.

## Works Cited

- Akkerman, Fokke. 1977. *Spinoza’s tekoor aan woorden*. Leiden: Brill. (Mededelingen vanwege het Spinozahuis, 36)
- Boros, Gábor. 2007. *The “Secularization” of Religious Emotions in Spinoza: Hope and Fear, Love, and Generosity*. Budel: Damon. (Mededelingen vanwege het Spinozahuis, 89)
- Camerer, Theodor. 1877. *Die Lehre Spinoza’s*. Stuttgart: Kessinger.
- Curley, Edwin. (ed. and trans.) 1985. *The Collected Works of Spinoza*. Vol. I–II. Princeton: Princeton University Press.
- Della Rocca, Michael. 1996. Spinoza’s Metaphysical Psychology. In: Garrett, Don (ed.): *Cambridge Companion to Spinoza*. Cambridge: Cambridge University Press.

- Garrett, Don. 2002. Spinoza's Conatus Argument. In: Koistinen, Olli I. – Biro, John I. (eds.): *Spinoza: Metaphysical Themes*. Oxford: Oxford University Press.
- Gebhardt, Karl. (ed.) 1925. *Spinoza Opera*. 4 vols. Heidelberg: Carl Winter.
- Marshal, David. 1986. *The Figure of Theatre: Shaftsbury, Defoe, Adam Smith, and George Eliot*. New York: Columbia University Press.
- Postema, Gerald. 2005. "Cemented with Diseased Qualities": Sympathy and Comparison in Hume's Moral Psychology. *Hume Studies* 31 (2): 249–298.
- Smith, Adam. 1982. *The Theory of Moral Sentiments*. Indianapolis: Liberty Classics.
- Spinoza, Baruch de. 1677. *Opera posthuma*. Amsterdam: Jan Rieuwersz. Complete photographic reproduction, ed. by F. Mignini (<https://www.quodlibet.it/libro/9788874621699>).
- Spruit, Leen – Totaro, Pina. 2011. *The Vatican Manuscript of Spinoza's Ethica*. Leiden: Brill.
- Steinberg, Justin. 2013. Imitation, Representation, and Humanity in Spinoza's Ethics. *Journal of the History of Philosophy* 51 (3): 383–407.
- van Vloten, J. – Land, J. P. N. (eds.) 1914. *Benedict De Spinoza: Opera Quotquot Reperta Sunt*. Vols. I–II. 3rd ed. The Hague: Martin Nijhoff Publishers.

# IX

Gábor Boros

## Life as Death in Spinoza

*XCII Immortals are mortal, mortals immortal, living the others' death, dead in the others' life. (71)<sup>1</sup>*

*XCIII The same . . . : living and dead, and the waking and the sleeping, and young and old. For these transposed are those, and those transposed again are these. (71)*

*CII For souls it is death to become water, for water it is death to become earth; out of earth water arises, out of water soul. (75)*

*XLI The death of fire is birth for air, and the death of air is birth for water. (47) (Fragments of Heraclitus)*

*Transitus ergo de hac vita mortali in aliam vitam immortalem, hoc est enim de morte ad vitam, in passione et in resurrectione Domini commendatur.<sup>2</sup>*

My decisive impulse to reflect upon Spinoza's thoughts on life and death came from Piet Steenbakkers, who presented and argued for his far-reaching emendation of caput V in the appendix to Book 4 of the Ethics. In the emended version the text reads as follows:

Nulla igitur *vita vitalis*<sup>3</sup> est sine intelligentia, & res eatenus tantum bonae sunt, quatenus hominem juvant, ut Mentis vita fruatur, quae intelligentia definitur. Quae autem contra impediunt, quominus homo rationem perficere, & rationali vita frui possit, eas solummodo malas esse dicimus.

1 When quoting Heraclitus I will refer to the page numbers of *Heraclitus* 2004.

2 Augustine: Epistola 55, in [http://www.augustinus.it/latino/lettere/lettera\\_055\\_testo.htm](http://www.augustinus.it/latino/lettere/lettera_055_testo.htm) (15/03/2017).

3 *Vita vitalis* instead of *vita rationalis*.

If we modify correspondingly Edwin Curley's rendering, we have the following translation:

V. No life, then, is worth living / lived in an eminent way without understanding, and things are good only insofar as they aid man to enjoy the life of the Mind, which is defined by understanding. On the other hand, those that prevent man from being able to perfect his reason and enjoy the rational life, those only we say are evil (Spinoza 1988, 589).

Piet Steenbakkens provided the community of Spinozists with an explanation of the emendation in his talk in London on the biannual meeting of ESEMP. The fundamental idea of my understanding of the whole issue appeared to me already during his talk, and this commencement of a kind of maturation process was furthered by the respective lectures of Ursula Renz and Olivér Tóth<sup>4</sup> during the same meeting. When I mentioned my idea to Piet he was far from being as enthusiastic as I was, so I decided to write this paper in order to convince him.

Regarding the perennial big questions of life and death, the views differ as irreconcilably as in all fundamental questions in which the age-old debate within European thought manifests itself: the debate between the ascetic and the hedonistic style of thinking. We are all familiar with the bold Platonic formula maintaining philosophy as a preparation for death and dying. However, the Epicurean formula, according to which nothing touches us less than death, is at least as universally known as the Platonic one. The Judeo-Christian-Greco-Roman culture of Europe has never really overcome this fundamental debate in any of its phases; in a number of respects it is closer, no doubt, to the ascetic views but in a number of other respects it has always been influenced considerably by hedonistic ones. Abelard asserts famously in his *Ethics* concerning the appropriate behavior of a servant lethally threatened by his master that he ought to have resigned himself to being killed, instead of killing the former for no other reason than saving his own life. In the Christian view exemplified here by Abelard's text, no concern for our own individual survival can legitimize the transgressing of the categorical commandment of "Thou shalt not kill!"

When we speak about philosophy's new beginning in the 17<sup>th</sup> century, however, we mean precisely the reversal of the perspectives and of the attitudes of the time. The new philosophy was built on the concept of an almost invincible striving to persevere in one's proper life, although this philosophy itself was also rooted in Judeo-Christian tradition, of which Descartes' metaphor of a tree is a reliable witness. As all parts of a tree are nourished and kept in existence by what they absorb through the roots from the soil, all parts of the newly formed system are permeated by the new philosophy's Christian-metaphysical concept of God which I referred to as "philotheist" in a recent article.<sup>5</sup>

The influence of both these theses – being attracted by a type of asceticism permeated by a metaphysical concept of God, and being based on the concept of a striving to persevere

---

4 See chapter X in this volume.

5 Cf. Boros 2016, 365–377, esp. 366 sq.

in the finite entity's proper being, linked to an at least moderately hedonistic attitude – can even more be maintained of Spinoza. His philosophy is rooted deeply in Jewish, Greco-Roman and Christian thinking, at the same time leading to a system that can be regarded as one of the first great accomplishments of modern European secular thought.

P7: The striving by which each thing strives to persevere in its being is nothing but the actual essence of the thing.

P8: The striving by which each thing strives to persevere in its being involves no finite time, but an indefinite time (E3 P7 & 8, Spinoza 1988, 499).

The absolute dominance of the striving to persevere in the individual's proper being could hardly be expressed less ambiguously; from this perspective, viewed from Spinoza's metaphysical ontology of essences there is scarcely any place left for death. Death, on the other side of the divide, is a phenomenon that can be well-grasped physically, expressing itself in the physical beings' inexorable falling apart. This is an obligatory ingredient of any epicurean physics, such as teaching on the bodies, Spinoza's physics included. From this physical perspective, survival proves to be atypical, as it were. The ever-stable life has no place in Spinoza's physics.

By singular things I understand things that are finite and have a determinate existence. And if a number of Individuals so concur in one action that together they are all the cause of one effect, I consider them all, to that extent, as one singular thing (E2 Def7, Spinoza 1988, 447).

I. The human Body is composed of a great many individuals of different natures, each of which is highly composite.

[...]

III. The individuals composing the human Body, and consequently, the human Body itself, are affected by external bodies in very many ways.

[...]

VI. The human Body can move and dispose external bodies in a great many ways (E2 Post 1, 3, 6, Spinoza 1988, 462).

There is no singular thing in nature than which there is not another more powerful and stronger. Whatever one is given, there is another more powerful by which the first can be destroyed.

[...]

P3: The force by which a man perseveres in existing is limited, and infinitely surpassed by the power of external causes.

[...]

P4: It is impossible that a man should not be a part of Nature, and that he should be able to undergo no changes except those which can be understood through his own nature alone, and of which he is the adequate cause (E4 Ax, P3 & 4, Spinoza 1988, 548).

Given these intrinsic and apparently opposing tendencies – asceticism and hedonism, eternal essence and corruptible existence – one of the most important tasks of Spinoza’s philosophy is to explain how the metaphysical ontology-based concept of life can be maintained in a system together with the physics-based concept of death, of disintegration. Furthermore, how can the concept of God play more than a purely ornamental role in the system? For the concept of God plays a crucial role in propositions that can well be interpreted within the horizon of the then contemporary mystical Christianity, rather than in a simple metaphysical or physical way.

P18: No one can hate God.

[...]

Cor.: Love toward God cannot be turned into hate.

[...]

P19: He who loves God cannot strive that God should love him in return.

[...]

P20: This Love toward God cannot be tainted by an affect of Envy or Jealousy: instead, the more men we imagine to be joined to God by the same bond of Love, the more it is encouraged.

[...]

P36: The Mind’s intellectual Love of God is the very Love of God by which God loves himself, not insofar as he is infinite, but insofar as he can be explained by the human Mind’s essence, considered under a species of eternity; i.e. the Mind’s intellectual Love of God is part of the infinite Love by which God loves himself.

[...]

Cor.: From this it follows that insofar as God loves himself, he loves men, and consequently that God’s love of men and the Mind’s intellectual Love of God are one and the same. [...]

Schol.: From this we clearly understand wherein our salvation, or blessedness, or Freedom, consists, viz. in a constant and eternal Love of God, or in God’s Love for men. And this Love, or blessedness, is called Glory in the Sacred Scriptures<sup>6</sup> – not without reason (E5 P18 & C & 19 & 20 & 36 & C & S, Spinoza 1988, 604 sq, 612).

In the first instance, one can obviously attempt to harmonize the paradoxical relation between life and death in Spinoza either by reducing life to the physical plane or by interpreting death on the metaphysical plane. The following passage is an obvious example of an attempt by Spinoza to do both, because one can easily see its link to the passage on the conditions of possibility of the survival of bodies I quoted earlier:

---

6 As for the passages one can refer to cf. Wolfson 1934/1962, II, 311–317.

I understand the Body to die when its parts are so disposed that they acquire a different proportion of motion and rest to one another. For I dare not deny that – even though the circulation of the blood is maintained, as well as the other [signs] on account of which the Body is thought to be alive – the human Body can nevertheless be changed into another nature entirely different from its own. For no reason compels me to maintain that the Body does not die unless it is changed into a corpse (E4 P39S, C 569).

Thus, we can speak of death physically when the proportion of motion and rest among the ingredients of the body – or to put it in another terminology, its capability to act as a self-identical cause bringing forth its typical effects – changes to an extraordinary measure. This event will unavoidably set in due course in every finite being's life as its proper extreme, its limit value. This is what Spinoza asserts very seriously at the beginning of Part 4 of the Ethics quoted above: the potency-in-act of the external bodies overcomes infinitely that of our own body.

At this point, it seems reasonable to insert a short excursus on the suicide, which Spinoza's official view treats as a special type of the being overcome by the potency of the external bodies.

The well-known sentences in E4 P18S leading to the pronouncement on suicide are as follows:

Since reason demands nothing contrary to nature, it demands that everyone love himself, seek his own advantage, what is really useful to him, want what will really lead man to a greater perfection, and absolutely, that everyone should strive to preserve his own being as far as he can. This, indeed, is as necessarily true as that the whole is greater than its part (see IIIP4).

Further, since virtue (by D8) is nothing but acting from the laws of one's own nature, and no one strives to preserve his being (by IIIP7) except from the laws of his own nature, it follows:

- (i) that the foundation of virtue is this very striving to preserve one's own being, and that happiness consists in man's being able to preserve his being;
- (ii) that we ought to want virtue for its own sake, and that there is not anything preferable to it, or more useful to us, for the sake of which we ought to want it; and finally
- (iii) that those who kill themselves are weak-minded and completely conquered by external causes contrary to their nature (Spinoza 1988, 555 sq.).

First of all, we must be aware of the difficulty the translator faces when attempting to avoid presenting Spinoza's view as appearing more dogmatic than it really is. In the deeper layer of the Latin expression *animo impotens* it refers to what happens if we fail

to preserve our own being physically, while fixing what happens to the one who commits suicide only as a physical being. However, if we want to understand Rubens' painting on Seneca's suicide we will clearly miss the point if we summarize the event by asserting that Seneca was "weak-minded and completely conquered by external causes contrary to [his] nature".

On the one hand, it can hardly be denied that if someone intends to remain healthy as long as she can, it will be necessary for her body to cooperate with the external bodies – "causes" – in a moderate, well-tempered way, controlled by herself. On the other hand, the immoderate cooperation, i.e. the one that gets out of her control will immediately turn against life, and not only against life as a physical-ontological phenomenon. This will be destructive for the respective proper life as well, insofar as the dynamic multi-level causal flow constituting her environment is no more governed by her, by her mind in the manner of a "spiritual automaton," but by the embodied external powers. The embodiment of external powers does not relate to external bodies only. The ingredients of her own body threaten to merit less and less the adjective "proper" in the sense in which Ethics 2 Ax 4 singles out the respective "proper" body from among all other bodies of the environment.

We feel that a certain body is affected in many ways (Spinoza 1988, 448).

This means, that the complex being consisting of a soul-or-mind and a body has come to the point where it either does not at all feel anything anymore, or where its sensible particles do feel – or even all the particles may perceive, like in Diderot, a century later – but they do not have any felt experience whatsoever of I. They are no longer capable of singling out a certain affected body against all others, or against the whole in which all bodies are dissolved in the final analysis.

So far so good. But nothing has been said concerning death up to now that regards suicide as a particular form of death. For, agony can well be described generally as our becoming impotens, incapable to govern for ourselves as spiritual automata the order of the influences of external causes – that could include the particles of our own bodies, in extreme cases of lethal illness. But Spinoza does tell us more about suicide than what he says about death in general, and I propose to search for this "more" in the qualification of "impotens" in our above quotation, namely in the adverb *animo*. Those who commit suicide are not simply impotens but precisely *animo* impotens, i.e.: the origin of the suicidal person's terrifying failure in preserving her being – in complying with the "normal" burden of external causes in one of the ways "normal" people do – is to be looked for in the soul or in the mind.

I do not want, however, to maintain that Spinoza worked out a sophisticated theory of the various species of death qua becoming impotens. One of the obvious reasons for not claiming this is that on the one hand, in Spinoza the differentiation between living and non-living things is not sufficiently clear. On the other hand, it is not an easy task to separate unambiguously the events that happen to the soul/mind from those happening to the

body. Everything that happens in any individual can equally be interpreted on the level of the body and on that of the soul/mind. So I would confine myself to a rather cautious interpretive statement. First, death as becoming impotens can always be interpreted as a process on both the bodily-physical, and on the metaphysical/ontological level of the soul/mind. In the case of suicide however, we have to reckon with a bifurcation on the metaphysical plane. When interpreting suicide, not only the bodily-physical dimension, but also the part of the metaphysical dimension that can be analyzed on the level of reason, without considering intellectually the individual qua this unique individual will appear precisely as irrelevant for those explaining the occurrences leading to the agony. In this way, and corresponding to the well-known differentiation between the three kinds of cognition, we can differentiate between three layers in Spinoza's interpretation of death. The first is the causal level of bodily events becoming fixed in the body/soul as traces of imagination and memory. The second is that of the processes in the soul that can be interpreted as a type of proto-psychology by way of reason's general laws similar to the adequate knowledge of the physical phenomena. The third layer is that of the ineffable individuality, as it were, informed by the intellect. Here the individual does not only feel the involuntary manifestations of the general causal system of bodies in a "certain body" – such manifestations being pleasure, pain, etc. – but in addition she is also able to express their belonging to her as a unique individual through reflective knowledge. In this latter case, the relevant issue is not that she can make an objective-scientific statement belonging to the second layer, but that she cannot help integrating all possible scientific statements concerning her body-soul-mind in her literal lifelong striving for her real individual happiness; for not only preserving her being but continually increasing her intellect-based power of acting. Only the presupposition of this third layer of knowing herself establishes the accentuated sense for individuality that appears first in the very short preface to Book 2 of the Ethics,<sup>7</sup> in order to unfold in those passages of Book 5 that belong to the so-called intuitive science as the highest mode of cognition.<sup>8</sup> All these passages aim at the attainment of true individual happiness, whose truth and individuality stem evidently from its being the farthest from any kind of egotism. This is the real happiness Spinoza is looking for from the very begin-

7 "I pass now to explaining those things which must necessarily follow from the essence of God, or the infinite and eternal Being – not, indeed, all of them, for we have demonstrated (IP16) that infinitely many things must follow from it in infinitely many modes, but only those that can lead us, by the hand, as it were, to the knowledge of the human Mind and its highest blessedness" (Spinoza 1988, 447).

8 "Again, because the essence of our Mind consists only in knowledge, of which God is the beginning and foundation (by IP15 and IIP47S), it is clear to us how our Mind, with respect both to essence and existence, follows from the divine nature and continually depends on God. I thought this worth the trouble of noting here, in order to show by this example how much the knowledge of singular things I have called intuitive, or knowledge of the third kind (see IIP40S2) can accomplish, and how much more powerful it is than the universal knowledge I have called knowledge of the second kind. For although I have shown generally in Part I that all things (and consequently the human Mind also) depend on God both for their essence and their existence, nevertheless, that demonstration, though legitimate and put beyond all chance of doubt, still does not affect our Mind as much as when this is inferred from the very essence of any singular thing which we say depends on God" (E5 P36S, Spinoza 1988, 613).

ning of his career as a philosopher, i.e. from the famous prefatory note of the Treatise on the emendation of Intellect: it belongs to the individual happiness to facilitate attaining the same happiness for as many people as possible. In Prop. 20 of Book 5 of the Ethics Spinoza rounds off the unfolding of this fundamental ethical-philosophical idea as follows:

This Love toward God cannot be tainted by an affect of Envy or Jealousy: instead, the more men we imagine to be joined to God by the same bond of Love, the more it is encouraged (Spinoza 1988, 605).

What can we say about the particularity of suicide after our excursus on the place of intellect-based individuality in Spinoza? In the final analysis, Spinoza interprets suicide as the failure to govern the causal flow directed to the life lived in the most “proper” way possible for a human being, to attain real individual happiness together with as many other individuals as possible. Failure leading to suicide consists, therefore, of the impossibility of communicative community, and between the given individual and her human entourage to be realized by way of the true insight in the true and highest good (to use the terminology of TIE), and consequently in truly loving God, where the failure lies on the individual’s side, or at least the individual is deeply convinced that the failure is hers.<sup>9</sup>

It was not my plan to focus my treatment on suicide as such. Nevertheless, I believe it was important to dwell on it for a while. I hope it has become more than obvious that the question of death cannot be reduced to the bodily-physical level, however important it might be. The analysis of suicide has opened up a symbolic space including such questions as that of the happiness achieved through reasonably ascetic virtue, of the “true” religion connected to the “true” love of God, and, finally, of the community, based on the factors mentioned earlier. Therefore, we must be extremely cautious when interpreting the famous assertion with an obviously Epicurean flavor that has by now almost become Spinoza’s trademark:

A free man thinks of nothing less than of death, and his wisdom is a meditation on life, not on death.

Dem.: A free man, i.e., one who lives according to the dictate of reason alone, is not led by Fear (by P63), but desires the good directly (by P63C), i.e. (by P24), acts, lives, and preserves his being from the foundation of seeking his own advantage. And so he thinks of nothing less than of death. Instead his wisdom is a meditation on life, *q.e.d.* (E4 P67, Spinoza 1988, 584).

It is certainly true that the basic attitude of Spinoza’s sage on death is not Platonic. He refuses decidedly the concept of a life starting before birth and continuing after the death of the body, whose advocates insist on the individual-personal immortality of the pure

---

<sup>9</sup> This might be interpreted in context of the self-hatred Keith Green analyzes in connection with P30 of E3; see chapter VIII in this volume.

soul based on imagination and memory. Spinozian wisdom is obviously not a meditation on life in the sense of a future, “after-death”. “truer” life than the present life of the body-soul-mind. Wisdom for Spinoza is not a meditation on an externally given or promised life but much more on life in the sense of a continually present renewal of life as a task. This implies also Spinoza’s answer to the question of why death is not to be feared – a question that was given a central role already in Descartes’ late philosophy. Descartes answered it relying on his concept of love. It is in the most crucial passages in the closing propositions of the Ethics that Spinoza develops his answer related also to the concept of love indirectly: the life of a man of wisdom is only possible in the horizon of the intellectual love of God.

The usual conviction of the multitude seems to be different. For most people apparently believe that they are free to the extent that they are permitted to yield to their lust, and that they give up their right to the extent that they are bound to live according to the rule of the divine law. Morality, then, and Religion, and absolutely everything related to Strength of Character, they believe to be burdens, which they hope to put down after death, when they also hope to receive a reward for their bondage, that is: for their Morality and Religion. They are induced to live according to the rule of the divine law (as far as their weakness and lack of character allows) not only by this hope, but also, and especially, by the fear that they may be punished horribly after death. If men did not have this Hope and Fear, but believed instead that minds die with the body, and that the wretched, exhausted with the burden of Morality, cannot look forward to a life to come, they would return to their natural disposition, and would prefer to govern all their actions according to lust, and to obey fortune rather than themselves (E5 P41S, Spinoza 1988, 615. Emphasis added).

On the one hand, this passage shows eloquently that what truly matters to Spinoza is not death in the sense of the physical disintegration of bodies – for this remains an unalterable fact howsoever successful our medical methods might prove to be in extending life-expectations. For him, the symbolic level is incomparably more important than the medical one. To this level belongs questions such as which kinds of moral, religious, political concepts we employ when grasping and integrating death in the layer of our thinking, through which we tend to acquire individual happiness. The passage quoted above shows unmistakably that on the symbolic level, the interpretation of physical destruction as an absolute caesura is for Spinoza a vital error. Parallel to the Appendix to Book 1 of the Ethics, the anthropocentric teleology as a general principle of interpreting the world is shown to be a fundamental prejudice. In the final passages of Book 5 he assures the reader that death as physical disintegration is of no relevance for the sage unless it is complemented with an interpretation on the symbolic level.

The contrast between the everyday conception of death and that of the Spinozian sage as a reasonably ascetic virtuous person can easily be grasped in Spinoza’s own terms. Everyday thinking is characterized and governed by imagination as the lowest kind of knowledge.

Most people are overwhelmed with desires whose objects can only be possessed by one singular individual; most people refrain themselves from scrounging up these goods only out of fear for the laws of the state and from religious commandments. Death as the final caesura means for this type of thinking that, mysteriously, by the separation of the soul from the body, no limits upon the enjoyment of the usually desired goods (paradoxically enough, mostly of corporeal character) will remain in force, and a happy state of total gratification of all passions and desires will set in. This state is structurally similar to the “state” Hobbes depicts, as that of the this-worldly happiness, except that Spinoza’s “multitude” imaginatively transfers this state precisely to the “felicity God hath ordained to them that devoutly honour Him”. “Continual success in obtaining those things which a man from time to time desireth, that is to say, continual prospering, is what men call FELICITY; I mean the felicity of this life. For there is no such thing as perpetual tranquility of mind, while we live here; because life itself is but motion, and can never be without desire, nor without fear, no more than without sense. What kind of felicity God hath ordained to them that devoutly honour Him, a man shall no sooner know, than enjoy; being joys, that now are as incomprehensible as the word of Schoolmen, beatifical vision is unintelligible” (Hobbes 1996, 41 sq.).<sup>10</sup>

The mysterious hope for the dissolution of all material constraints as a consequence of death, this absolute caesura, does in no way constitute the aim of virtuous life for Spinoza’s sage. His fundamental aim is a this-worldly aim, but not in the sense of the limitless satisfaction in possessing more and more bodily goods. On the contrary, his aim is that already the life in the finite human body be governed by the love of God originating in the reason- or intellect-based cognition that outweighs and overcomes the imaginative ways of cognition, the basis for the anthropocentric illusions that fuel the “moral” of the multitude. For him, the virtuous life does not consist in obeying externally given laws and commandments in order to turn one’s back to virtuous life as soon as those laws and commandments cease to be given or valid. His virtuous life comes from inner ethical insights built on the eternal laws of the love of God.

---

10 The following formulation belongs also to this context inasmuch as Spinoza summarizes here his opinion about the mentality and moral attitude of the everyday people – which is similar to Hobbes’ – and his characteristically Spinozian idea of true happiness – that he does not posit in the heaven like Hobbes does. “And even if it can happen that a greedy, ambitious, or timid man abstains from too much food, drink, and sexual union, still, Greed, Ambition, and Timidity are not opposites of gluttony, drunkenness, or lust.

For the greedy man generally longs to gorge himself on another’s food and drink. And the ambitious will not be moderate in anything, provided he can hope he will not be discovered; if he lives among the drunken and the lustful, then because he is ambitious, he will be the more inclined to these vices. Finally, the timid man does what he does not wish to do. For though he may hurl his wealth into the sea to avoid death, he still remains greedy. And if the lustful man is sad because he cannot indulge his inclinations, he does not on that account cease to be lustful.

Absolutely, these affects do not so much concern the acts of eating, drinking, etc., as the Appetite itself and the Love. Therefore, nothing can be opposed to these affects except Nobility and Tenacity, which will be discussed later on.” (E3 AD48S, Spinoza 1988, 541 sq.)

This explanation is, however, not at all new for those who take the Ethics Part 5 seriously. Nor do I want to emphasize the result itself. Instead it is the transition that is worthwhile being emphasized, the transition we must presuppose from the state of the man governed by imagination to that of the sage. Spinoza belongs in this respect to the representative thinkers of the third big European tradition, to those who conceive of death in a symbolic sense not as the absolute caesura of destruction but as a caesura that transfers the low level life into a higher level life, “that can lead us, by the hand, as it were, to the knowledge of the human Mind and its highest blessedness” (E2 Pref. Spinoza 1988, 446). This transition, transformation and transference can also be viewed as rebirth, not bound to any change in the physical composure of the body nor to the absolute separation of the soul from the body. Spinoza renews on the one hand the philosophical-mystical tendency in Heraclitus, while on the other hand he transposes in the Early Modern philosophical context a strong chain of Christian thinkers that starts with St. Paul and St. Augustine, continues in the mystically oriented Christian and Jewish Platonists in the patristic and scholastic periods in order to arrive finally to the 17<sup>th</sup> century in the works of authors such as Pascal, or the partisans of “pure love”. The purists of love aim at the re-birth of the ego as the new, pure ego, giving birth within himself to Jesus Christ, by the death, the annihilation, of the egotist ego. Pascal conceives the conversion to a stricter Christian belief and life-praxis also as the death of the old man and the birth of a man renewed and rescued from death.

I do not want to claim that Spinoza was directly influenced by these thinkers in a philologically demonstrable way, although it is tempting to point out the well-known fact that Spinoza owned the complete works of Augustine. Instead of this, I intend to convey one more example in support of Hans Blumenberg’s theory of legitimate re-appropriation in history; and with this aim in mind, I interpret the relevant passages of the Ethics as the (re-)occupation, the philosophical re-casting (*Umbesetzung*) of a place, of a topos that had become empty after a philosophical “paradigm shift”. As the above list of authors show, the topos played a dominant role in earlier theology and philosophy oriented to the transcendent God, but it became empty through the appearance of the immanence-oriented, pantheistic way of thinking that became more and more the “natural system of philosophy in the 17<sup>th</sup> century” (Dilthey). Spinoza was certainly not the only thinker who tried to re-occupy this topos. There were both theological and laical-religious movements who tried to do the same, partly even parallel to and profiting from Spinoza’s philosophical re-occupation, as in the case of the collegian circle around him.<sup>11</sup> The emendatio of the text of Caput V from *vita rationalis* to *vita vitalis* by Piet Steenbakkers with its newly gained Augustinian momentum is clearly a witness to this affinity on the side of Spinoza.

Let us return to Spinoza’s interpretation of death on the symbolic level without identifying it with bodily destruction. Earlier I quoted the passage where Spinoza defines death almost formally as the destruction of the body. I have already hinted at the fact that this is

---

<sup>11</sup> This affinity between religious movements and at least some segments of Spinoza’s philosophical system is shown by Schmidt-Biggemann 1988.

not the most important “considered” interpretation of death we have from Spinoza. Now I will quote the same passage further, where the amnesia of a Spanish poet and the relation between being a child and being an adult for the same person appear in a peculiar way as examples of death:

And, indeed, experience seems to urge a different conclusion. Sometimes a man undergoes such changes that I should hardly have said he was the same man. I have heard stories, for example, of a Spanish Poet who suffered an illness; though he recovered, he was left so oblivious to his past life that he did not believe the tales and tragedies he had written were his own. He could surely have been taken for a grown-up infant if he had also forgotten his native language. If this seems incredible, what shall we say of infants? A man of advanced years believes their nature to be so different from his own that he could not be persuaded that he was ever an infant, if he did not make this conjecture concerning himself from [NS: the example of] others (E4 P39S, Spinoza 1988, 569).

I find this passage crucial with respect to its implicit differentiation between two kinds of death, neither of which is absolute in the sense of total destruction. It seems in no way superfluous even to note that the differentiation does not run parallel to the one applied in today’s bioethics between the death of the brain and the death of the whole person, used to establish the legal foundation for an organ transplant. Spinoza goes in a metaphysical direction without a transcendental-theological dimension. The former relates him to the above mentioned thinkers’ views; the latter separates him from them. He joins here the tradition that considers radically changed life as death from the point of view of the respective previous life. In contrast with the philosophical-theological tradition, Spinoza does not conceive of death as the moment after which the question might arise, when the soul, in her bodiless state, or wrapped in a body finer than her earthly one, changes its nature transmuting to a transcendent status. In this respect, we cannot have any doubts: Spinoza is convinced that being bothered with this question is not worthy of the sage. To consider the radically changed life as death does not unconditionally mean a departure from the sphere of immanence – even relative to a Christian context – nor does it render the philosopher applying this conception of death unworthy of the title of a sage.

The Heraclitian lines quoted as our motto shows that the concept of life as death can appear on an elementary level as value-neutral:

XLI [The death of fire is birth for air, and the death of air is birth for water.] (47)  
Fire lives in the death of earth, air lives in the death of fire, water lives in the death of air, and earth in the death of water.<sup>12</sup>

---

12 Cf. 25 B76(?), The G.W.T. Patrick translation: <http://classicpersuasion.org/pw/heraclitus/herpatu.htm> (2016/09/10).

At the same time, the attentive reader comes across, among Heraclitus' lines, the moment of evaluation which Spinoza would eventually employ in a similar way.

For it is death to souls to become water, and death to water to become earth. But water comes from earth; and from water, soul (DK B36).

I already mention the Spinozian parallel here, since the soul's becoming water does not mean death in an absolute sense, either in Heraclitus or in Spinoza. For both thinkers, this means something pleasurable for those who are far from being a sage, whereas both thinkers believe the sage must avoid it:

CVIII [It is delight, not death, for souls to become moist.] (77)

CVI A man when drunk is led by a beardless boy, stumbling, not perceiving where he is going, having his soul moist. (77)

CI To the soul belongs a report [logos] that increases itself. (75)

CIX A gleam of light is the dry soul, wisest and best. (77)

In no way would I like to suggest that Spinoza's concept of life as death is identical with that of Heraclitus. Among many other differences, I would especially like to draw attention to a peculiar feature of the Spinozian idea. For in the fragments quoted above we find the obvious persuasion that death is something negative, which easily suggests an analogy, or even a metaphorical connection, between the two transmutations; the living becoming death and the sober soul becoming drunken, wet. To see this connection it's not even necessary to compare the wet-soul to the sage; a comparison with average people strikes us already. In Heraclitus, the adult becomes more childish than a true child, when touched by this type of death, because he is guided by a true child, not even by an adolescent of the everyday type. However, the opposite case does not seem to appear in Heraclitus: he does not consider death if someone becomes sage from the status of possessing but an everyday wisdom.

Now, precisely this happens in Spinoza. I quote again:

If this seems incredible, what shall we say of infants? A man of advanced years believes their nature to be so different from his own that he could not be persuaded that he was ever an infant, if he did not make this conjecture concerning himself from [NS: the example of] others (Spinoza 1988, 569).

The transition from childhood to the status of an adult appears in this light as an everyday, secular example of life as death – as opposed to religious examples such as the taking of monastic vows – so much so that there is no need to become a true sage, even though ideally, the alteration aims at such a full transmutation.

And really, he who, like an infant or child, has a Body capable of very few things, and very heavily dependent on external causes, has a Mind which, considered solely in itself, is conscious of almost nothing of itself, or of God, or of things. On the other hand, he who has a Body capable of a great many things, has a Mind which, considered only in itself, is very much conscious of itself, and of God, and of things.

In this life, then, we strive especially that the infant's Body may change (as much as its nature allows and assists) into another, capable of a great many things and related to a Mind very much conscious of itself, of God, and of things. We strive, that is, that whatever is related to its memory or imagination is of hardly any moment in relation to the intellect (as I have already said in P38S) (E5 P39S, Spinoza 1988, 614).

I am convinced that even if Spinoza does not explicate this step, we can complete the examples of the two types of life as death we have found up to now, both in Heraclitus and Spinoza, by a third type. What I have in mind is the transition from the average adult living in the state of everyday thinking into that of the philosopher who lives her wisdom, as it were; who is eternal but not immortal in the everyday sense of the word. It seems evident that in the early modern European context, it is the terminology and liturgical praxis of Christian tradition that renders it understandable if a philosopher conceives of the transition from a lower to a higher quality of life in terms of death, even if the thinker, in our case Spinoza, does not develop a traditional Christian philosophy. Precisely at this point, I think it is very useful to consider his case as a case in point with regard to Blumenberg's theory of *Umbesetzung*.

Since I cannot point out any strict philological support in the text of Spinoza, I must confine myself to underpinning my thesis by highlighting some structural convergences.

In the very year of Spinoza's excommunication in 1656, Pascal corresponded with Mlle Roannez, for whom he was a spiritual leader, keen on guiding her to Jansenist spirituality. In the first letter, Pascal identifies the conversion straightforwardly with death, with the destruction of the old man:<sup>13</sup>

I should think, that the prediction of the overthrow of the ancient temple, which prefigures that of the old man in each of us, and wherein it is said, that 'not one stone shall be left upon another', indicates that none of the former passions of the corrupt nature shall remain; (36)

Pascal goes on in this very letter to hint at the theodicy-like apology of death as the extermination of the root of the sin lurking necessarily in the body, in the flesh. Obviously, however, what he writes is true of the death of the old man destructed in the conversion:

---

13 I will quote from the 19<sup>th</sup> century translation, where the numeration of the letters differs from the later editions. The numbers in brackets will refer to Pascal 1849.

“Then death becomes necessary, for the entire destruction of this fatal root of corruption; and the Christian is led to desire it for that purpose”. (37) The idea of life as death takes on a transcendent meaning in his thought, thus becoming virtually “death as life”.

God never abandons his own, even in the grave; then their bodies, though dead to the eyes of men, are even more alive than before with the Lord, because sin is extinguished in them; (37)

Pascal identifies again metaphysically the peace of God in Jesus Christ with death desirable because of the destruction of the body, of the flesh corrupted through the root of sin.

[I]t will not be perfected till the body shall be destroyed; and this it is which makes us desirous of death, while we nevertheless manfully endure life, through love to Him who for our sakes endured both life and death. (26)

From our point of view, the most pertinent formulation can be found in letter 5 dated on 5<sup>th</sup> November 1656. This is the most pertinent for two reasons. On the one hand, Pascal makes here an explicit reference to the main source of these very thoughts: St. Paul’s letter to the Colossian, verse 3, lines 1–17. On the other hand, because we can discover its parallel structure to Spinoza’s chain of ideas – supposing obviously that we do not expect or postulate in Spinoza’s case the immortality of man and a transcendent supporter of it. It is striking in a sense to realize the possibility of reading St Paul in this rather immanentist manner, whereas Pascal’s strong emphasis on the desirability of death renders the same text much more exalted than the original.

“Our old man perishes”, says St. Paul; “yet the inward man is renewed day by day”; and it will be perfectly renewed only in eternity; where we shall sing without ceasing that new song of which David speaks – that is, the melody which springs from a renewed and loving spirit. (24)

The intonation of the passage from the letter to the Colossian titled “The old and the new man” addresses first death taken in its everyday sense. Afterwards it employs a formulation that makes it evident that its true reference is to the status of the homo viator who perishes from day to day, on the one hand, qua “old man” and, on the other, will become renewed, qua “new man”.

3 Therefore, if you were raised with Christ, look for the things that are above where Christ is sitting at God’s right side. 2 Think about the things above and not things on earth. You died, and your life is hidden with Christ in God.<sup>14</sup>

---

<sup>14</sup> Biblical quotations are from the „Common English Bible”.

As a multiple warning, command follows this intonation that can only be related to the this-worldly wanderer. Otherwise, the verbs would be in the future tense instead of the imperative as they stay in reality.

So put to death the parts of your life that belong to the earth, such as sexual immorality, moral corruption, lust, evil desire, and greed (which is idolatry). But now set aside these things, such as anger, rage, malice, slander, and obscene language. Don't lie to each other. Take off the old human nature with its practices and put on the new nature, which is renewed in knowledge by conforming to the image of the one who created it. [...] put on compassion, kindness, humility, gentleness, and patience. Be tolerant with each other and, if someone has a complaint against anyone, forgive each other. As the Lord forgave you, so also forgive each other. And over all these things put on love, which is the perfect bond of unity. The peace of Christ must control your hearts – a peace into which you were called in one body.

The text itself can evidently be read as prophesizing the annihilation of the this-worldly state of the soul resurrected without a body or in another kind of body. If we think of the first and second letters of Pascal, of the desirable death, then we must say that he interpreted Saint Paul's text in this manner. However, if we investigate the quotation from letter 5, then we find an interpretation of the basic text, which is an unambiguous hint for the this-worldly reader, which can run parallel to the immanent perspective of eternity that dominates the closure of Ethics. Those warnings in the form of propositions, corollaries and scholiums intend to reach the sober thinkers among everyday people and to provide them with an idea of the truly human life in the intellectual love of God that can only be acquired through a radical change in the governance of life, comparable to the emptying of the soul from the egotist self and yielding to the true governor: for the religious conversion to God, in the person of Jesus Christ, in the case of Spinoza's sage to God, in the form of His infinite intellect with much similarity to Platonic philosophical theologies, Christian and Jewish alike. When in the end of his letter published in Florence in 1675, Steno urged Spinoza to convert to Christianity and thus become a new Augustine,<sup>15</sup> he might have fallen back to earlier conversations about parts of his Opera they might have discussed in the circle around Spinoza. He seems, however, to have failed to notice that Spinoza did make a conversion. It was not he who underwent a conversion to Christianity, but who enabled one of the most important categories of Augustinian spirituality to undergo a conversion to a confessionally neutral, philosophical concept of a *vita vitalis* that could and can even today be accepted by religious and agnostic philosophers alike.

---

<sup>15</sup> Spinoza 2016, 458.

## References

- Boros, Gábor. 2016. The Anti-metaphysical Turn in the Philosophy of Love (From the Cartesian Perinatal Perspective on Love to Diderot's Free Love in His Supplement to Bougainville's Journeys). *Fudan Journal of the Humanities and Social Sciences* 9 (3): 365–377. esp. 366 sq. (DOI 10.1007/s40647-016-0134-y).
- Heraclitus. 2004. *The art and thought of Heraclitus*. An edition of the fragments with translation and commentary by Charles H. Kahn. Cambridge: Cambridge University Press.
- Hobbes, Thomas. 1996. *Leviathan*. Oxford: Oxford University Press.
- Pascal, Blaise. 1849. *The miscellaneous writings of Pascal; consisting of letters, essays, conversations, and miscellaneous thoughts*. London: Longman, Brown, Green, and Longmans.
- Schmidt-Biggemann, Wilhelm. 1988. Veritas Particeps Dei. Der Spinozismus im Horizont mystischer und rationalistischer Theologie. In: *Theodizee und Tatsachen. Das philosophische Profil der deutschen Aufklärung*. Frankfurt a/M.: Suhrkamp.
- Spinoza, Benedictus de. 1988. *The Collected Works of Spinoza*. Vol. I. Edited and translated by Edwin Curley. 2nd ed. Princeton: Princeton University Press.
- Spinoza, Benedictus de. 2016. *The Collected Works of Spinoza*. Vol. II. Translated by Edwin Curley. Princeton, N.J.: Princeton University Press.
- Wolfson, Harry Austryn. 1934/1962. *The Philosophy of Spinoza: Unfolding the Latent Processes of His Reasoning*. Vol. I–II. Cambridge: University Press.

# X

*Olivér István Tóth*

## Is Spinoza's Theory of Finite Mind Coherent? – Death, Affectivity and Epistemology in the *Ethics*<sup>1</sup>

*Imperious Caesar, dead and turned to clay,  
Might stop a hole to keep the wind away.  
Shakespeare: Hamlet. Act V, Scene 1.*

According to most interpretations, one of the central features of Spinoza's philosophy of mind is that ideas are at the same time logical and psychological entities. Ideas receive their logical, as well as psychological roles in the same way: both are defined by the causal history of the idea. While it is certainly true that Spinoza's philosophy of mind admits only one kind of thing, namely modes of thought (Renz forthcoming), conflating the logical and psychological roles of ideas naturally leads to identifying epistemology with affectivity. On its face, this identification seems to be a natural consequence of Spinoza's philosophy of mind. Every conscious idea that has a psychological role is related to the conatus of the subject and thereby qualifies as affective (Marshall 2014; Malinowski-Charles 2009). Also, knowledge by definition is an action of the mind and thereby an active affect. If the affective aspect of an idea and the knowledge claim of that idea are simply different descriptions of the same thing, it means that ultimately, truth is just a special kind of feeling.

This identification of truth and affectivity can be reconciled in either of two ways. First, by emphasizing that truths are cognitive counterparts of the affective working of our mind. There is no mind-independent truth; what we hold to be true is what we feel to be beneficial for us or for our social community (Lenz 2013). Alternatively, by emphasizing that

---

1 An earlier version of this paper was presented at the "Life and Death in Early Modern Philosophy" conference organized by the British Society for History of Philosophy and the European Society for Early Modern Philosophy on 14–16 April 2016. I would like to thank the audience there for their feedback, especially Julie R. Klein and Mogens Lærke. I would like to thank Ursula Renz and Gábor Boros for their comments on an earlier version of the manuscript, although any errors are my own and should not tarnish the reputations of these esteemed persons.

affects are confused expressions of knowledge claims, wise passions, so to speak. Agreeing on reasons for actions entails having the same affective functioning (Sangiaco 2015a, 2015b). Therefore, a successful political project can be achieved by aligning our feelings just as well as by deliberating about the right policies.

Sangiaco has argued that this identification of psychological and logical aspects of the idea results from Spinoza's identification of will and intellect entailing the identification of certainty and truth: if having an idea and believing the idea is the same thing, then error is nothing else than having a confused and obscure idea. In his reading, the ethical intellectualist project of the young Spinoza failed because this identification cannot be maintained (cf. Renz 2015): akratic behavior shows that sometimes knowledge does not guarantee virtuous action. Therefore, Spinoza revised his philosophy of mind in the *Theological-Political Treatise* and the *Ethics*, where ideas have their own power independent of their epistemic status, and therefore, inadequate but powerful ideas can defeat less powerful rational ideas (Sangiaco 2015c). Spinoza describes this power of an idea in affective terms (cf. Della Rocca 2003): the power with which an idea affects our affective life is the same as the power of the idea with which it affects our deliberations and reasoning. This revision of Spinoza's philosophy is what prompted the identification of the psychological and logical aspects of the idea, whereas in Spinoza's early works, ideas were representations automatically qualifying as beliefs and reasons for action, where a change in belief entailed a change in representation. In his mature works, ideas are representations and reasons for actions, of which the most powerful qualify as beliefs.

In this paper, I argue that accepting this interpretation of Spinoza's mature philosophy on its own does not solve the problem generated by the identification of will and intellect. I claim that even if for Spinoza, affectivity and epistemology are the same, and representations qualify as beliefs, not simply because of their being representations, but because of their being the most powerful representations, the problem of ethical intellectualism is preserved. Given Spinoza's identification of epistemology and affectivity reflected in his identification of ethical good with usefulness for persevering in being, this ethical intellectualism is markedly different from its usual forms. Since virtue is self-preservation for Spinoza, what is necessary and sufficient for virtue (i.e. having the right kind of ideas) is also necessary and sufficient for self-preservation. Although I am not sure that we can only have the right kind of ideas by having knowledge, by having knowledge we necessarily can have the right kind of ideas. Given that nothing can follow from our nature that destroys us and that knowledge follows from our nature alone, knowledge is always conducive to self-preservation.

This, however, generates an unwelcome consequence for Spinoza: his adherence to universal intelligibility combined with this peculiar form of ethical intellectualism rules out the necessity of death. This is shown by the difficulties Spinoza faces when trying to demonstrate the necessary finitude of human life. As I will show at the end of this paper, problems related to this identification are not novel to Spinoza's philosophy: in the scholastic tradition they were part of the problem of material intellect. Since Spinoza was aware of this tradition (Nadler 2001; Klein 2014; Adler 2014), comparing his solutions to the traditional ones can help us better understand the genesis and the program of Spinoza's philosophy, and in it the role of what Sangiaco calls the temptation of the intellect.

My argument, in short, is that Spinoza embraced all four of the following inconsistent claims:

- (1) Life is the preservation of one's essence; death is the destruction of this essence brought about by a harmful external cause.
- (2) Everything is intelligible: there is no such thing – and therefore no such external cause – which an actual human mind cannot, in principle, form an adequate idea of.
- (3) Ethical intellectualism: of which we form an adequate idea cannot be harmful.
- (4) It is necessary that human individuals die.

Claims (1)–(3) imply that knowledge can always avert death, while claim (4) is simply the denial of this implication. Since accepting the identification of the logical and epistemological aspects of the idea leads to this contradiction, these two aspects of the idea have to be determined independently, and therefore they both provide independent sources of knowledge (Boros 1997).

In section 1, I introduce Spinoza's definition of death according to which it is the disruption of the body's essence by a harmful external cause. In sections 2 and 3, I argue that Spinoza indeed embraced universal intelligibility and ethical intellectualism. In section 4, I present Spinoza's demonstration of the necessity of death and argue that he is unable to prove it. In section 5, I argue that Spinoza cannot easily reject either ethical intellectualism, or universal intelligibility. Therefore, this incoherence is not just the result of the careless acceptance of a superficial statement, but rather a deeply embedded feature of his philosophy of mind. In section 6, I will place this incoherence in a historical context and argue that a similar problem existed in the Medieval philosophy of mind which influenced Spinoza.

## **1. Definition of Death: Death is the destruction of the body's essence by a harmful external cause**

The most famous treatment of life and death in Spinoza's *Ethics* comes in E4p39,<sup>2</sup> where Spinoza defined good and evil by the influence external causes have on the proportion of motion and rest with which parts of the human body communicate. This proportion is, by the definition of the Physical Digression, the essence of the human body: those external

---

2 All references to the English translation of works by Spinoza are from Curley's edition with the usual abbreviation: prae – preface, a – axiom, p – proposition, s – scholium, c – corollary, app – appendix, d – definition if it is immediately after the number of the part and demonstration in all other cases. TTP to the *Theological-Political Treatise* followed by the number of chapter and paragraph (Spinoza 1988, 2016).

causes are good which preserve this essence, or even enhance its power in order to be capable of doing many things at the same time (cf. Sangiacomo 2013). On the other hand, external causes bringing about a radical enough change in this proportion, alter the essence of the human body. Since by changing essence the human body loses its identity and is thereby destroyed, the external causes bringing about this change are evil. Thus, preservation of the essence of the human body is life, which is good, while destruction of this essence is death, which is evil.

In the scholium to the proposition, Spinoza elaborates on his definition of death. Since death is the loss of identity due to change in essence, the death of one individual is the birth of another. This definition of death is a revisionary and not descriptive one: as the famous case of the Spanish poet shows, the circulation of blood and other features of the body by which the layman identifies the poet as still living is maintained, yet since he has changed essence he has died and a new person has been born. It must be noted that however perfect and powerful is the resulting new individual born from the death of the previous one, for the previous individual its death is always evil. For a horse, it is equally evil to change into an insect or into a man (E4prae).

Since the essence of the human body, the persistence of which is life, is the conatus (E3p9s), death, the destruction of the conatus, cannot come about by a cause internal to the human body, only external to it (E3p10). As Spinoza explains in E4p20s, this external cause could destroy the essence of the body in many ways: by directly destroying it (e.g.: when a sword is plunged in one's chest), by affecting its imagination though hidden external causes changing its nature (e.g.: when someone believes that it is better for her to be dead than alive), or by creating such an environment that a quick death is preferable to dying slowly (e.g.: when Nero orders the suicide of Seneca).

So, propositions E4p39 and E3p10 together provide the definition of death, which is: *the destruction of the essence of the human body brought about by a harmful external cause.*

## **2. Universal intelligibility: there is no such external cause of which we cannot form an adequate idea**

It is a general feature of Spinoza's philosophy that he is committed to universal intelligibility, which was noted by scholars emphasizing his use of the Principle of Sufficient Reason (Della Rocca 2008), as well as his commitment to realist rationalism (Renz 2010). This strong commitment is the most evident in E1a1–2: here Spinoza states that everything is either in itself, or in another (E1a1), which allows for its understanding either through itself or through another (E1a2). Since we have an adequate idea of what is in itself – God's essence – (E2p47) we are able to know everything that depends on and follows from God's

essence, which is of course everything there is (E1p25c and its use in E5p24d). Therefore, we can form in principle an adequate idea of any modification of the body (E5p4) and of external bodies (E5p14).

These propositions together imply the principle of universal intelligibility: since we understand an effect through its cause (E1a2), and we have an adequate idea of God's essence (E2p47), and that everything follows from God's essence (E1p25) – including external causes – *there is no external cause of which we cannot form in principle an adequate idea*. Of course, the question remains whether we actually form adequate ideas, but this question is independent of whether we have epistemic access to the idea.

### **3. Ethical intellectualism: that from which we form an adequate idea cannot be harmful**

Spinoza states in E5p3 that as soon as we form an adequate idea of a passion – because something being a passion entails having an inadequate idea of it (E3d2–3) – it ceases to be passion and becomes an action. An action follows from our nature and therefore is always useful and not harmful (E4p38–39, E4app3, E4app6). What is harmful is evil and is the result of inadequate ideas (E4p64). These propositions together imply ethical intellectualism: whatever *we have an adequate idea of cannot be harmful* or evil.

If we take these three claims together the following conclusion seems inevitable: if an individual  $y$  has  $x$  as its cause of death,  $x$  has to be an external cause to  $y$  (by 1),  $y$  has to be able to form an adequate idea of  $x$  (by 2) and therefore  $y$  has to be able to turn  $x$  into an action (by 3) which rules out  $x$  as a cause of death (by 1). Since this can happen to any  $x$ , it is not necessary that  $y$  dies.

### **4. It is necessary that human individuals die**

The problem is that according to Spinoza humans necessarily die. That he wants to maintain this claim is evident from E2p10, which states that the being of substance does not pertain to the essence of man, and from E2p31c, which states that all particular things are contingent and corruptible. However, Spinoza does not state clearly that it is necessary to die. I assume that the general line of argument shows that in Spinoza's view, humans necessarily do die. I will consider the option that he thought that humans are contingent particulars capable of an indefinitely long life in section 6. His unwillingness to state the necessity of death might be the consequence of his views that the free man should not think about death (E4p67), and that the mind strives not to imagine those things that diminish the body's power of acting (E3p12–13). More probably, this can be the consequence of his

inability to prove the necessity of death with the resources of his system. He comes close to proving it in E4p4:

It is impossible that a man should not be a part of Nature, and that he should be able to undergo no changes except those which can be understood through his own nature alone, and of which he is the adequate cause. Dem. [...] if it were possible that a man could undergo no changes except those which can be understood through the man's nature alone, it would follow (by IIIp4 and P6) that he could not perish, but that necessarily he would always exist. [...]

Here Spinoza wants to demonstrate precisely the impossibility of the scenario presented at the end of the previous section, i.e.: that humans can avert death indefinitely. According to the demonstration humans could avert death either by their own power, or because it follows from the common order of nature as a kind of "accident". The refutation of both options is problematic.

The refutation of the second option goes as follows:

...if it were possible for a man to undergo no changes except those which could be understood through the man's nature alone, so that (as we have already shown) he would necessarily always exist, this would have to follow from God's infinite power; and consequently (by IP16) the order of the whole of Nature, insofar as it is conceived under the attributes of extension and thought, would have to be deduced from the necessity of the divine nature, insofar as it is considered to be affected with the idea of some man. And so (by IP21) it would follow that the man would be infinite (E4p4d).

The claim is that if someone would not have any passions because of the common order of nature, then the common order of nature would be deducible from her nature and therefore would be an infinite mode.

This argument has two shortcomings. First, E1p21 does not say that the whole order of nature can be deduced from infinite modes, even less so the claim that only from infinite modes can the whole order of nature be deduced. The entailment of the common order of nature by infinite modes is itself a hotly debated topic (cf. Garrett 1991; Curley–Walski 1999).

Second, this move implies that the practically impossible ideal of the free man who has only adequate ideas (E4p66cs) is also theoretically impossible. Since the free man has only adequate ideas he would be free of passions and would not be acted upon. This is problematic because given Spinoza's treatment of modal terms, impossibility implies either contradiction in essence (e.g. square circle), or unactualized possibility. Spinoza does not seem to allow for unactualized possibilities (E1p17c2s). But then the only option left is that the free man, our ethical ideal, is an inconsistent concept, like a square circle! But even if we accept that the free man is an unactualized possibility, he could hardly serve as an ethical ideal,

since he would be of a different metaphysical category. We humans are finite modes, while he would be an infinite mode.<sup>3</sup>

The first option, that a human being by her own power could avoid passivity, is refuted by appealing to E4p3 according to which “The force by which a man perseveres in existing is limited, and infinitely surpassed by the power of external causes”. The problem with this proposition is that its demonstration refers back only to E4a1 which, being an axiom, is not argued for, and simply states that however powerful a mode, if enough other modes join together they can be more powerful. So Spinoza is not able to provide systematically grounded reasons for ruling out this option.

There is a further concern: the proposition says that “it is impossible that a man [...] be able to undergo no changes except those which can be understood through his own nature alone, and of which he is the adequate cause” (E4p4). Making the claim this way only rules out the possibility that one is absolutely free, namely that one has no passions at all. However, it leaves open the possibility that while one always has some mild passions and is saddened here and there, one never suffers from such a dramatic external effect that would bring about her death. That is, even if Spinoza’s demonstration would be successful, the best he could demonstrate is that it is necessary that humans have passions.

## 5. Could Spinoza reject universal intelligibility or ethical intellectualism?

So far I have argued that Spinoza embraces claims that together imply that he cannot account for the necessity of death. Also, his account for the necessity of death suffers from three main shortcomings: (1) it does not demonstrate the necessity of death, only the necessity of passivity, (2) the refutation of the option that one does not die because of the common order of nature implies that the ideal of the free man is inconsistent, (3) the lack of refutation of one’s potential option to avert death indefinitely by way of one’s own power. I argue that this shows that accepting universal intelligibility and ethical intellectualism when combined with the identification of the logical and psychological aspects of the idea, results in Spinoza’s inability to explain the necessity of death. In this section I argue that both doctrines are central tenets of Spinoza’s philosophy and therefore he could not give them up without giving up Spinozism.

(1) Giving up ethical intellectualism would give us the following picture: an individual can know her cause of death, but knowing does not allow her to avert it. This view might

---

3 To be fair, this would be an intended result in the Medieval context, since there the ethical ideal is the conjunction with the active intellect (Black 1999) which has the same metaphysical category as the infinite modes in Spinoza. Yet, the free man and the active intellect play quite different ethical roles: in conjunction the subject becomes identical with the active intellect, while the free man as an ideal only helps to align our actions with our rational interests (Kisner 2010).

seem plausible given Spinoza's necessitarianism (cf. E4app32). Also, the common sense implausibility of the claim that pain ceases as soon as we form an adequate idea of it has been stated (Alanen 2012, 250). Given Spinoza's formulations, he could modify his system in two ways: either by claiming that the set of actions and adequate ideas are not coextensive, or by giving up the claim that actions are never harmful. The first would make the definition of action meaningless (E3d3), since then there would be no distinction between action, passion and affect. Also, this could not be reconciled with the parallelism doctrine (E2p7). In this doctrine, for every mode of thought there is an extended mode, with which it is identical. Therefore, for every adequate mode of thinking following from the essence of the human mind alone there will be an extended mode following from the essence of the human body alone. Because of the conatus doctrine, these extended modes cannot be harmful and thus have the same systematic role as action on the official theory. Therefore, the distinction between action and passion follows from parallelism, the conatus doctrine and the identification of epistemic value with epistemic autonomy. The second option, namely, admitting harmful actions, might have some plausibility given Spinoza's account of rational suicide and the wording of E4p59 (Nadler 2016). But this option entails giving up the conatus doctrine, since in this case something would follow from our nature alone that is harmful to us (Grey 2016). That is, Spinoza's ethical intellectualism is entailed by his conatus doctrine, his parallelism doctrine and his identification of epistemic value with epistemic autonomy.

(2) Spinoza could give up universal intelligibility. This would give us the following picture: an individual can know everything except her cause of death. Spinoza could argue for this restriction of universal intelligibility in two ways: first, by restricting the scope of possible objects of understanding; and second, by restricting the time-frame of understanding. The first option would be that some objects are by definition unintelligible for the subject. A good candidate for such a restriction would be objects of the idea which excludes the existence of the subject's body. The second option might be called the Epicurean solution: death is such a violent disruption of the essence of the body that we do not have time to understand it; once it is here we are gone.

The problem with these proposals is twofold. First, they are hard to reconcile with the textual evidence. E5p4 states that we can form adequate ideas of every modification of the body. Since death was defined as a disruption of the essence of the body, it qualifies as a modification of the body. Also, in E4p59 Spinoza states that "To every action to which we are determined from an affect which is a passion, we can be determined by reason, without that affect". Here action could not mean the technical term "what follows from our nature alone". Rather, it must refer generally to an event happening in or outside of us (cf. E3d2). Since death is a bodily modification, it is an action in this sense and therefore we can be motivated to it by reason, i.e. adequate ideas. (This is also compatible with Spinoza's analysis of different types of suicide, see: E4p20s.)

Second, there is the more general problem that restricting intelligibility would introduce a brute bifurcation in the system. Although the exact role of the Principle of Sufficient Reason in Spinoza is debated, the debate is not about the question of whether Spinoza admitted a brute bifurcation in his system, rather, about what qualifies as a brute bifurca-

tion (Della Rocca 2008, 2010; Renz 2010; Melamed 2010, 2011, 2012, 2013). Restricting the scope of intelligibility this way could qualify as a brute bifurcation and would certainly disqualify Spinoza from being a rationalist (Renz 2010, 14–15).

To sum up, in this section I have argued that Spinoza's inability to account for the necessity of death comes from his acceptance of universal intelligibility and ethical intellectualism, together with the identification of the logical and psychological aspects of the idea. In my view, the doctrine of universal intelligibility is so fundamental to Spinoza's project, that giving it up would entail giving up Spinozism. Ethical intellectualism, especially when coupled with the identification of the logical and psychological aspects of the idea, does not seem to be either very plausible or necessary for Spinozism. In fact, as I have shown in the introduction, Sangiacomo has argued that it is not even part of the philosophy of a mature Spinoza. While it is true that the intellectual aspect in Spinoza's mature thinking is affective and not purely conceptual, in this section I have argued that giving up ethical intellectualism is only possible by abandoning either the conatus doctrine, or the identification of epistemic value with epistemic autonomy. And both are fundamental tenets of Spinozism.

## 6. Analogous problems in the medieval philosophy of intellect

The relationship of intellect and imagination – that is, the relationship of epistemological values and the psychological states in which they were embodied – was also a relevant problem for the Arabic and Hebrew philosophy that partially constituted Spinoza's philosophical context. These philosophers – Maimonides (Nadler 2014), Gersonides (Klein 2003; Melamed forthcoming; Harvey 2012; Klein 2014), Shem Tov ben Shem Tov (Adler 2014), Elijah Del Medigo (Licata 2013; Fraenkel 2013, 2011) – tried to solve the problem generated by the seemingly inconsistent claims of Aristotle's *De anima*.<sup>4</sup> There Aristotle distinguished two types of intellect: the active and the material intellect. The material intellect does not have a nature but can become anything (DA III.5 430a10–15). That is, the material intellect can be informed by every form and thus it can potentially understand anything. In contrast, the active intellect is distinct, unaffected, unmixed and in essence, activity: the giver of those forms which inform the material intellect (DA III.5 430a15–20). Aristotle elsewhere also claimed that unqualified intellect alone can survive death (DA I.4. 408b20, II.1. 413a5–10, III.5. 430a20–25), and that it is unmixed and unaffected (DA III.4 429a18–20).

There are two ways in which these claims can naturally be understood (Davidson 1992). The first way was proposed by Alexander of Aphrodisias (Aphrodisias 2014a, 2014b). He focused on the metaphysical implications of Aristotle's claim about the unmixed nature of the intellect. In Aristotle's physics, if something is unmixed with and distinct from matter,

---

<sup>4</sup> References to Aristotle's work is to this edition with the usual abbreviation (Aristotle 1993).

it is an incorporeal and eternal form. Given its eternity, this interpretation goes along well with Aristotle's claim that this is what survives death. However, it faces the problem that in Aristotle's metaphysics, instances of the same species are individuated by matter. Given that this eternal intellect is unmixed, it cannot be individuated, thus it is one of a kind. But there are multiple epistemic subjects in the universe, and something has to distinguish their mental operations. Therefore, in Alexander's view, the intellect that is unmixed and eternal is the active intellect only, which he identifies with God. And the material intellect, which does not have a nature, is in his view, the disposition of the corruptible soul to accept the intellectual forms provided by the active intellect.

The alternative view was proposed by Themistius (1990). In Alexander's theory the intellect that remains after death is the active intellect, which is actually God, and therefore the human soul is corruptible. Themistius wanted to avoid this conclusion and therefore he focused on Aristotle's claim that the unqualified intellect is unmixed and therefore eternal. This solves one problem: if the material intellect, as well as the agent intellect is eternal, then everyone has a unique immortal soul. But then he has to explain how is potentiality possible in an eternal substance and what individuates the numerically different intellects.

These two interpretations also have bearings on the question of the mind's ability to understand. Aristotle clearly linked the unmixed character of the intellect to its ability to understand everything. In Aristotle's theory of knowledge, understanding something is to become identical with it: just as the eye becomes red when perceiving red, the intellect instantiates the (intellectual) form of apple when understanding the apple. Therefore, in the same way as the eye must lack color in order to be able to perceive all colors, the intellect cannot have any form in order to be able to know all things, i.e. to acquire any possible form. The problem is that neither of the two interpretations can account for this formless pure potentiality, which constitutes the problem of the material intellect that generated much of the discussion in the scholastic philosophy of mind, in both Arabic (Davidson 1986) and Hebrew (Visi 2012). Themistius cannot explain how something purely potential can be eternal, while Alexandros cannot explain how a disposition of a corruptible being can lack nature.

I have argued elsewhere that Spinoza's distinction of intellect and imagination was influenced by his Medieval predecessors (Tóth 2016a, 2016b). Here I would like to argue that Spinoza's problem stems from a problem similar to the one of the material intellect. Although, as we have seen, he tried to solve the difficulties of his early ethical intellectualism by turning concepts into affects (Lenz 2013); this did not rule out ethical intellectualism, but only gave it an affective twist.

Originally, the problem was that Spinoza identified intellect and will, and therefore he could not explain error: as we have seen in section 2, everyone has the source of knowledge (i.e. the idea of the essence of God) which is sufficient for knowing everything. So why are people mistaken? Why are we not omniscient? This problem was solved by turning ideas into affects: their epistemic status, i.e. their role in our web of belief and our reasons for action, became a function of their power. But this solution came at a high price. Now Spinoza could explain the source of error with the power of ideas, i.e. with affectivity: the

powerful inadequate idea can defeat the less powerful adequate idea. However, he could not explain how we can understand everything. Since ideas ceased to be mere representations, a class of ideas became the wrong kind of ideas: those ideas that represent harmful objects are just the wrong kind of idea to have. And those ideas that represent lethal objects are just the kind of ideas that one cannot have.

Spinoza accepts that there is a conformity of causes and effects: the same object produces the same kind of effect in the same subject with the same constitution. The same music will always be delightful to the same subject, unless the constitution of the subject has changed, e.g. she begins to mourn (E4prae).

Spinoza also accepts that causes can be beneficial or harmful depending on their effects on the human body. This is, of course, to a large extent determined by the current state of the human body. That is, based on how the human body is constituted, the potential causes can be categorized either as harmful, or as beneficial.

Some of the harmful causes produce such effects that they exclude the existence of the human body. These are the lethal effects. Perhaps the lethal nature of some of these effects are contingent on the condition of the human body, and therefore can turn into beneficial effects by appropriately modifying the state of the human body. This could happen in a similar manner to the change that occurs when the mourning person turns into a melancholy person and therefore the music that was previously harmful turns into music that is beneficial.

One might argue that all of the lethal effects are such. One way in which this change in the body's constitution may come about, is by acquiring knowledge. Thus, one could say that with knowledge, all lethal effects can be mitigated. The problem with this reading is that it implies that we do not necessarily have to die. Also, it seems implausible that there are no effects that are actually contrary to human nature: when the big fish eats the small fish (TTP 16.2), it is not the result of a terrible misunderstanding; the big fish has a nature that is actually contrary to the nature of the small fish.

To be fair, it is not obvious that this option was all that counter-intuitive for Spinoza. As we have seen in section 4, he might not have stated explicitly that human individuals necessarily die because he really thought that all lethal actions can be mitigated by knowledge. He might have embraced the claim that in a perfect world, big fish do not eat small fish and humans do not necessarily die.<sup>5</sup> In religious thinking neither of these claims is unheard of. Maimonides asserts in a much-discussed place of his *Guide* that intellectual knowledge saves the sage even on the battlefield.

If man frees his thoughts from worldly matters, obtains a knowledge of God in the right way, and rejoices in that knowledge, it is impossible that any kind of evil should befall him while he is with God, and God with him. When he does not meditate on God, when he is separated from God, then God is also separated from him; then he is exposed to any evil that might befall him; for it is only that intellectual link with

---

5 I would like to thank Ursula Renz for raising this objection.

God that secures the presence of Providence and protection from evil accidents. [...] if you should happen to pass on your way a widely extended field of battle and even if one thousand were killed on your left and ten thousand on your right, no evil at all would befall you (Maimonides 1974 III. 51; cf. Nadler 2014).

Also, in the Scripture, end times are often characterized by the metaphor of carnivores and herbivores living together peacefully. E.g.:

The wolf also shall dwell with the lamb, and the leopard shall lie down with the kid; and the calf and the young lion and the fatling together; and a little child shall lead them. And the cow and the bear shall feed; their young ones shall lie down together: and the lion shall eat straw like the ox (Isiah 11:6–7).

Therefore it might be the case that Spinoza took these claims literally and developed a philosophy according to which acquiring knowledge really has such wonderful consequences. But in my view this is not the case: since contrariety and agreement in nature correlate with similarity and dissimilarity, the cow and the bear should cease to be two different species in order to live peacefully together (cf. Wilson 2002).

The other option is to accept that some lethal effects are generated by causes that will always generate lethal effects in humans no matter what. But then the human body will always strive against these effects and the human mind will strive against having their ideas. And since we know external bodies by the ideas of their effects (cf. Lenz 2012), the human mind will strive not to know them. That is, the fact that the human body has a particular nature – on the basis of which natures with beneficial effects and natures with harmful effects on the human body can be distinguished – makes it impossible for the human mind to acquire ideas of those natures which have lethal effect on the human body. In fact, the human mind will do everything in its power not to know them.

I claim that this problem is analogous with the problem of the material intellect. In scholastic philosophy, the material intellect had to be free of any nature in order to be able to know everything; and then, no one was able to account for this pure potentiality. In Spinoza's philosophy, the mind had to be able to conform to all natures in order to know them, but only something devoid of nature can conform to any nature. In both cases the determinate nature of the subject precludes universal intelligibility. This problem, because of the identification of epistemology and affectivity, manifests itself in Spinoza's system through his inability to demonstrate the necessity of death.

## References

- Adler, Jacob. 2014. Mortality of the Soul from Alexander of Aphrodisias to Spinoza. In: Nadler, Steven (ed.): *Spinoza and Medieval Jewish Philosophy*. Cambridge: Cambridge University Press. 13–35.

- Alanen, Lilli. 2012. Spinoza on Passions and Self-Knowledge: The Case of Pride. In: Pickavé, Martin – Shapiro, Lisa (ed.): *Emotion and Cognitive Life in Medieval and Early Modern Philosophy*. Oxford: Oxford University Press. 234–254.
- Aphrodisias, Alexander of. 2014a. *Alexander of Aphrodisias: On the Soul. Part I: Soul as Form of the Body, Parts of the Soul, Nourishment, and Perception*. Translated by Victor Caston. Reprint edition. London: Bristol Classical Press.
- Aphrodisias, Alexander of. 2014b. *Alexander of Aphrodisias: Supplement to On the Soul*. New York: Bloomsbury Publishing.
- Aristotle. 1993. *De Anima Books II and III*. Translated by D. W. Hamlyn. New York: Oxford University Press.
- Bennett, Jonathan. 1984. *A Study of Spinoza's Ethics*. Indianapolis: Hackett.
- Black, Deborah L. 1999. Conjunction and the Identity of Knower and Known in Averroes. Edited by Robert E. Wood. *American Catholic Philosophical Quarterly* 73 (1): 159–184. (doi:10.5840/acpq19997317).
- Boros, Gábor. 1997. *Spinoza és a filozófiai etika problémája*. Budapest: Atlantisz.
- Curley, Edwin. 1988. *Behind the Geometrical Method: A Reading of Spinoza's Ethics*. Princeton: Princeton University Press.
- Curley, Edwin – Walski, Gregory. 1999. Spinoza's Necessitarianism Reconsidered. In: Gennaro, Rocco – Huenemann, Charles (ed.): *New Essays on the Rationalists*. New York: Oxford University Press. 241–262.
- Davidson, Herbert A. 1986. Averroes on the Material Intellect. *Viator* 17 (January): 91–138.
- Davidson, Herbert A. 1992. *Alfarabi, Avicenna, and Averroes, on Intellect: Their Cosmologies, Theories of the Active Intellect and Theories of Human Intellect*. New York: Oxford University Press.
- Della Rocca, Michael. 1996. *Representation and the Mind-Body Problem in Spinoza*. New York: Oxford University Press.
- Della Rocca, Michael. 2003. The Power of an Idea: Spinoza's Critique of Pure Will. *Nous* 37 (2): 200–231. (doi:10.1111/1468-0068.00436).
- Della Rocca, Michael. 2008. *Spinoza*. New York: Routledge.
- Della Rocca, Michael. 2010. PSR. *Philosophers' Imprint* 10 (7): 1–13.
- Fraenkel, Carlos. 2011. Spinoza on Philosophy and Religion: The Averroistic Sources. In: Fraenkel, Carlos – Perinetti, Dario – Smith, Justin E. H. (eds.): *The Rationalists: Between Tradition and Innovation*. Dordrecht: Springer Netherlands. 27–43.
- Fraenkel, Carlos. 2013. Reconsidering the Case of Elijah Delmedigo's Averroism and Its Impact on Spinoza. In: Akasoy, Anna – Giglioli, Guido (eds.): *Renaissance Averroism and Its Aftermath: Arabic Philosophy in Early Modern Europe*. Dordrecht: Springer Netherlands. 213–236.
- Garrett, Don. 1991. Spinoza's Necessitarianism. In: Yovel, Yirmiyahu (ed.): *God and Nature. Spinoza's Metaphysics*. Leiden: Brill. 191–218.
- Grey, John. 2016. Reply to Nadler: Spinoza and the Metaphysics of Suicide. *British Journal for the History of Philosophy* 25 (2): 380–388. (doi:10.1080/09608788.2016.1230539).
- Harvey, Warren Zev. 2012. Gersonides and Spinoza on Conatus. *Aleph* 12 (2): 273–297.

- Kisner, Matthew J. 2010. Reconsidering Spinoza's Free Man: The Model of Human Nature. In: Garber, Daniel – Nadler, Steven (eds.): *Oxford Studies in Early Modern Philosophy*. Vol. V. Oxford: Oxford University Press.
- Klein, Julie R. 2003. Spinoza's Debt to Gersonides. *Graduate Faculty Philosophy Journal* 24 (1): 19–43.
- Klein, Julie R. 2014. "Something of It Remains": Spinoza and Gersonides on Intellectual Eternity. In: Nadler, Steven (ed.): *Spinoza and Jewish Philosophy*. Cambridge: Cambridge University Press. 177–203.
- Lenz, Martin. 2012. Intentionality without Objectivity? Spinoza's Theory of Intentionality. In: Salice, Alessandro (ed.): *Intentionality*. Munchen: Philosophia Verlag. 29–58.
- Lenz, Martin. 2013. Ideas as Thick Beliefs: Spinoza on the Normativity of Ideas. In: Lenz, Martin – Waldow, Anik (eds.): *Contemporary Perspectives on Early Modern Philosophy: Nature and Norms in Thought*. Dordrecht: Springer. 37–50.
- Licata, Giovanni. 2013. *La via della ragione: Elia del Medigo e l'averroismo di Spinoza*. Macerata: Edizioni Università di Macerata.
- Maimonides, Moses. 1974. *The Guide of the Perplexed*. Vol. 1. Translated by Shlomo Pines. Chicago: University of Chicago Press.
- Malinowski-Charles, Syliane. 2009. Emotions as the Engines of Spinoza's Ethics. *Kritika & Kontext* 38–39: 120–125.
- Marshall, Eugene. 2014. *The Spiritual Automaton. Spinoza's Science of the Mind*. New York: Oxford University Press.
- Melamed, Yitzhak Y. (forthcoming). Gersonides and Spinoza on God's Knowledge of Universals and Particulars. In: Freudenthal, Gad – Wirmer, David – Elijor, Ofer (eds.): *Gersonides Through the Ages*.
- Melamed, Yitzhak Y. 2010. Acosmism of Weak Individuals? Hegel, Spinoza, and the Reality of the Finite. *Journal of the History of Philosophy* 48 (1): 77–92.
- Melamed, Yitzhak Y. 2011. Why Spinoza Is Not an Eleatic Monist (Or Why Diversity Exists). In: Goff, Philip (ed.): *Spinoza on Monism*. New York: Palgrave Macmillan. 206–222.
- Melamed, Yitzhak Y. 2012. The Sirens of Elea: Rationalism, Monism and Idealism in Spinoza. In: Lolordo, Antonia – Stewart, Duncan (eds.): *Debates in Early Modern Philosophy*. London – New York: Blackwell.
- Melamed, Yitzhak Y. 2013. *Spinoza's Metaphysics: Substance and Thought*. New York: Oxford University Press.
- Nadler, Steven. 2001. *Spinoza's Heresy*. Oxford University Press.
- Nadler, Steven. 2014. Virtue, Reason, and Moral Luck: Maimonides, Gersonides, Spinoza. In: Nadler, Steven (ed.): *Spinoza and Jewish Philosophy*. Cambridge: Cambridge University Press. 152–176.
- Nadler, Steven. 2016. Spinoza on Lying and Suicide. *British Journal for the History of Philosophy* 24 (2): 257–278. (doi:10.1080/09608788.2015.1084491).
- Renz, Ursula. (forthcoming). Spinoza's Epistemology. In: Garrett, Don (ed.): *The Cambridge Companion to Spinoza*. 2nd ed. Cambridge: Cambridge University Press.
- Renz, Ursula. 2010. *Die Erklärbarkeit von Erfahrung*. Frankfurt am Main: Vittorio Klostermann.

- Renz, Ursula. 2015. From the Passive to the Active Intellect. In: Yitzhak, Melamed Y.: *The Young Spinoza: A Metaphysician in the Making*. Oxford: Oxford University Press. 287–299.
- Sangiaco, Andrea. 2013. What Are Human Beings? Essences and Aptitudes in Spinoza's Anthropology. *Journal of Early Modern Studies* 2 (2): 169–190.
- Sangiaco, Andrea. 2015a. Teleology and Agreement in Nature. In: Campos, Andre Santos (ed.): *Spinoza: Basic Concepts*. Exeter: Imprint Academic.
- Sangiaco, Andrea. 2015b. The Ontology of Determination: From Descartes to Spinoza. *Science in Context* 28 (4): 515–543.
- Sangiaco, Andrea. 2015c. Fixing Descartes: Ethical Intellectualism in Spinoza's Early Writings. *The Southern Journal of Philosophy* 53 (3): 338–61. (doi:10.1111/sjp.12113).
- Spinoza, Benedictus de. 1988. *The Collected Works of Spinoza*. Vol. I. Edited and translated by Edwin Curley. 2nd ed. Princeton: Princeton University Press.
- Spinoza, Benedictus de. 2016. *The Collected Works of Spinoza*. Vol. II. Translated by Edwin Curley. Princeton, N.J.: Princeton University Press.
- Themistius. 1990. Paraphrase of Aristotle De Anima 3.4–8. In: *Two Greek Aristotelian Commentators on the Intellect*. Translated by Frederick M. Schroeder and Robert B. Todd. Toronto: Pontifical Institute of Mediaeval Studies. 77–134.
- Tóth, Olivér István. 2016a. “az értelem nincs olyképpen alávétve a véletlennek, mint a test” – Spinoza az értelem szabadságáról egy középkori vita kontextusában. *Magyar Filozófiai Szemle* 60 (2): 83–99.
- Tóth, Olivér István. 2016b. Inherence of False Beliefs in Spinoza's Ethics. *Society and Politics* 10 (2): 74–94.
- Visi, Tamás. 2012. A szférák zenéje és az angyalok éneke. In: Takó, Ferenc – Tóth, Olivér István (szerk.): „...de van benne rendszer” *Tanulmányok az Eötvös Collegium Filozófia Műhelye fennállásának 15. évfordulójára*. Budapest: Eötvös Collegium, Filozófia Műhely. 164–184.
- Wilson, Margaret D. 2002. For They Do Not Agree with Us. In: Gennaro, Rocco – Huenemann, Charles (ed.): *New Essays on the Rationalists*. New York: Oxford University Press. 336–352.

# XI

*Brian Glenney*

## Spinoza on Misperception

In a rather astounding passage, Spinoza considers misperceiving the sun as near<sup>1</sup>:

For example, when we look at the sun, we imagine it to be about two hundred feet away from us. In this we are misled so long as we remain ignorant of its true distance. But when its distance is learned, the error is removed, not the imagination, that is, the idea of the sun that explains its nature only insofar as the body is affected by it. And so, although we come to know its true distance, we shall nevertheless imagine it as near to us (E4p1s).<sup>2</sup>

Learning that the sun is distant fails to revise our misperception that the sun is near. How so? As Spinoza continues, “[W]e do not imagine the sun to be near just because we are ignorant of its true distance but because the mind conceives the sun’s size insofar as the body is affected by the sun” (E4p1s). For Spinoza, we perceive objects like the sun as they affect the perceiver’s body (see also E2p16–17, 26). Insofar as the body is concerned, the sun *is* near.

The case of the misperceived yet persisting near sun is particularly beguiling for understanding Spinoza who claims further that the faculty of imagination through which we perceive the world as we know it (E17ps) is also error free (E2p17s). All ideas, including the near sun misperception, are modes of a perfectly true God who admits of no error. Yet, as Wilson (1985) correctly notes of the misperceived near sun, “But, one is tempted to insist, *this* idea is false, period” (Wilson 1985, 110). Thus, one is not right to misperceive the sun as near. Other

---

1 Whether this is a reference to Descartes mention of the limits of seeing distance (see Garrett 2013) or to Aristotle’s discussion of seeing the sun as small (*De Anima* 428b3–4) as a case of an inaccurate common sensible, or based on Spinoza’s own experience is not clear.

2 References to Spinoza use the translation of Curley (1994) and indicate the text: (E) for *Ethics* and (TIE) for *Treatise on the Emendation of the Intellect*, then part or section and number of the proposition (p), corollary (c), scholium (s), demonstration (dem), definition (d), appendix (app), preface (praef), or axiom (a).

commentators, bemused by Spinoza's comments, have tried other ways to piece together this near sun case (Bennett 1984; LeBuffe 2009, 2010; Steinburg 2009; Lenz 2013). In this chapter, I offer another attempt that considers the role of doubt in cases of misperception.

I begin by discussing separately the persistence of misperception and the Spinoza's claim to perfect perception, focusing on some sharp critiques by Bennett (1984) and Shapiro (2012a, 2012b) who leave Spinoza with a non-doxastic account of perception. I respond, following Lenz (2013), that Spinoza's account of perception is doxastic and develop an account that allows for belief revision based on the critical role of "doubt", one that is particularly significant in Spinoza's early work, TIE.

## Perceptual Persistence

Spinoza considers his near sun example of persistent and true misperception to be generalized across all cases of perception. In the same passage as quoted above, he continues:

Thus, when the rays of the sun falling upon the surface of the water are reflected toward our eyes, we imagine it just as if it were in the water, even if we know its true place. And so it is with the rest of the imaginations by which the mind is misled, whether they indicate the natural constitution of the body or that its power of acting is increased or diminished. They are not contrary to the true, and they do not disappear in its presence (E4p1s).<sup>3</sup>

This passage is a confounding account of perception for numerous reasons, but two stand out: 1) misperception *persists* and 2) misperception is an error that can be removed, but is itself "not contrary to the true". Let me consider these in turn, beginning with the persistence of misperception.

Rather than considering the sun's perceived nearness or its perceived mislocation on the water's surface, both of which seem less convincing today, consider the persisting misperception of geocentric "sunrises" and "sunsets" by sky gazers. Perceiving the sun as moving is a persuasive persisting experience even though a sky gazer may understand and believe the sun to neither rise nor set. Even when a Copernican-believing sky gazer attends to the fact that the earth orbits the sun and rotates on its axis, they may not be able to help but see the sun as rising (Sommers 2009). The moon illusion is a similarly persisting experience. We know that the moon neither changes in size nor distance in its ascent from horizon to zenith, yet we cannot but see the moon at the horizon as larger than when at its zenith (Ross-Plug 2002). The Müller-Lyer illusion of two parallel lines with opposite facing arrows produces a similar experience of different sizes even as we measure their equivalent sizes and behave as if they

---

<sup>3</sup> Spinoza cites this problem of perception at least one other time in the Ethics (E35p1s) and twice in *Treatise on the Emendation of the Intellect* (TIE: 21 and 30).

were equivalent (Carey 2001). These are all cases of persisting misperception – cases of non-veridical sensory experiences that endure in the midst of incompatible beliefs and knowledge.

Lenz (2013) discusses a similar parallel between Spinoza's misperceived nearness of the sun case and the Müller-Lyer illusion as evidence for his view that, for Spinoza, beliefs are "thick": they include normative evaluative and emotional content patterned on our past experiences and the customs of society. We have attitudes about the sun that influence our behavior; the sun is very hot – enough to burn us – and thus we have a propensity to shade ourselves. It is attitudes like these, likely, that provide the basis for the sun's perceived nearness, attitudes that Lenz likens to Gibsonian affordances, relational properties that dispose us to act in relation to our bodily survival.

We find support for Lenz's view of a normative basis of perception in Spinoza's account of *conatus*, where all the various modes of being human are in service to a subject's striving toward a state of preservation (E2p26). The misperception of the sun's nearness persists in spite of its conflict with the intellectual knowledge of the sun's distance because the affective force of the sun's heat overpowers any intellectually grounded belief that it is in fact distant. Hence, for Spinoza, perception is something that has emotional strength. As Lisa Shapiro, who emphasizes this emotional content by calling Spinoza's account "passionate perception" writes, "A sensation of an object ... requires an emotion as a way of conceiving *that* object at all" (Shapiro 2012b, 214). If sensory perception of the sun is "passionate", then emotional and sensory content motivate behavior towards a subject's wellbeing.

A contemporary analysis of persisting perceptual experience in the midst of contrary belief is called "cognitive impenetrability", and can further help account for Spinoza's claim to persistence of misperceptions, where one's beliefs do not influence their experiences (Pylyshyn 1999). It might be said that our perceptions of sunrises and sunsets themselves are impervious to revision, likewise the sun's perceived nearness and mislocation on the water's surface, because the intellectual beliefs cannot "penetrate" the perception. In fact, it might be said that Spinoza is the first to anticipate an account of impenetrable experience with his general claim to persisting misperceptions.

Thickening the content of belief (Lenz 2013) or emphasizing its passionate content (Shapiro 2012b), however, does not fully resolve Spinoza's account of the near sun case. For, on both views, there would be a close relationship between one's perception and behavioral propensity. Garrett (2013), another commentator who shares a view similar to that of Lenz and Shapiro, writes that the perceived nearness of the sun, "will guide one to act *as though* it were 200 feet distant [...] Suppose, for example, one desires to alter the sun in some respect... one will be guided to look for materials to build a 200-foot ladder... or to build a cannon with a 200-foot range". But as the sun is not 200 feet away, this case of misperception would upend the relationship between perception and behavior and the perceiver's state of preservation or health. If the sun *is* near, then the behavioral propensities for both increasing a perceiver's warmth or avoiding heat would involve moving short distances toward and away from the sun. The perceiver would quickly change their behavior once their anticipated warmth and cooling failed. It is not clear how a theory of thick or passionate perception, which seems to entail unrevisable belief, can account for why and when a sub-

ject misperceives an object in a manner that undermines while still directly influences their *conatus*, maintaining its fundamental aims of a subject's well being. This worry puts strain on the second part of Spinoza's account of perception, that the imagination as 'error free'.

## Perceptual Perfection

For Spinoza, the imagination itself is understood to be error free:

The mind does not err from the fact that it imagines, but only insofar as it is considered to lack an idea which excludes the existence of those things which it imagines to be present to it. For if the mind, while it imagined non-existent things as present to it, at the same time knew that those things really did not exist, it would of course attribute this power of imagining to a virtue of its nature, not to a vice [...] (p17cs).

Spinoza concludes that the misperception of the sun as near is a "virtue". How? We sidestep error in perception by attributing error to the intellect. In particular, Spinoza explains that it is an error of 'lacking an idea', or a lack of knowledge (p17cs), what he calls elsewhere a "privation" (E35p).<sup>4</sup> Before discussing this claim in greater detail, we can explain Spinoza's argument here for an error-free imagination, and thereby perception, as follows:

- 1) We perceive non-existent things (e.g. a near sun).
- 2) Belief in non-existent things is an error.
- 3) Errors are part of a lack of knowledge, not perception.
- 4) Thus, perception itself does not err.

In all four of Spinoza's discussions of misperceiving the sun as near (E2p1s, E2p26, TIE 21, TIE 30), involving both his early and later periods,<sup>5</sup> Spinoza presents the case of error as a lack of knowledge.<sup>6</sup> Consider this most revealing discussion in TIE 30:

---

4 I read both LeBuffe (2009, 2010) and Steinberg (2009, 164f) as holding this basic view of Spinoza's account of error.

5 There is noted controversy about how much and what is distinctive about Spinoza's early and later work (Renz 2014; Sangiacomo 2015). Thus, it is significant that Spinoza uses the same 'near sun' example in both for, as I assume here, similar purposes of exploring the possibility of perceptual error. This assumption of coherence is restricted to Spinoza's use of these examples and is given some support by the fact that in both early and later uses of this example in particular Spinoza views the intellect as "active", contrary to Renz (2014) who argues that Spinoza had a "passive" understanding of the intellect in the early period.

6 In his commentary on Spinoza, Jonathan Bennett finds premise (3) of Spinoza's argument problematic: "he treats an error as *consisting in* a lack of knowledge rather than as *arising from* it, and it seems clear that he is wrong in this" (Bennett 1984, 169). Bennett's analysis seems correct, but such a consideration would take us too far afield from the central question of error as privation.

If someone has never been led, either by experience or by anything else, to think about the deceptiveness of the senses, he will never doubt whether the sun is larger or smaller than it appears to be. So Country People are generally surprised when they hear that the sun is much larger than the earth (TIE, 30).

It is neither some failed mystic cosmology that causes one to misperceive a near sun, nor some false understanding of senses, but rather the naivety of the “Country People” who “never reflected” nor “never doubt”, making their ignorance – their lack of knowledge – the problem, not perception itself.<sup>7</sup> On this view, however, Spinoza is still committed to the claim that the ‘country people’ falsely believe themselves to be in the presence of a near sun when in fact they are in the presence of a distant sun; misinformed or uninformed, they are in a state of false belief. But, this cannot be if the mind is “error free”.

Bennett (1984) argues that Spinoza’s theory of perceptual error can be saved if premise (2) is rejected and perception is non-doxastic; “‘imaginings’ really do seem to be sensory states and nothing like beliefs” (171). Analogous examples of the geocentric experiences or the impenetrable experiences of the Moon and Müller-Lyer illusions support this reading of Spinoza’s theory of perceptual error because these misperceptions are contrary to the beliefs that we hold of the objects. Again, we believe the lines and the moon to be the same size, though our perceptual experience tells us otherwise. Sensory states, for Bennett, have no doxastic purchase and are as such, ‘nothings’. If we accept Spinoza’s third premise that privation is the basis of error, then error is itself a privation, a nothing. “I submit”, writes Bennett, “that Spinoza wants to ensure that ‘not P’ does not occur in his description of a universe in which P is the case” (172).

On Bennett’s account of Spinoza’s theory of perceptual error, however, perception has nothing to do with external reality – a literal nothing. Thus, it is not clear whether Spinoza’s account of perception can itself have *any* cognitive purchase on reality. Perception, on Bennett’s view of Spinoza, gives up too much in being more of an imaginative exercise, a fictional universe of suns that you can reach out and touch or even bathe with. This seems incompatible with Spinoza’s account of perception, which is fundamentally behavioral – in support of *conatus*. Error *should* matter for an account of perception that forms the basis for the body’s welfare and preservation – error can not be nothing. This acknowledgement weakens Bennett’s interpretation of Spinoza and motivates a closer examination of Shapiro’s account of “passionate” perception.

---

7 Note that while privation may in fact account for near sun misperceptions, and that this is likely Spinoza’s own view, it seems perfectly possible that a misinformed but highly reflective account of cosmology *could also* share in this misperception. The “New Realism” school of thought is a modern contender for affirming the existence of a near sun. Led by Edwin Bissell Holt (1912), the New Realists argued that for every perceived appearance, there exists a matching object.

# Passionate Perception

When we imagine the pains or joys of those about us, or characters in a good story, we become emotionally invested. Similarly, Shapiro (2012a) argues, when we experience an object, like the seeming proximity of the noonday sun, we perceive with emotional investment – with affect and relevant motivational states and strivings toward our preservation. “Being affected involves apprehending a thing as impacting our continued existence, and that intrinsically involves imagining the object of our affection, that is, it intrinsically involves imaginative content” (Shapiro 2012b, 211). Passages such as E3p12dem, E2p7, and E2p17s suggest that this is Spinoza’s own view. For instance, Spinoza writes:

If the human body is affected with a mode that involves the nature of an external body, the human mind will regard the same external body as actually existing, or as present to it, until the body is affected by an affect that excludes the existence or presence of that body (E2p17).

Such passages show the imagination, including perception, to be itself a kind of affect. This seems particularly evident in Spinoza’s discussion of misperceiving the sun’s distances as near, a conclusion that, “so long as the mind imagines things that increase or aid our body’s power of acting, the body is affected with modes that increase or aid its power of acting” (E3p12dem). Shapiro concludes, “It is not the case that we first take things to exist and then find ourselves affected by them. Rather we take as existing the things we do because of how we are affected” (2012b, 97). Like Lenz (2013), Shapiro’s passionate view places emphasis on behavioral correlations to perception. Thus, the imaginings are far from Bennett’s ‘nothings’, as they contribute to the foundational nature of mind, the *conatus*. Let’s consider this option in terms of Spinoza’s own example of perceiving a near sun as a case of persisting incorrigible and error-free imaginative perception.

## Passionate Misperception

We see a midday sun as almost near enough to reach out and touch, though we know, or can come to know, it must be quite far from reach. In other words, our intellect corrects our perception of the sun as near even though, as Spinoza notes, we continue to see the sun as near. What can account for such a commanding misperception? Shapiro’s interpretation of Spinoza is insightful; passion rather than sensation shapes sensory experience. If the sun’s affection is the basis for our perception of the sun, then the various aspects relative to this affection would be primary in our sensory experience of the sun.<sup>8</sup> This may be what

---

<sup>8</sup> Note that some affections are “neutral”, neither increasing nor decreasing the power to act. See Shapiro 2012b.

Spinoza means by writing, “For as we have said in 2p35s, we do not imagine the sun to be near just because we are ignorant of its true distance but because the mind conceives the sun’s size insofar as the body is affected by the sun” (E2p1s). This seems intuitive. For instance, given the sun’s extreme heat and light, it appears to be more proximate to our body than other familiar sources of heat and light: fires, candles, light bulbs, etc., even though its distance is so great. The sun’s affection *is* proximate, providing a basis for how it is that Spinoza’s account can explain persistent misperceptions of objects as error free.

Yet, passionate perception, like Bennett’s account, views persisting misperceptions without doxastic and epistemic import. In fact, Shapiro’s account may be worse off than Bennett’s, justifying persisting misperceptions of objects in the world that have crucial behavioral importance resulting in moth-to-the-flame type problems. Again, if a ‘country person’, misperceiving the near proximity of the sun becomes cold, they may run toward it in a hopeless quest for greater heat, or if overheating they might haplessly run away to cool off. Misperception generates behavior that may significantly harm one’s body, resulting in a failure of the intended purpose of the senses for the body’s welfare. Shapiro neither recognizes nor discusses this predicament for her interpretation of Spinoza’s account of perception, but it undermines its plausibility significantly to my mind.

## Doubtful Passionate Misperception

Consider an alternative account centered on the context for the case of near sun misperception in Spinoza’s earlier period.<sup>9</sup> In the passage from TIE 30 quoted above, Spinoza suggests that misperception generates a tendency toward belief revision that most suits successful striving – that benefits one’s *conatus*. Spinoza continues as follows, “And if someone, after doubting, acquires a true knowledge of the senses and of how, by their means, things at a distance are presented, then the doubt is again removed” (TIE: 30). What the tendency toward belief revision motivates is unclear, but seems to be tied to acquiring ‘a true knowledge’ or a more adequate idea, i.e. questioning, investigating, and gaining understanding. Spinoza seems to argue that doubt undermines itself by provoking behavior that rids the subject of doubt. However, the purpose is not intellectual contentment but rather contending with one’s *conatus*. The connection between misperception, doubt, and knowledge is also presented in Spinoza’s other discussion of misperceiving the sun in TIE:

Or after we have come to know the nature of vision, and that it has the property that we see one and the same thing as smaller when we look at it from a great distance than when we look at it from close up, we infer that the sun is larger than it appears to be, and other things of the same kind (TIE: 21).

---

<sup>9</sup> See fn 5.

This passage makes the specific claim that misperception leads to greater understanding, which leads to consequent correction of misperception.

Spinoza's four discussions of misperceiving the size of the sun exemplify "doubt", a cognitive state that he defines as agnostic and disinterested in nature, caused by *recognized* lack of knowledge. In the *Ethics*, Spinoza emphasizes the doubtful mind as in a state of equilibrium between ideas. "This constitution of the Mind which arises from two contrary affects is called vacillation of the mind, which is therefore related to the affect as doubt is to the imagination" (E3p17s). For the early Spinoza, doubt is, "nothing but the suspension of the mind concerning some affirmation or negation, which it would affirm or deny if something did not occur to it" (TIE: 80). Doubt "always arises from the fact that things are investigated without order" (TIE: 80). Yet, it must be noted that in the context of doubt, misperception generates some form of doxastic and/or epistemic content as things are "investigated". Doubt, being a doxastic state, though a confused one, promotes a novel basis for appreciating misperception distinct from Bennett and Shapiro as doubt motivates inquiry. How?

The account of doubt in *Ethics* provides a clue for how it motivates inquiry: doubt is a sad state of mind, instigating a natural tendency to rid ourselves of doubt upon its arrival. However, given the disordered state provoked by doubt's vacillation, doubt also has a neutralizing affect on the perceiver lacking "doxastic power" to motivate the perceiver to rid themselves of doubt. Thus, it must be that doubt is not resolved without some extrinsic motivating force, or "affective power". For instance, the negative affects associated with the sun's proximity or distance like feeling too hot or too cold would produce an affective power to rouse investigation.<sup>10</sup>

Imagine a naïve country person running towards the sun in hopes of gaining warmth and realizes that the sun is giving them a literal cold shoulder. They might begin to doubt their belief in the sun's nearness based on the sensory appearance – they might find themselves in a state of vacillation over the location of the sun. This doubt motivates a reconsideration of sensory belief – to question, investigate, and attempt to understand why and how approaching the near appearance of the sun failed them. In a word, the country person would acquire an intellectual belief in the sun that is distinct from the sensory belief, an intellectual belief that benefits behaviour consistent with the country person's welfare. We might state this doubt-based version of passionate misperception as follows:

- 1) We perceive non-existent things (e.g. a near sun).
- 2) Belief in non-existent things is an error.
- 3\*) When cognitive faculties are influenced by a privation of knowledge, doubt is produced.
- 4\*) Doubt generates behavior of acquiring knowledge to produce more beneficial behavior.
- 5\*) Thus, when perception errs, more beneficial behavior is produced.

---

<sup>10</sup> See Steinberg 2017 for further discussion of doxastic and affective power.

*Doubtful* passionate misperception provides a basis for Spinoza's otherwise beguiling account of misperception.

I am not the first to think of Spinoza's account of error as prompting revision. Lenz (2013) arrives at a similar conclusion when considering Spinoza's later work. "What we have now are rather conflicting beliefs. And eventually it will be the emotionally and conatively stronger belief (that is: the one having the greater affective force with regard to our striving) that overpowers the other one" (Lenz 2013, 49). Della Rocca calls such beliefs "winning beliefs", arguing that the alternative losing beliefs remain a "live psychic force" (Della Rocca 2003, 210).<sup>11</sup> The result is, again, not some intellectual discovery of "truth" as normally understood, but rather belief shifting from being less adequate to more adequate in terms of how the perception benefits *conatus*.<sup>12</sup> As Morrison (2015) argues, "[A]dequate' and 'true' are just different ways of picking out the same kind of idea, and therefore the real definition of adequate idea is the same as the real definition of true idea" (Morrison 2015, 91).

Bennett's and Shapiro's accounts of error suggested problems with treating Spinoza's account of perception as non-doxastic. For, perception has behavioral consequences that might be contrary to a perceiver's *conatus* – a naïve perceiver may act on their sensory idea of the near sun and engage in hapless, if not dangerous behavior. But, this leaves a significant problem of explaining how doxastic states are not in error when produced by misperceptions. I have followed Lenz (2013) in suggesting that misperceptions produce an important doxastic state of doubt of *suspended* belief that produces more reliable and thus beneficial behavior. I conclude by considering a further tension when doubt leads to belief revision: holding persisting contrary ideas simultaneously.

## Conclusion

I have argued that Spinoza's theory of perception recognizes that the adequacy of an idea is relative to its intellectual content. For example, were naïve country people to understand the process by which sight is achieved and the relevant planetary knowledge of our solar system, they would acquire a more adequate idea of a distant sun, though also maintaining the less adequate yet persisting perceptual experience of the near sun. But, this is an odd consequence as both less adequate "near sun" and more adequate "distant sun" perceptions exist simultaneously. One possible way to avoid the existence of simultaneous yet contrary perceptions is to distinguish between the manner in which intellectual and sensory content is presented: say, between "objective" and "subjective" manners. The early Spinoza suggests

---

11 Steinburg (2017) has recently demonstrated some problems with this 'dominance' model of belief, as Spinoza presents cases where, all things being equal, a subject believes a weaker idea over the stronger.

12 Tóth (2016) also presents Spinoza's account of error in terms of 'inadequate' belief. "So, the imagination, the product of inadequate causation, as a representation is not true or false in itself" (Tóth 2016, 86).

this response, arguing, “we know that those activities by which imaginations are produced happen according to other laws, wholly different from the laws of the intellect, and that in imagination the soul only has the nature of something acted on” (TIE, 86). Here, the ‘laws’ of the intellect and sensory contents are what distinguish their perceptions, allowing them to be simultaneous yet contrary.

To understand this possible strategy, the analogous example of perceiving sunrises and sunsets is informative. With respect to the sky gazer impressed with a geocentric experience, the sunrise and sunset are, lawfully, impressed on the mind. It is only the intellect that informs the sensory experience with a Copernican view – a distinct set of laws – that reshape the experience to be dual in nature: the earth in rotation orbiting the sun intellectually, the sun in orbit of the earth experientially. Yet, by including the intellect in sensory perception, we do not have the resources to account for how passionate misperception can produce reliably beneficial behavior.

We are left with the following situation. Just as the Copernican sky gazer perceives the sun to rise and set, so too is the perceiver, informed of the great distance of the sun intellectually while maintaining the ‘country persons’ experience of the near sun has two simultaneous but contrary cognitive contents and are ‘of two minds’ about the sun.

## Bibliography

- Aristotle. 1995. *De Anima. The Complete Works of Aristotle: The Revised Oxford Translation*. Edited by Jonathan Barnes. New Jersey: Princeton.
- Bennett, Jonathan. 1984. *A Study of Spinoza’s Ethics*. Indianapolis: Hackett.
- Carey, David P. 2001. Do action systems resist visual illusions? *Trends in Cognitive Sciences* 5 (3): 109–113.
- Della Rocca, Michael. 2003. The Power of an Idea: Spinoza’s Critique of Pure Will. *Nous* 37 (2): 200–231.
- Garrett, Don. 2013. Representation, Misrepresentation, and Error in Spinoza’s Philosophy of Mind. In: Della Rocca, Michael (ed.): *The Oxford Handbook of Spinoza*. New York: Oxford University Press.
- Holt, Edwin B. 1912. The Place of Illusory Experience in a Realistic World. In: *The New Realism: Cooperative Studies in Philosophy*. New York: Macmillan Company. 303–373.
- LeBuffe, Michael. 2009. The Anatomy of the Passions. In: Koistinen, Olli (ed.): *The Cambridge Companion to Spinoza’s Ethics*. Cambridge: Cambridge University Press. 188–222.
- LeBuffe, Michael. 2010. Theories About Consciousness in Spinoza’s Ethics. *Philosophical Review* 119 (4): 531–563.
- Lenz, Martin. 2013. Ideas as Thick Beliefs: Spinoza on the Normativity of Ideas. In: Lenz, Martin – Waldow, Anik (eds.): *Contemporary Perspectives on Early Modern Philosophy: Nature and Norms in Thought*. Dordrecht: Springer. 37–50.
- Morrison, John. 2015. Truth in the Emendation. In: Melamed, Yitzhak Y. (ed.): *The Young Spinoza: A Metaphysician in the Making*. New York: Oxford University Press. 66–91.

- Pylyshyn, Zenon Walter. 1999. Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences* 22 (3): 341–423.
- Renz, Ursula. 2015. From the Passive to the Active Intellect. In: Melamed, Yitzhak Y. (ed.): *The Young Spinoza: A Metaphysician in the Making*. New York: Oxford University Press. 287–299.
- Ross, Helen – Plug, Cornelius. 2002. *The Mystery of the Moon Illusion: Exploring Size Perception*. Oxford: Oxford University Press.
- Sangiaco, Andrea. 2015. Fixing Descartes: Ethical Intellectualism in Spinoza’s Early Writings. *The Southern Journal of Philosophy* 53 (3): 338–361.
- Shapiro, Lisa. 2012a. Spinoza on Imagination and Affect. In: Ebbersmeyer, Sabrina (ed.): *Emotional Minds*. Berlin: De Gruyter. 89–104.
- Shapiro, Lisa. 2012b. How we Experience the World: Passionate Perception in Descartes and Spinoza. In: Pickavé, Martin – Shapiro, Lisa (eds.): *Emotion and Reason in Medieval and Early Modern Philosophy*. Oxford: Oxford University Press. 193–216.
- Sommers, Fred. 2009. Dissonant Beliefs. *Analysis* 69 (2): 267–274.
- Spinoza, Baruch. 1984. *The Collected Writings of Spinoza*. Vol. I. Translated by Edwin Curley. New Jersey: Princeton.
- Steinberg, Justin. 2017. Two Puzzles Concerning Spinoza’s Conception of Belief. *European Journal of Philosophy* 25 (2) (DOI: 10.1111/ejop.12218).
- Steinburg, Diane. 2009. Knowledge in Spinoza’s Ethics. In: Koistinen, Olli (ed.): *The Cambridge Companion to Spinoza’s Ethics*. Cambridge: Cambridge University Press. 140–166.
- Tóth, Olivér István. 2016. Inherence of False Beliefs in Spinoza’s Ethics. *Society and Politics* 10 (2): 74–94.
- Wilson, Margaret. 1996. Spinoza’s Theory of Knowledge. In: Garrett, Don (ed.): *The Cambridge Companion to Spinoza*. 2nd ed. Cambridge: Cambridge University Press. 89–141.
- Wolf-Devine, Celia. 1993. *Descartes on Seeing: Epistemology and Visual Perception*. Carbondale: Southern Illinois University Press.

# XII

*Christopher Davidson*

## A Spinozist Aesthetics of Affect and Its Political Implications

### **PART ONE: Imaginative Affects, Human Complexity and Health**

#### **SECTION 1A: Introduction**

Benedict Spinoza was familiar with and seemingly enjoyed the fine arts and artists. Spinoza kept a sketchbook of his own drawings (Colerus 1706, 33–34), lived in the same area as Rembrandt (Bal–Vardoulakis 2011, 277), had the painter Hendrik Van Spyk as his landlord in the Hague (Spruit–Totaro 2011, 1), and also had the painter Daniel Tydemann as his landlord in Voorburg (Nadler 1999, 203). He learned some of his Latin by acting out scenes from the plays of Terence (Israel 2007, xliv) and Seneca (Spinoza 2016, 144, Curley’s footnote 18), and often quoted Ovid (Spinoza 1985a, 492; Curley’s footnote 4), simply referring to him as “the Poet”, e.g. at *Ethics* 4P17Sch and 3P32Cor (Spinoza 1994).<sup>1</sup> He could expect an educated reader to recognize unattributed quotes from Ovid since he lived during the sophistication of the Dutch Golden Age, well steeped in literature as well as painting.

1 Citations from the *Ethics* follow this standard format: 1P8Sch2 refers to the scholium 2 of proposition 8 in Part One. 1Ax4 refers to axiom 4 of Part One; 2Lem1 refers to the first lemma of Part Two; 3P32Cor is Part Three’s corollary to proposition 32; 4P17Sch refers to the scholium of proposition 17 of Part Four; 4P42Dem is the demonstration of proposition 42 in Part Four. 1Appx refers to the appendix after Part One; 4Pref is Part Four’s preface; 3DefAffXIIIExp is the explication of the thirteenth definition of the affects after Part Three; 4AppxXXVII is entry 27 of the appendix after Part Four. Other texts by Spinoza are cited as follows: TIE par. 52 refers to the *Treatise on the Emendation of the Intellect*, paragraph 52; TTP 2.8 is the *Theological-Political Treatise*, chapter 2, section 8; Ep. 54 is letter 54 from *The Letters*; TP 2.26 is chapter 2, section 26 of the *Political Treatise*.

But his philosophical works say very little about art. His direct references are quite limited; further, even when he explicitly mentions art, he typically uses it to make a point that has little to do with art as such. For instance, when Spinoza names painting and architecture as complex human activities that still must be explained naturalistically (3P2Sch), his points are general rather than narrowly aesthetic: that complex human behavior can be explained through purely corporeal motion, and that mental events do not cause bodily events. Likewise, his comments about the amnesiac Spanish poet (4P39Sch) are about memory and identity, not poetry. Spinoza says elsewhere that whether a song would be considered good or bad depends not on the song itself, but on the person who hears it: “for example, music is good for one who is melancholy, bad for one who is mourning, and neither good nor bad to one who is deaf” (4Pref, II/208). While melancholy will be decisive later, for now, I note only that Spinoza again is making a very general point (things are not good or bad inherently but only relatively to each person) rather than speaking about art as such.

Similarly, in the unfinished *Treatise on the Emendation of the Intellect* (Spinoza 1985b), his remarks about deliberately feigning fictions which are nonetheless not false (TIE par. 52–67) make epistemological points: feigning hypotheticals should not be confused with error, for imagining strictly contains no error. His statement that an architect imagining a building has a true idea about the building is intriguing, but since he repeats this exact claim in reference to the genetic definition of a circle (TIE par. 69–72), his point is once again quite general: “the form of the true thought must be placed in the same thought itself without relation to other things, nor does it recognize the object as its cause, but must depend on the very power and nature of the intellect” (TIE par. 71). Neither here, nor at 2P17Sch (where he combines these two points), does he develop the idea into an aesthetic point.

In the *Theological-Political Treatise* (Spinoza 2007), he tells us that the Old Testament says that music was sometimes required for Elisha to have prophetic visions (TTP 2.8), that the Old Testament contains a relatively accurate national history of the Ancient Hebrews (even if written down far after the events) because singing their history as psalms aided their memory (TTP 12.12), and that music can ease melancholy (TTP 1.25). The affect of melancholy is a near-total inability to act, so music is quite powerful in Spinoza – but these passages do not say why music aids memory, eases the paralyzing affect of melancholy, and jumpstarts the “very rare” (TTP 1.30) power of prophecy – let alone establish what art in general is for Spinoza. We do not have much to go on, then, and it is not clear what, if anything, unites these comments. What are we to do if we want to turn Spinoza’s philosophy to aesthetic purposes?

These seem to be comments made almost at random. However, while Spinoza does not directly say why music can cure melancholy either time he makes that claim, a closer look at the definition of melancholy will give an initial hint at the content of a Spinozist aesthetics. Melancholy’s definition shows that art is a crucial tool to keep the complex *ratio* of our many parts in balance, which Spinoza calls health. When Spinoza defines the affects, he begins with desire, which, “related to the mind and body together [...] is nothing but the very essence of man, from whose nature there necessarily follow those things that promote his preservation” (3P9Sch). Desire, joy, and sadness are the three “primary” affects, and all other affects “arise from these three” (3P11Sch). Joy and sadness are thus necessary to

understand any passion. “By joy, therefore, I shall understand in what follows that passion by which the mind passes to a greater perfection. And by sadness, that passion by which it passes to a lesser perfection. The affect of joy which is related to the mind and body at once I call pleasure or cheerfulness, and that of sadness, pain or melancholy” (3P11Sch). Melancholy is specified further at 4P42Dem as a species of “sadness, which, insofar as it is related to the body, consists in this, that the body’s power of acting is absolutely diminished or restrained. And so (by P38) it is always evil, q.e.d.”. While 4P42Dem comments only on melancholy’s absolute restraint and diminishment of the body, its mental effects are comparable insofar as melancholy is “related to the mind and body at once” (3P11Sch).<sup>2</sup> Since our power of acting is all that we essentially are according to Spinoza, nothing (other than death) could be worse than melancholy: absolute diminishment of perfection, power, or existence itself. Anything which could ease this nearly absolute evil would itself be a nearly absolute good for those with melancholy, and a most useful tool for reducing sadness in general. The only treatment of melancholy Spinoza ever named was music.

While Spinoza’s references to art are not numerous, they are consistent and link to many of his key concepts. Anchored by the concepts of health (from the *Ethics*) and prophecy (in the *Theological-Political Treatise*), this paper establishes the broad outlines of what Spinoza’s philosophy can tell us about art. First, linking art to Spinoza’s sense of health, as the joy-producing restoration of one’s capacity to perform many actions, shows us the power of the pleasures of art. To explain why art can restore health, we must explain how Spinoza defines human beings through their complexity, show that this complexity is identified with a *ratio* of all our parts, define health as the restoration of that *ratio* which leads to joyful pleasures, show that pleasures which affect the whole are more valuable than those that affect only parts, and finally establish music as a uniquely valuable path to health.

The paper has a second main point, which is of a social or political nature. Like the vivid visions of prophecy, artwork deploys intense affects on groups of people. Politically, prophecy’s affects are used to form and shape, constitute and control, the nation as a whole. When the use of social affects is repeated through public ceremonies, the nation eventually develops a relatively consistent character or set of affective habits. However, small sub-genres of art expose only small groups (not the nation in general) to their intense affects. If this small group of people (what I call a “sub-genre”)<sup>3</sup> consistently undergoes different affects, they over time will develop different habits and ways of life than the norm. That is, the production of similarity within an artistic sub-genre is simultaneously the production of a difference relative to the nation at large: small groups partially differentiated from the multitude at large.

---

2 3DefAffIIIExp states that melancholy, pain, cheerfulness, and pleasure are “chiefly related to the body”. However, by the crucial propositions 2P7 and 3P2, while any mode’s affection may be conceived under either attribute (thought or extension), that mode necessarily expresses the affection under both attributes simultaneously. Put otherwise, melancholy renders the mind unable to think just as much as it renders the body unable to move.

3 I use the phrase “sub-genre” to highlight the distinction between the smaller number of adherents of a specific type of music and the larger generic culture of the “people” (*gens* in Latin; see TTP 2.39 for *gentes* as “peoples”, though often it means Gentiles or pagans, as at TTP Pref.7).

## SECTION 1B: Beauty, No; But an Aesthetics, Yes

Before we go any further, I must acknowledge a serious obstacle to any attempt to develop a Spinozist aesthetics.<sup>4</sup> Often, discussions of art revolve around some concept of beauty – but Spinoza emphatically states that beauty is an empty concept which does “not indicate the nature of anything” (1Appx, II/83) referred to as beautiful. Hence, one of the few attempts to locate an aesthetics in Spinoza is titled “Why Spinoza Has No Aesthetics” (Morrison 1989). No wonder that some have thought this, since the Appendix to Part 1 of the *Ethics* states that beauty is not in the thing, but is only an inadequate judgment about our own bodies:

when things are so disposed that, when they are presented to us through the senses, we can easily imagine them, and so can easily remember them, we say that they are well-ordered; but if the opposite is true, we say that they are badly ordered, or confused [...] For example, if the motion the nerves receive from objects presented through the eyes is conducive to health, the objects by which it is caused are called beautiful; those which cause a contrary motion are called ugly (1Appx, II/82).

For Spinoza, each sensation produces an idea (a mental event) which is a mixture of the nature of our own body with the nature of the external body, mixed together in an inaccurate way. He calls this imagination (2P17Sch) and details its epistemic limitations at 2P11–2P26. “Imagination” refers only to ideas caused by an external body interacting with a human body, and in Spinoza’s technical vocabulary has no connotations of artistic creativity or freedom of thought.<sup>5</sup>

We never sense the object as it is, but instead sense how our own body is affected or changed by the object. The experience of the object varies person to person, due to the fact that no two human bodies are identical. Different nerves and different abilities to remember<sup>6</sup> will produce different judgments about order and beauty.<sup>7</sup> Judgments about art, then, will vary as much as individuals do: “each one has judged things according to the disposition of his brain”, and “although human bodies agree in very many things, they still differ in very many. And so for that reason [...] what seems ordered to one, seems confused to another; what seems pleasing to one, seems displeasing to another, and so on [...] This is why we have such sayings as [...] ‘there are as many differences of brains as of palates’” (1Appx, II/83). Beauty cannot be rooted in any adequate understanding of the object (2P25), ideal, transcendental forms (2P40Sch), or ordered harmony (1Appx).

---

4 “Aesthetics” is an anachronistic term when applied to Spinoza, and its meaning is also highly contested. However, I use the term anyhow since a common meaning of the term, how one’s body or sensation is affected by artworks, is central to the paper.

5 As de Deugd notes (de Deugd 1966), that Spinoza partly inspired German Romanticism is really quite surprising. Romanticism, which casts the creative genius as a mystical, transcendental force, hardly seems compatible with Spinoza’s immanent determinism.

6 Different lifestyles generate different mental associations as well, as Spinoza notes at 2P18 and TPP 2.7. Both passages contrast farmers’ and soldiers’ associations.

7 This idea is central to his ethical theory as well: goodness is not an inherent property of actions or of things, based on the same reasoning. See Ep. 19, Ep. 54 (Spinoza 1995), 3P15, and 3P39Sch.

Since Spinoza insists that “things regarded in themselves, or as related to God, are neither beautiful nor ugly” (Ep. 54), will it even be possible to think about artworks within Spinoza? In short, yes: art is not exhausted by the concept of beauty. Aristotle showed that successful artists do not need to craft beautiful objects; spectators’ delight in mimetic (imitative or representative) art forms comes mainly from plots which teach us about the nature of human actions (Aristotle 1984). Therefore art can just as properly be ugly or painful. Many contemporary art theorists also avoid relying primarily on an aesthetics of beauty. Danto describes how painting in the 1970s refused to be hemmed in by a narrow concept of beauty he traces back to Kant (Danto 2001). Lyotard, in many influential works (e.g., Lyotard 1994), uses Kant’s concept of the sublime – a category which Kant insists is non-artistic and only is experienced in nature – to argue that the best art is sublime and hence to minimize the importance of artistic beauty. Deleuze counters a Platonic tradition in the philosophy of art (which wants truth, beauty, and goodness to be united in a recollected ideal form) by arguing that aesthetic experience must generate *sensibilia* which force us to think in new, often painful or ugly, ways (Deleuze 1994; Deleuze–Guattari 1994). These thinkers, and others, have shown that aesthetics can have foundations other than beauty: that beauty is denied by Spinoza need not exclude Spinozists from thinking about art. Very recently, the scholarship connecting Spinoza to art has dramatically expanded: see Vardoulakis’ (ed.) *Spinoza Now* (Egenhofer 2011; Uhlmann 2011; Bal–Vardoulakis 2011), and Lord’s (ed.) *Spinoza Beyond Philosophy* (Rawes 2015; Cimini 2015; Calder 2015), which also includes artworks inspired by Spinoza. James writes on “Narrative as the Means to Freedom” (James 2010), while Gatens discusses George Eliot’s Spinozism in multiple pieces (e.g., Gatens 2012) and shows artists’ similarities to prophets (Gatens 2015). There is also Vinciguerra’s book on Spinoza and semiotics (Vinciguerra 2005) and a related article (Vinciguerra 2012). This is a sharp uptick in contributions considering that there was little prior discussion of these topics other than books by Mignini (Mignini 1981) and de Deugd (de Deugd 1966) which both dealt with Spinoza and art throughout. These contributions show that there is much of value in this developing field of Spinozist aesthetics.<sup>8</sup> Indeed, avoiding the well-worn path of beauty (as most of these contributions do) allows a Spinozist conversation about art to proceed in new directions.

## SECTION 1C: Essence of a Human Individual Defined as Its Complex Ratio

The universe is a nothing but a nexus of causes and effects which can be explained in terms of motion. Therefore, what follows below is true of all things, including human and non-human things. There is no human exceptionalism in Spinoza: all human activities “follow from the same necessity and force of nature as the other singular things” and thus must be treated “by

---

8 Interest is also shown by recent international conferences on the topic: e.g., “Spinoza et les arts”, May 15–17, 2014, Paris, and “The Arts of Spinoza + Pacific Spinoza”, May 26–28, 2017, Auckland. *Interstices: A Journal of Architecture and Related Arts* plans to publish an issue with papers from the latter.

the same method”, allowing us to “consider human actions and appetites just as if it were a question of lines, planes, and bodies” (3Pref). Hence, Spinoza says that “paintings [...] should be able to be deduced from the laws of nature alone” (3P2Sch). If sexual jealousy, in its all-too-human messiness, can be explained by Spinoza as if it were lines, planes, and bodies (as at 3P35), and if methods akin to natural science can be applied to revelation and the history of the Jewish people (as in the TTP), then art ought to be susceptible to the same method.<sup>9</sup>

Spinoza says that “there must be, for each existing thing, a certain cause on account of which it exists [...] For example, if 20 men exist [...] there must necessarily be a cause why each [particular man] exists. But this cause [...] cannot be contained in human nature itself, since the true definition of man does not involve the number 20” (1P8Sch2). Since “knowledge of an effect depends on, and involves, the knowledge of its cause” (1Ax4), each thing (including each person) must be understood as a whole composed of parts, one which does not precede in time, or otherwise exist independently of, its causally determined parts. The existence of actual particular things “does not follow from the nature of these things, but from the order of the whole of corporeal nature” (1P11Dem). I take these passages to affirm that essence is individual, rather than there being a shared essence typically called “human nature”.<sup>10</sup> That is, a person is nothing other than the many parts which make her up, their motions (causally affected by external objects or “corporeal nature”), and the interactions of these parts’ motions.

Under the attribute of extension, the simplest physical bodies are distinguished only by their motion (2Lem1). When many of these small bodies “communicate their motions to each other in a certain fixed manner, we shall say that those bodies are united with one another and that they all together compose one body or individual, which is distinguished from the others by this union of bodies” (2Def, II/100). Such individuals can then become

9 Mignini says much the same (Mignini 1981), showing that aesthetic imagination neither itself gives adequate knowledge, nor is it pure chaos which is utterly immune to analysis. Insofar as all affection or sense experience (imagination) is causally determined, there is truth to it at the level of what really happens to our body. There are adequate ideas in God of the external object, of the human body, and of their causal interaction. That is, all experience is in principle knowable in some way (5P4), including aesthetic experience as a species of imagination. Since all imagination (including aesthetic experience) causes pleasure or joy when it truly increases our power of acting, through it we could in principle gain some information about ourselves. Thanks to Olivér István Tóth for an illuminating discussion of Mignini.

10 A Spinozist aesthetics need not rely on a universal human essence, making it a promising alternative to Kantian aesthetics based on universal faculties. However, there is debate whether essence in Spinoza is individual or universal. Hübner summarizes the debate and presents her own subtle position on “resemblance” or what humans have “in common” (Hübner 2014, 2016), which is reason (the ability to have adequate ideas) and is universal. But Spinoza’s nominalism denies that universal terms, including “man”, can have any adequate content (1Appx, 2P40Sch1, and 4Pref). Essence is also defined as a thing’s striving “to persevere in its being both insofar as it has inadequate ideas and insofar as it has adequate ideas” (3P7–9) which differs for each person (3P57). Hübner acknowledges such passages but establishes her claim using others. However, despite the textual support, there are difficulties with reason as human essence. It is said to be our real essence, yet held up as an ethical model of what we are not yet. It is said even of those who are highly irrational. It is said to be universal, yet each has it in different degrees or ways. Finally, you and I shaking hands (or undergoing the same affect) have a resemblance, yet these are not said to be our essence.

composed into larger, more complex individuals called “composite individuals” (2L7Sch). A human body is itself composed of many such composite individuals. All of these composite individuals, taken together and including all their interactions, define what a thing truly is. Spinoza calls this the *ratio*, the characteristic fixed proportion of motion and rest or “certain and fixed manner” (2L7Sch) in which a thing moves: what the thing can really do and characteristically does is a combination of all the part-individuals’ motions which interact to form the whole-individual.

After affirming that every thing strives to persevere in its existence (3P6), Spinoza then identifies essence with such striving and every actual effect which follows from it. “So the power of each thing, or the striving by which it (either alone or with others) does anything, or strives to do anything [...] is nothing but the given, or actual, essence of the thing itself” (3P7Dem). So, a thing’s essence is just the striving to preserve the sum total of its various motions, the relations between its various ideas, or the interactions of its affects. This striving defines all that one is and all that one can do. I am nothing but my power of acting which is my *conatus*; that power is determined by the whole complex *ratio* of all interacting parts that make up my body. (The same is true for the mind of that body, but we will restrict our discussion to the attribute of extension.) Then Spinoza identifies striving with desire at 3P9Sch. Thus, affections and the modifications of desire or striving that follow from them are modifications to the very core essence of what one is, rather than incidental fluctuations in one’s accidents.

This complexity is also what makes humans relatively unique. That is, the *ratio* defines each individual thing as a unique balance of various motions, and the generally high level of complexity of human *ratio* distinguishes humans from other things. Spinoza has told us that composite individuals are made up of parts which are themselves individuals, that humans are made up of very many highly composite individuals as our parts (2P15Dem), and that the whole of the universe is itself an even more highly composite individual (2Lem-7Sch). Individuality, then, as the complex *ratio* of motion and rest of very many interacting parts, is not unique to humans – “below us”, our organs are individuals, and “above us”, the universe itself is an individual.<sup>11</sup>

For the things we have shown so far are completely general and do not pertain to man more than to other individuals [...] However, [...] the difference between the human mind and the others, and how it surpasses them [is] that in proportion as a body is more capable than others of doing many things at once, or being acted on in many ways at once, so its mind is more capable than others of perceiving many things at once (2P13Sch).

Humans are themselves nothing but natural objects. Spinoza denies any specific qualitative content to human individuality – but we do excel other things solely by the complexity

---

11 See Ep. 32 for the important “worm in the blood” analogy, which elaborates how individuals can be nested within other larger individuals.

of our bodily motion (or, under the attribute of thought, by the number, persistence, and complexity of our ideas).

## SECTION 1D: Health as Maintaining the Whole Person

Human complexity is power: it allows us to do more. But this complexity requires extra maintenance, just like a machine with many complex parts. Think of the difference between a sports car and a child's bicycle. One is complex and capable of varied things and high speed but perpetually requires nourishment in the form of gas to do anything at all and replacement parts to function well. The other is simple, requiring no fuel and needing repairs rarely but limited in its speed and in the number of tasks it can perform.

Spinoza explains how vital it is for us to constantly readjust the many parts which make us up. While it is most succinctly stated at 4P38, 4P39, and 4P45Sch, Spinoza connects the terms “health”, “whole”, “composed of many parts”, “capable of many things”, “all parts equally affected”, “requires many other things to be preserved”, and “useful for preserving the proportion of parts” in many additional places: e.g. 2Post1, 2PostIII–IV, 2Lem4–7, 4P31, 4P42–44, 4P60–62, and 4AppxXXVII–XXX. At 4P38, he states: “Whatever so disposes the human body that it can be affected in a great many ways, or renders it capable of affecting external bodies in a great many ways, is useful to man; the more it [does so], the more useful it is; on the other hand, what renders the body less capable of these things is harmful”. He links this directly to the *ratio*: “Those things are good which bring about the preservation of the proportion of motion and rest the human body's parts have to one another”, while if “the human body's parts acquire a different proportion [then] the human body is destroyed, and hence rendered completely incapable of being affected in many ways” (4P39 and 4P39Dem). To “die” in Spinoza apparently means to “acquire a different proportion” permanently while “illness” means a temporary disruption of the proportion (4P39Sch). Spinoza continues: “the human body is composed of a great many parts of different natures, which constantly *require* new and *varied* nourishment, so that the *whole* body may be equally capable of *all the things* which can follow from its nature, and hence, so that the mind also may be equally capable of understanding *many things* at once” (emphasis added, 4P45Sch). If a part is not replaced, the particular power that it contributes to our *whole* essence (our overall *ratio*) is lost: that is, we can do *less*, we become less. If enough complex parts go without replacement or we go unnourished, we lose power, fall ill, are restrained, reduced, and simplified. To replace our varied complex parts, we *require* a *variety* of interactions with many kinds of objects.

4P45Sch also lists a wide variety of things which can restore health, including multiple forms of art. Art does not consist of things which are themselves beautiful, but it does provide pleasure: specifically, the pleasures of restored health, which is a core principle of our Spinozist aesthetics. As Spinoza explains,

Nothing forbids our pleasure except a savage and sad superstition. For why is it more proper to relieve our hunger and thirst than to rid ourselves of melancholy? [...] On the contrary, the greater the joy with which we are affected, the greater perfection to which we pass [...] To use things, therefore, and take pleasure in them as far as possible [...] this is the part of a wise man. It is the part of the wise man, I say, to refresh and restore himself in moderation with pleasant food and drink, with scents, with the beauty of green plants, with decoration, music, sports, the theater, and other things of this kind [...] This plan of living [...] is best, and to be commended in every way (4P45Sch).

As important as praising joy and condemning sad superstition are, the core of a Spinozist aesthetics consists of the arts' ability to "restore" the whole body's power through "new and varied nourishment" (4P45Sch; see also 4DeAffXXVII–XXXI). A human is defined by the complexity of her *ratio*, which allows her to do and think many things – but this complexity requires variety to be maintained. Art, among other things, exposes humans to variety, which is necessary to maintain our health. To not maintain health leads to inactivity and limitation of possible actions, thoughts, and affects, and potentially to melancholy or even destruction. Since we are nothing other than what we can do, for Spinoza, less art leads to reduced complexity, which leads to the essential self being reduced.

However, since food, drink, and even looking at plants also cause nourishment – what is special about the arts?<sup>12</sup> How arts excel other pleasures is that art (at least in the case of music) has the potential of positively affecting the whole person, rather than just a part. Recall that Spinoza says that "music is good for one who is melancholy, bad for one who is mourning, and neither good nor bad to one who is deaf" (4Pref, II/208). For music to be good for the one who has a completely disrupted *ratio* (melancholy is when one's parts "all are equally affected" with sadness, 3P11Sch), it must be able to affect the whole person at once: what we might call a "wholesome pleasure". Since we are identified with our *ratio*, the whole balance of our many moving parts, to affect this whole *ratio* positively is a benefit at the deepest, strongest level possible. "Health" for Spinoza is the life-sustaining support of that whole-self, complex, essential *ratio*, and music has a unique power to produce that overall health. Thus, music – and Spinoza only named music as a cure for melancholy – can produce cheerfulness, a "joy which [...] consists in this, that all parts of the body are equally affected" which "is always good" (4P42Dem).

Further, when music affects the whole, it is by definition not affecting only one or a few parts. That is, music avoids the risks of other, localized pleasures. When other kinds of pleasure affect only one or a few parts, they run the risk of being harmful. "Pleasure can be excessive and evil [...] Pleasure is a joy which [...] consists in this, that one (or several) of its parts are affected more than the others [...] The power of this affect can be so great that it surpasses the other actions of the body [...] and so prevents the body from being capable of being affected in a great many other ways" (4P43 and 4P43Dem). We know from

---

12 I appreciate Daniel Schneider for pressing me on this point in questions responding to my paper.

experience that food, drink, and other localized pleasures can easily harm by negatively impacting the whole; pleasures can also lead to obsession (as in the case of sexual desires, Spinoza says) when the mind is occupied by one idea so strongly that it cannot think any others. It is hard to imagine music or theater causing these problems.<sup>13</sup> Spinoza notes “there seem to be very few things of this kind in nature” that can aid the “preservation of our body [in] all its parts” (4AppxXXVII). That is, music appears to offer a very rare kind of joy that not only avoids the risk of a local pleasure in one part that causes a larger harm to the total self, but also holds the singular advantage of easing the greatest evil (the most passive state) a living person can experience.

These points are consistent with Spinoza’s other comments about health in the *Treatise on the Emendation of the Intellect*. Spinoza there maintained that we should “enjoy pleasures just so far as suffices for safeguarding our health” (TIE par. 17). In fact, this point is one of only three practices that we should “assume [as] rules of living [which are] good” (TIE par. 17). Again, at TIE 13–17, Spinoza lists health as one of 5 components of the highest good which will produce true happiness, once again showing that he takes health utterly seriously. Similarly, he contrasts useful, healthy pleasures with harmful pleasures at 4P38–44 and 4AppxXXVII–XXX, saying that we must take care not to obsess over money at the expense of overall bodily health (4AppxXXIX and 4P44Sch), just as he notes how easily we become “distracted” by money so that we are “quite prevented from thinking of anything else” (TIE par. 3–4). Therefore, Spinoza is consistent about the importance of overall health, and about why some pleasures (e.g., the arts) are most useful for health: while other pleasures in their one-sidedness can derange our minds and disorder our bodies, art maintains the power of the whole person and thus is crucial. Art is less likely to distract or unbalance. As a corollary, art’s unusual ability to work on the whole complex of parts which we are explains music’s unique ability to reverse melancholic ills which have “absolutely diminished or restrained” (4P42Dem) a person in each of her parts.

It would seem at first glance that linking art’s pleasures to the health of the whole complex human being requires that we say that all art which heals must be complex – but this is not in fact required. Human complexity consists in having a wide variety of parts first and foremost, with some (but not all) of these parts themselves being complex. So, some of the parts we lack (and find healthy pleasures in the replacement of) could be simple parts replaced by simple art-objects. Variety, having many different kinds of parts, is not the same as complexity of each part. Additionally, even for our complex parts, healing them may only require that one of their simple sub-components be replaced. A complex five-point seatbelt harness in a race car might be fixed by replacing just one bolt. Therefore, the usefulness of art seems to be as much about its variety as it is about any particular artwork’s complexity – hence why art is included in a list with things which do not appear to be relatively complex (like green plants),<sup>14</sup> and hence why that list is itself highly varied. Perhaps a refined aesthete could reflect on the particular mood (set of affects) that she had

---

13 Sports, however, can more easily become obsessional, but I speculate that happens when sports activate the same affects of nationalism – “us” versus “them”, and hope mixed with fear – analyzed in the TTP.

14 Thanks to Tobias Sebastian Dreher for giving me occasion to reflect on this point in our conversations.

at any given moment, and tailor a specific “cure” – sometimes a simple scent, sometimes the complexities of Joyce’s *Ulysses* – for whichever particular parts happened to need replacement or nourishing.<sup>15</sup>

Though music seems to have a unique power, it is not *the most* useful thing. Spinoza says that other people have the greatest potential to be useful (see TTP 5.7, 16.5–16.7, and the political sections of the *Ethics* at 4P18Sch and 4P31–37). “For what is most useful to man is what most agrees with his nature (by P31C), i.e. (as is known through itself), man. But a man acts entirely from the laws of his own nature when he lives according to the guidance of reason (by IIID2), and only to that extent must he always agree with the nature of the other man” (4P35Cor1). Spinoza qualifies this: “Still, it rarely happens that men live according to the guidance of reason” (4P35Sch). Briefly, a few points: first, fully rational people are the most useful. But they are a quite rare resource, while art is much more abundant resource, which hence might be of particular usefulness. Second, agreement occurs in degrees or to different extents: perfectly rational people agree always in their natures and so are most useful, people of average rationality agree less often (but still sometimes agree) and so are somewhat useful. As non-rational, affects in general can pit people against each other, but affects do not always fully disagree: they can agree to some extent, and insofar as they do, they bind rather than divide. As 4P34Sch notes, one’s affects, “insofar as they agree in nature” with another’s affects, can encourage joy and so “it is far from true that they are troublesome to one another insofar as they love the same thing”. Third, there are ways to encourage a higher degree of similar joyful affects in a group, such that they more often (though not always) agree in nature and thus become more useful to each other. These ways of encouraging group cohesion are discussed next.

## PART TWO: Political Implications Of Affects and Ceremonies

Companions in pathos, who barely murmur  
[...] A new mystery sings in your bones. Cultivate your legitimate strangeness.

(René Char, *Partage Formel* XXII, quoted in Foucault 2009, xxxvi.)

The second part of the paper discusses the formation of new capabilities which can occur when similar affects are caused in a group through repeated exposure to public ceremonies. Spinoza’s comments on religious ceremonies and rituals helps us understand why we

---

15 Huysmans’ novel, *À Rebours* (*Against Nature* or *Against the Grain*), is about a decadent aesthete who employs such techniques, including gilding a tortoise and covering it with carefully selected jewels (Huysmans 1998). This particular experiment was not a success, to put it mildly.

commonly see that groups of people who listen to similar music strongly resemble each other in their bodily expressions of clothing, mannerisms, and dance. Enough exposure to a musical genre can reshape its listeners and bind them into a small, but powerfully united, community. The focus in this part of the paper differs from the first in two main ways: one, the level of analysis now is the group, not the individual. Two, while health is about maintaining one's already existing *ratio*, social binding through ceremonies changes what one is. Health is based on the individual body or unique *ratio* of a person, while joining a niche artistic community (by consistently following specific kinds of artists) exposes one to shared experiences in the form of public ceremonies which slowly modify the community. Undergoing ceremonies in a musical sub-genre, beyond restoring health, can lead to a rebirth in a new kind of life.<sup>16</sup> Art in general can produce individual health (the result of preserving the existing self), while the less common expansion of oneself through repeated exposure to unique ceremonies can increase one's capacities. The "third nature" developed by a small community of likeminded art lovers expands them to a new level of power, founding a new community trained through affectivity to cultivate selves which do not yet exist.

## SECTION 2A: Harnessing Affectivity in Order to Move the Generic Multitude

My discussion of prophets will be restricted to just those points regarding their repertoire of affective methods that might also explain the effects artists have on their audience. These effects are political, if we understand politics as Spinoza does: a "gathering of people" (TTP 16.8) by directing or conducting the behavior of the multitude such that they have the ability and desire to move as a more or less unified body. Politics induces affects and spreads them through the social body in order to provoke behavior and desires in the multitude, causing determinate effects in the real actions of people. The understanding of "politics" presented here is largely in line with Balibar's reading of Spinoza (e.g. Balibar 1998, and 1997). That is, like Balibar, I emphasize that politics certainly works through laws and institutions but also through the imitation of affects; that politics is the shaping of affects which constrains but also enables; and that this significantly molds – and can even create – a people.

On this reading, politics involves so much more than laws, and overlaps with a wide range of interpersonal relations and the social world generally. A full range of affective

---

16 There is a central ambiguity regarding the nature of an individual's striving: can something strive to change its essential *ratio*, to negate what it is? Seemingly not; but if not, then how can it strive to increase its power? Does a thing's striving mean being "opposed to everything that can take away its existence" (3P6) as maintenance, or rather striving to "increase or aid [its] power of acting" (3P12) as expansion? The question has many angles; see a variety of treatments in Miller 2015, Chapter 3, especially 120–143; Davidson 2015; Marshall 2014, 78–86; Ravven 2004; Ansaldi 2003, and multiple works by Garrett (e.g. Garrett 1994).

techniques sustains political, religious, and social matters: ambition, fear, hope, and love (directed at God, neighbor, religious leaders, and politicians) are all essential to politics.<sup>17</sup> Spinoza's analysis of ancient prophets (and, by extension, of modern religious leaders' use of prophetic texts) shows us the imaginative techniques used, how effective they are at winning the hearts of the masses, and the actions that they cause. The political stakes of prophetic imagination are central to the project of the TTP as a whole, as Spinoza repeatedly unifies or identifies the functions of politics and religion: both are concerned with nothing but the obedience of the multitude, which Spinoza also calls "morality". That is, the use of prophecy and the goal of the politician are simply social cohesion and the peace which follows from it. Binding disparate individuals together as "a people" cannot primarily be accomplished through reason, and imaginative affects are an effective method of doing so. This binding and cohesion theoretically aims to approximate universality, but Spinoza says this is impossible and undesirable in practice (TTP 17.1–7, 20.6, 20.13).<sup>18</sup> Unification is always partial and obedience does not make slaves of people (TTP 16.10), nor does differing in minor ways necessarily make one a rebel against the state.

The preceding summary of theological-political use of imaginative affects will be detailed below, but allow me to briefly state now the relevance for art (also expanded below). Artists use social affects, as do prophets, though not to the end of obedience. Art can eventually engrave itself deeply into the bodies of the audience, especially when it relies on public performances such as live music, and especially for distinctive musical niches or sub-genres which repeatedly echo specific, unique affects. Much like the public ceremonies that founded and sustained the Ancient Hebrews, ceremonies unique to a niche or sub-genre of art can produce and continually reproduce a small group. Since affects produce behavior, and ceremonial repetition of affects can produce habits and ways of life, repeated exposure to similar affects will over time produce a specific kind of people, simultaneously distinguished from the multitude at large while unified with each other. This distinction will never be absolute, and need not be one which rejects or negates the broader community. If the ceremonial shaping of a people through theological-political ceremonies can be thought of as a "second nature" (TTP 17.23), then we can say that those who find themselves repeating a sub-genre's unique ceremonies acquire a "third nature". This third nature will expand what they are, what they can do, what they desire, and the range of affects they experience, rather than simply negating what they have been.

The multitude that make up a nation are guided by prophet and politician to become, as Spinoza says in the *Political Treatise*, "as if they were one mind" or body (TP 2.26, 2.21, 3.2,

---

17 Notably, prophets and politicians primarily use affects of hope and fear (TTP Pref.1), while artists use the full range of affects. Followers of art are more likely to experience what Spinoza calls "devotion" specifically and love generally. While this is crucial, I merely note it here rather than treating it in full.

18 Thank you to Steph Marston for her questions after my paper, in response to which I have reduced my original over-emphasis on the degree of universality of the multitude's affects or their degree of obedience.

3.5, and 3.7), insofar as they adhere to the same social-moral laws.<sup>19</sup> The TTP also shows that the state's management of affects through "morality", just as much as its institutional organization, determine "how much [...] the state could guide men's minds" (TTP 17.16) and that stable states are those whose citizens are "far more closely bound to each other" by having a similar morality (TTP 17.27). What Spinoza calls "morality" in the TTP strictly means any shared behavior which, preventing neighbor from turning on neighbor, protects the peace: hence its identification with "obedience". Even Moses' ban on adultery is about unification, not about preventing sin or evil. Morality is "only relevant to the temporal prosperity of the state [and] relates only to the interest of the commonwealth" (TTP 5.3); that is, the commandment not to commit adultery is only called "morality" because it prevents socially disruptive fights over patrimony. "Morality" as defined in the TTP consists of whatever shared actions tend to preserve a group; when a mass of singular individuals begins to act in the same way, they begin to cohere as a people.

Neither "morality" nor nations are natural, for Spinoza; humans do not automatically act similarly. No two people will automatically desire the same behaviors because of the natural differences between the *ratio* of each individual. Thus, we must be formed and re-shaped until our desires are less dissimilar – not identical, but at least no longer directly opposed. "Nature certainly does not create peoples, individuals do, and individuals are only separated into nations by differences of language, law and morality. It can only be from these latter factors, namely law and morality, that each nation has its unique character, its unique conditions" (TTP 17.26). A group comes to be constituted similarly over time as the effect of determinate causes, or comes to approximately resemble each other via forces of historical constitution. There is no essential spirit of a people, no timeless national character.

## SECTION 2B: Fiction, Not Truth, Moves the Masses

Spinoza says that it is prophets and politicians who convince the people to act in ways that bind them more tightly, that better prevent dissociation. And how do they do so? Precisely through causing changes in the affects of the multitude by deploying imaginative techniques. Prophets affect the general multitude "by means of the imagination, namely by means of words or visions" (TTP 1.20). Religious prophets, as defined by Spinoza, are those who have particularly "vivid" imaginations (TTP 1.20), that is, those who experience intense affects or sensations (prophetic visions) that they then transmit to others "in a language and a style designed to influence the minds of the common people" (TTP 13.1). In Spinoza's deterministic account, to cause the affections of others is to cause some kind of an effect: to determine their actual behavior. The truth of prophetic vision is immaterial: the masses are moved, in fact, more readily by superstition than by true reason. Notably, all kinds of books could inculcate obedience, not

---

<sup>19</sup> Unification will never be total, of course. Even a tyrant cannot force all subjects to have identical desires: see TTP 17.1–3.

only revealed texts – such is the power of “vivid” affects like those stirred up by art. “It is due to the salutary opinions that follow from them that the narratives of the Old and New Testaments are superior to other, non-sacred legends, and even among these, some are superior to others. Hence if anyone reads the stories of holy Scripture [...] without amending his life, he might just as well read the Koran or the dramatic plays of the poets, or at any rate the common chronicles” (TTP 5.19). There are a few points I would like to underscore in this interesting passage. First, all texts are politically valuable to the extent that they promote obedience: this is the sole theological-political goal. Second, any book could do this in theory: all of them have the power to stir up affects and amend lives. Third, they are all legends: their truth is not at all the issue. Art, prophecy, and religion are effective, regardless of their truth: “it was not the purpose of the Bible to teach any branch of knowledge [...] it requires nothing of men other than obedience” (TTP 13.3). Art in Spinoza need not be beautiful, nor does it need to contain truths, whether truths of realism or of Platonic ideal Forms. “For this reason, it speaks in a wholly inexact manner [...] precisely because it is not seeking to sway men’s reason but to influence and captivate their fancy and imagination [...] it is very appealing to them when everything is narrated poetically” (TTP 6.15). Artists deploy powerful affects that can send visceral affects directly into the social body, with an eye toward moving the people, not toward truth.

## **SECTION 2C: Politicians and Prophets Are Nearly Identical, While Artists Differ Primarily Due to Amorality**

Spinoza’s *Theological-Political Treatise* treats religion and politics as nearly identical forces. There is slim difference between religious “prophets” (the analysis of whom Spinoza applies to religious leaders generally) and politicians: both recommend behaviors to their subjects or followers and seek to cause that behavior with greater frequency through a combination of laws and affective levers such as public ceremonies. Machiavelli explicated the similarity before Spinoza,<sup>20</sup> referring to successful politicians like Moses as “armed” prophets in the *Prince* (Machiavelli 2009, chap. 6). Spinoza certainly follows Machiavelli here: both analyze Moses as the political founder of the Hebrew nation, and Spinoza begins his analysis of Moses’ rule (TTP 17) by referring to *love* and fear – fundamental Machiavellian political categories (e.g, Machiavelli 2009, chap. 17) – not Spinoza’s standard paring of *hope* and fear (TTP Preface, 3DefAffXIIIExp, and 3P18). Prophet and politician have the same power to bind together the community because they deploy

---

20 Scholars of Republicanism support this Machiavelli–Spinoza connection, seen more clearly in the *Discourses on Livy* (Machiavelli 2003) than in the *Prince*. On Machiavelli’s Republicanism, see numerous works by Pocock and Skinner. Specific connections of Machiavelli and Spinoza can be found in (Lucchese 2009; Mulier 1993) as well as multiple pieces by Morfino.

the same highly affective formative techniques, producing shared fears, hopes, loves, and hatreds.

To restate, prophets have a political effect through imagination which modifies the affects of the people. Therefore, the truth of prophecy is not relevant. Affects felt by prophets and conveyed to the multitude with intensity are most effective at modifying real desires and real behaviors. The aim of these modifications is obedience only, defined as making individuals' ideas, desires, and behaviors just similar enough that they do not conflict. Politicians and prophets are nearly identical in their use of affects to instill obedience.

All of the same can be said of artists – except that artists do not produce obedience, piety, or “morality”. It is precisely on this point that artists diverge from prophets and politicians. Artists' transmitted affects lead to other desires and other behaviors, thus differentiating their followers from the nation rather than reinforcing conformity to the nation (i.e., obedience.) In short, politics exposes the generic multitude to the same affectively-charged ceremonies to secure national unity; art forms smaller groups or “sub-genres” within the nation which, insofar as they are exposed to other ceremonies, have different behaviors based on different affects. If it is fair to refer to politicians as armed prophets, and thus to prophets as unarmed politicians, then we can say that artists are “amoral prophets” since they deploy affect-based group-making practices but lack the “moralizing” or universal content.

## **SECTION 2D: The Necessity of Public Ceremonies for Causing a Shared Way of Life**

There is no people without rituals or ceremonies (repetitive social practices), Spinoza tells us. As Spinoza notes in a chapter devoted to ceremonies, “these ceremonies served to establish and preserve the Jewish state” (TTP 5.1), creating a new people from almost nothing. “As soon as [the Hebrews] departed from Egypt, they [...] had the freedom to enact new laws or make new rules as they pleased and to establish a state [...] However, they were not in any way fit to make laws wisely or organize a government [...] for they were all [...] down-trodden by the miseries of slavery [and] obstinate” (TTP 5.10), meaning they were disobedient. And yet they became a people, and even a successful people for a time. Spinoza says that, among other causes, “a unique system of morals and worship [...] served to harden the minds of the Hebrews in bearing all things with singular constancy and courage on behalf of their country” (TTP 17.24) – a singular constancy which they, in their obstinacy, lacked prior to Moses' institution of new practices. Ceremonies (and laws) gave them group cohesion, but also implanted specific new virtues, capabilities, or powers within them.

Politics should not only discourage behavior through fear of punishments, but also encourage certain affects such as devotion (TTP 5.8). “This is why Moses [...] introduced religion into the commonwealth, so that the people would do its duty more from devotion

than from fear” (TTP 5.11). Like all people, the Hebrews’ group cohesion was strengthened to the extent that they had reason to desire or love that community rather than only fearing it. Lacking a nation until that point, they of course lacked a love of nation until that point: ceremonies (with a distinctly aesthetic tenor) helped cause that desire, and the unity that came from it. Ritual accompanied every aspect of their lives: “[Moses] did not permit them, accustomed as they were to slavery, to do anything at their own pleasure [...] They were not permitted to plough or sow or reap as they pleased, nor could they eat or dress or shave their heads or beards as they pleased, but all in accordance with a fixed and specific ordinance of the law. Not only that, but they were obliged to have certain symbols on their doorposts, in their hands and between their eyes, to remind them continually of their obedience” (TTP 5.11). While acting solely at the pleasure of the legislator is not the freest status, neither was the post-slavery status of the Hebrews: they could neither tolerate full freedom (since they were unaccustomed to it), nor manifest bondage. Spinoza gives a religious “festival” as an example which can lead to “cheerfulness [and] honest enjoyment”: “I do not think anything can be devised which is more effective than this for swaying men’s minds. Nothing captivates minds more effectively than the cheerfulness arising from devotion, i.e., from love and wonder together. They were unlikely to become bored with it” (TTP 17.25) since wonder results from novelty, holding the attention fast (3P52). Images in private and in public; clothing on, food in, and utterances issuing from one’s own body; novelty to captivate and joy to cause love – by these affective means, ceremonies gave shape to a people.

Ceremonies are so powerful at uniting a group that they can maintain one even once national independence has been lost. Spinoza notes that “the Jews were scattered everywhere for so many years [...] and yet they still survived, as no other nation has” (TTP 3.11). What others call God’s special election of the Hebrew people, Spinoza explains as the effect of shared ceremonial symbols of belonging. Circumcision, for instance, stimulated affects required for social cohesion; the Jewish people were maintained “not only because of their external rites [...] but also by the sign of circumcision which they zealously maintain [...] I think the sign of circumcision has such a great importance as almost to persuade me that this thing alone will preserve their nation for ever” (TTP 3.12). Signs of belonging do not merely represent or symbolize a people’s unity or cohesion: they actually make possible and support that cohesion.

To repeatedly modify the behavior of a people is to reshape them over time: repeated similar affectation causes habit or disposition, and can increase the ability to act in that particular manner. Spinoza says that if our bodies’ “softer parts” are “often” impacted by external bodies, then this leads to a “change” that renders us more liable “than [...] before” to think in certain ways or undergo certain affects: “So although the external bodies by which the human body has once been affected [are no longer present], the Mind will still regard them as present, as often as this action of the body is repeated” (2P17Dem). That is, if we have often interacted with a certain object while experiencing a particular affect, future interactions with that object will reproduce that affect even if we would not otherwise experience it.

Repetitive rituals take advantage of this potential power, shaping our soft parts in ways that leave permanent traces after the ritual ends. Spinoza treats repetitive ceremonies almost like athletic or even military training. The Hebrews were “so nourished and *inflamed* by *daily* worship that it must have *become second nature*. For their daily worship was [...] completely different (which made them *altogether unique and utterly distinct* from others) [and was] bound to generate a *ceaseless* [repetition of affects] firmly entrenched in their minds” (TTP 17.23; emphasis added). Daily repetition is absolutely central here, causing intense affects, long-lasting changes, and significantly distinguishing the group from others. Ceremonies gain cumulative affective force through repetition. There are other examples which show how repetition increases the force of affects. Sacrifices also were effective since “they had been accustomed to sacrifices from childhood [...] which hence stimulated their devotion more than anything else” (TTP 5.5). That is, lifelong repetition allowed the tradition to stimulate the affect of devotion the most, because it can draw on a lifetime of memories as affective resources.

The repetitive nature of ceremonies eventually leads to an automatic affective response which was not possible before. Althusser was fascinated by these passages, and linked them to what he called Pascal’s “theory of the apparatus of the body: ‘Kneel and pray [and only then can you believe], [...] (see what Michel Foucault appropriately calls the ‘disciplines of the body’ in the seventeenth century; they have obviously not disappeared since) [...] The TTP [similarly reveals] the *materiality of the very existence of ideology*” (Althusser 2008, 3 and 10). According to Althusser, the state uses these methods in order to “interpolate” an ideology into its subjects. However, if the *method* (repetition) is what is truly indispensable to ceremonies’ effectiveness, then the *content* need not be ideology or obedience. Artists can harness the same methods of ceremonies (their repetitive and social nature) while supplying other affective content.

## **SECTION 2E: Artists Give Shape to a Sub-Genre: Alternative Affects Founding an Alternative Way of Life**

My claim is that, much as political ceremonies can produce a “second nature” (TTP 17.23) in their followers, artists can produce a “third nature” in their audience. When we are devotees of a genre consisting of similar bodies of work, repeated exposure to these artworks leaves traces in our soft parts as it moves us. However, there is this key difference: artwork’s lack of uniform “moral” content leading to obedience. Morality in Spinoza requires consistent obedience to laws and customs that keep all of society together through similarity. The content of art, however, can refine a niche within the nation, a sub-genre of people as it were, by affirming a difference or a mark of distinction (like circumcision did for the Hebrews). A genre of art repeats the ceremonies and customs of its sub-group, over time refining the group’s shared divergence from the nation at large.

I think, of all the arts, this is especially true for music, since public performance is more central to music than to most forms of art. Kinds of music and kinds of people (even whole

ways of life) are often identified: hip-hop, goth, punk, Rastafarian, rave(r)s, metal(-heads) – but not so for visual arts (what art lover identifies as an “impressionist” or “abstract expressionist”?), for poetry (can a person be “epic”?),<sup>21</sup> and so forth, because these arts as currently practiced have fewer public ceremonies. More so than any other art (other than perhaps performance art), live music concerts involve a “concerted” performative endeavor on the part of the audience. Audience members undergo many simultaneous affects: they are always listening at the same time, often singing and dancing together, and possibly surrounded by similarly dressed people. As an essentially group experience, as a ceremony, the repetition of a specific sub-genre of live music would be the form of art best able to bind a small group today by reshaping their bodies and giving them a “ceaseless” or “singular constancy” (17.23–24) in affects unique to that group.

Nobody is born a punk, a raver, and so on. The process of becoming punk includes repeatedly suffusing one’s soft parts with the music, adjusting one’s dress and hair, and following the ritual of slam-dancing at basement shows. The sub-genre’s motions become ever more a part of the individual as initial affects of pleasure become a new habitual comportment which distances them ever more from the generic multitude at large. Interaction with art one time would cause merely temporary affects, but repeated exposure shapes us more permanently. Performing one sacrifice was not enough to establish the Ancient Hebrews: but doing so for a lifetime is part of what constituted each of them as members of that group. Similarly, one all-night dance party in a warehouse or one basement show does not produce a raver or a punk. However, I would claim that nobody becomes a raver without repeating the shared “festival” of all-night dancing (TTP 17.25), nor a punk without the “dress” of torn and dirty denim (TTP 5.11).<sup>22</sup> Nobody is born a devotee or adherent to a musical sub-genre: rather, it will come to include her over time, as she establishes new affective capacities that she initially lacked. Adherence to the discipline of the musical sub-genre eventually will produce a new and improved *ingenium* for her, which she will share with others in the sub-genre but will not share with the nation at large. That is, members of a niche or sub-genre of people become similar to each other – which distinguishes them from everybody else. As those in the genre repeat their distinctive practices, as they move and mark their bodies in similar ways (TTP 3.12), they increase their distinction from the multitude and gain their own atypical powers.

Exposure to a variety of art gives healthy pleasure; but sometimes, repeated exposure to the same type of artistic ceremonies (such as many live musical performances of the same style) has effects beyond health, beyond merely replacing worn-out prior parts. It then introduces new affects (and perhaps even new parts) and differentiates you from your past self, reducing your conformity to the generic national culture. Art causes pleasure by “refreshing” us and is useful for healing the bodies which we already have. In addition, if

---

21 Of course, non-musical arts can also produce new ways of life: e.g. Beat poets, Situationists, the avant-garde generally, dandies, etc. Perhaps pre-classical “Homeric culture” indeed was “epic” (though if so the role of public performance and regular repetition surely played a decisive role).

22 Musicology, sociology, and anthropology could provide concrete details on identity formation within group settings. A more detailed Spinozist account would refer to the imitation of affects.

we repeatedly undergo exposure to the ceremonies of a specific musical genre, this whole-person practice can reform our whole selves and refine our difference from the generic multitude. It can impel us in a direction that others do not take, and it can accelerate our speed in that direction.

## References

- Althusser, Louis. 2008. The Only Materialist Tradition. Part I: Spinoza. In: Stolze, Ted – Montag, Warren (eds.): *The New Spinoza*. Minneapolis: University Of Minnesota Press. 3–19.
- Ansaldi, Saverio. 2003. Love, Perfection, and Power in Spinoza. *Graduate Faculty Philosophy Journal* 24 (2): 59–74.
- Aristotle. 1984. Poetics. In: *The Complete Works of Aristotle*. Edited by Jonathan Barnes. Translated by Ingram Bywater. Revised Second Edition. Vol. 2. Princeton: Princeton University Press.
- Bal, Mieke – Vardoulakis, Dimitris. 2011. An Inter-Action: Rembrandt and Spinoza. In: Vardoulakis, Dimitris (ed.): *Spinoza Now*. Minneapolis: University of Minnesota Press. 277–303.
- Balibar, Étienne. 1997. *Spinoza: From Individuality to Transindividuality*. Delft: Eburon. (Mededelingen vanwege het spinozahuis, 71)
- Balibar, Étienne. 1998. *Spinoza and Politics*. Translated by Peter Snowdon. London: Verso.
- Calder, Simon. 2015. George Eliot, Spinoza and the Ethics of Literature. In: Lord, Beth (ed.): *Spinoza Beyond Philosophy*. Edinburgh: Edinburgh University Press. 168–187.
- Cimini, Amy. 2015. The Secret History of Musical Spinozism. In: Lord, Beth (ed.): *Spinoza Beyond Philosophy*. Edinburgh: Edinburgh University Press. 87–107.
- Colerus, John. 1706. *The Life of Benedict de Spinoza*. London: D.L.
- Danto, Arthur. 2001. Postmodern Art and Concrete Selves: The Model of the Jewish Museum. In: *Philosophizing Art: Selected Essays*. Berkeley: University of California Press. 125–143.
- Davidson, Christopher. 2015. Spinoza as an Exemplar of Foucault’s Spirituality and Technologies of the Self. *Journal of Early Modern Studies* 4 (2): 111–146.
- Deleuze, Gilles – Félix Guattari. 1994. *What Is Philosophy?* Translated by Hugh Tomlinson and Graham Burchell. New York: Columbia University Press.
- Deleuze, Gilles. 1994. The Image of Thought. In: *Difference and Repetition*. New York: Columbia University Press. 129–167.
- Deugd, Cornelis de. 1966. *The Significance of Spinoza’s First Kind of Knowledge*. Assen: Van Gorcum.
- Egenhofer, Sebastian. 2011. Image and Machine: Introduction to Thomas Hirschorn’s Spinoza Monument. In: Vardoulakis, Dimitris (ed.): *Spinoza Now*. Minneapolis: University of Minnesota Press. 237–261.
- Foucault, Michel. 2009. *History of Madness*. Edited by Jean Khalfa. Translated by Jonathan Murphy and Jean Khalfa. London: Routledge.
- Garrett, Don. 1994. Spinoza’s Theory of Metaphysical Individuation. In: Barber, Kenneth – Gracia, Jorge (eds.): *Individuation and Identity in Early Modern Philosophy*. Albany: SUNY Press. 73–101.

- Gatens, Moira. 2012. Compelling Fictions: Spinoza and George Eliot on Imagination and Belief. *European Journal of Philosophy* 20 (1): 74–90.
- Gatens, Moira. 2015. Mark Sacks Lecture 2013: Spinoza on Goodness and Beauty and the Prophet and the Artist. *European Journal of Philosophy* 23 (1): 1–16.
- Hübner, Karolina. 2014. Spinoza on Being Human and Human Perfection. In: Kisner, Matthew – Andrew, Youpa (eds.): *Essays on Spinoza's Ethical Theory*. Oxford: Oxford University Press. 124–142.
- Hübner, Karolina. 2016. Spinoza on Essences, Universals, and Beings of Reason. *Pacific Philosophical Quarterly* 97: 58–88.
- Huysmans, J.K. 1998. *Against Nature*. Translated by Margaret Mauldon. Oxford: Oxford University Press.
- Israel, Jonathan. 2007. Introduction. In: Israel, Jonathan (ed.): *Theological-Political Treatise*. Translated by Michael Silverthorne and Jonathan Israel, Cambridge, UK: Cambridge University Press. viii–xxxiv.
- James, Susan. 2010. Narrative as the Means to Freedom: Spinoza on the Uses of Imagination. In: Melamed, Yitzhak – Rosenthal, Michael (eds.): *Spinoza's Theological-Political Treatise: A Critical Guide*. Cambridge: Cambridge University Press. 250–267.
- Lucchese, Filippo Del. 2009. *Conflict, Power, and Multitude in Machiavelli and Spinoza: Tumult and Indignation*. London: Continuum.
- Liotard, Jean-François. 1994. *Lessons on the Analytic of the Sublime: Kant's Critique of Judgment*. Sections 23–29. Stanford: Stanford University Press.
- Machiavelli, Niccolò. 2003. *Discourses on Livy*. Translated by Julia Conaway Bondanella and Peter Constantine. Oxford: Oxford University Press.
- Machiavelli, Niccolò. 2009. The Prince. In: *The Essential Writings of Machiavelli*. Translated by Peter Constantine. New York: Random House Publishing Group.
- Marshall, Eugene. 2014. *The Spiritual Automaton: Spinoza's Science of the Mind*. Oxford: Oxford University Press.
- Mignini, Filippo. 1981. *Ars imaginandi: apparenza e rappresentazione in Spinoza*. Naples: Edizioni scientifiche italiane.
- Miller, Jon. 2015. *Spinoza and the Stoics*. Cambridge: Cambridge University Press.
- Morrison, James. 1989. Why Spinoza Had No Aesthetics. *The Journal of Aesthetics and Art Criticism* 47 (4): 359–365.
- Mulier, Eco Haitsma. 1993. A Controversial Republican: Dutch Views of Machiavelli in the Seventeenth and Eighteenth Centuries. In: Bock, Gisela – Skinner, Quentin – Viroli, Maurizio (eds.): *Machiavelli and Republicanism*. Cambridge: Cambridge University Press.
- Nadler, Steven M. 1999. *Spinoza: A Life*. Cambridge, UK: Cambridge University Press.
- Ravven, Heidi Morrison. 2004. Spinoza's Ethics of the Liberation of Desire. In: Tirosh-Samuelson, Hava (ed.): *Women and Gender in Jewish Philosophy*. Bloomington: Indiana University Press. 78–105.
- Rawes, Peg. 2015. Spinoza's Architectural Passages and Geometric Compartments. In: Lord, Beth (ed.): *Spinoza Beyond Philosophy*. Edinburgh: Edinburgh University Press. 66–86.

- Spinoza, Benedict. 1985a. Ethics. In: *The Collected Works of Spinoza*. Vol. I. Edited and translated by Edwin M. Curley. Princeton: Princeton University Press.
- Spinoza, Benedict. 1985b. Treatise on the Emendation of the Intellect. In: *The Collected Works of Spinoza*. Vol. I. Edited and translated by Edwin M. Curley. Princeton: Princeton University Press.
- Spinoza, Benedict. 1994. Ethics. In: *A Spinoza Reader: The Ethics and Other Works*. Edited and translated by Edwin M. Curley. Princeton: Princeton University Press.
- Spinoza, Benedict. 1995. *The Letters*. Edited by Steven Barbone, Lee Rice, and Jacob Adler. Translated by Samuel Shirley. Indianapolis: Hackett Publishing.
- Spinoza, Benedict. 2000. *Political Treatise*. Translated by Samuel Shirley. Indianapolis: Hackett Publishing.
- Spinoza, Benedict. 2007. *Theological-Political Treatise*. Edited by Jonathan I. Israel. Translated by Michael Silverthorne and Jonathan I. Israel. Cambridge, UK: Cambridge University Press.
- Spinoza, Benedict. 2016. Theological-Political Treatise. In: *The Collected Works of Spinoza*. Vol. II. Edited and translated by Edwin Curley. Princeton: Princeton University Press.
- Spruit, Leen – Totaro, Pina. 2011. *The Vatican Manuscript of Spinoza's Ethica*. Leiden: Brill.
- Uhlmann, Anthony. 2011. Spinoza, Ratiocination, and Art. In: Vardoulakis, Dimitris (ed.): *Spinoza Now*. Minneapolis: University of Minnesota Press. 263–275.
- Vinciguerra, Lorenzo. 2005. *Spinoza et le signe: la genèse de l'imagination*. Paris: Vrin.
- Vinciguerra, Lorenzo. 2012. Mark, Image, Sign: A Semiotic Approach to Spinoza. *European Journal of Philosophy* 20 (1): 130–144.

# XIII

*Zsolt Bagi*

## Emancipation of the Body in Spinoza's Political Philosophy. The Affective Integration

### 1. Empowerment in Spinoza

#### **Spinoza's philosophy is a philosophy of emancipation**

It is easy enough to notice that the ultimate aim of the *Ethics* is to reach personal freedom from passions (passive affects) through reason, while the ultimate aim of political philosophy is to reach freedom of the community through a rationally instituted and directed state (“*quae ratione fundatur et dirigitur*” [Geb III. 287]). However, it is rather difficult – for the contemporary interpreter – to establish a firm correlation between these goals and Spinoza's evident hostility toward the concept of free will. One must realize that personal freedom and the freedom of the community means first and foremost autonomy in his philosophy and by no means free will. Autonomy and free will being two very different, indeed contrary concepts for Spinoza, one has to clarify what he means by autonomous, free action.

It would be just as misleading to simply say that Spinoza uses a scholastic concept – that of spontaneity – when he speaks about freedom, in which case necessity and freedom are not mutually exclusive, but rather mutually constitutive parts of the concept. Luis de Molina for example distinguishes three meanings in the concept of freedom (*libertas*). According to the first meaning it is opposed to servitude, but Molina rules this meaning out, stating that he is concerned only with the meaning which is based on the word “he/she likes to” (*libet*). This restricted concept in turn can be differentiated as well according to two different meanings: spontaneity and free will. According to the first (which concerns us here, although for Molina this meaning has to be discarded as well) something happens spontaneously (as opposed to by constraint [*in oppositur coactioni*])

“whether it is happening by natural necessity or not (*sive fiat naturali necessitate, sive non*)”. This is similar to the Father generating the Son freely (without constraint) but also necessarily (Molina 1876, 10). In fact, Spinoza’s concept owes a great deal to this interpretation of freedom. He writes in the 56<sup>th</sup> letter to Boxel: “That the necessary and the free are two contraries seems no less extravagant and contrary to reason. For no one can deny that God knows himself and all other things freely; nevertheless everyone grants, by common agreement, that God cannot fail, or cease, to know himself”. Also in the 58<sup>th</sup> letter to Schuller: “You see, then, that I place freedom not in a free decree, but in a free necessity”. Nevertheless his concept of freedom is far from being identical to that of scholastic spontaneity. The definition of freedom in the *Ethics* 1D7 is based on scholastic spontaneity, but it is only a nominal definition. The real definition of freedom can only be found in the 5<sup>th</sup> part of the *Ethics*.

Freedom’s real definition for Spinoza – as well as Leibniz – is a genetic one. It requires (in case of a created thing) the “proximate cause” (Spinoza 1985a, 39). For example in the case of a circle it is not enough to define it as “a figure in which the lines drawn from the center to the circumference are equal”. The proximate cause is also required: “it is the figure that is described by any line of which one end is fixed and the other movable” (Spinoza 1985a, 39–40). To define a free thing it does not suffice to say (E1D7) “that thing is called free which exists from the necessity of its nature alone, and is determined to act by itself alone” (Spinoza 1985b, 409). It is also required to include its proximate cause. In case of the rational way of life this proximate cause of a free subject is the detachment of the passion from its accidental cause by the way of the intellect, thus forming an active affect out of a passive one. It is not an exaggeration to say that the whole process of philosophical machinery included in the *Ethics* serves the goal of reaching a “perfect definition” – that is a real definition providing the proximate cause – of “the power of the intellect”, in other words “human freedom”. Spinoza is in accord with Pascal (and the *Port Royal Logic* which is based on Pascal’s view) that a real definition is in fact a theorem and it has to be proved. Providing the proximate cause for a thing so complex as the free human being requires an ethics. According to the fifth part of the *Ethics*, freedom consists – for the rational way of life at least – in gaining power (*potentia*) by way of detaching the affectivity of the human consciousness (*mens*) from the accidental external causes. Freedom is possible due to the fact that passive affects can be eliminated by forming a clear and distinct idea of them (E5P3).

The fact that freedom is based on a “power” of the consciousness (*potentia mentem*), and at the same time freedom by empowerment is also a subject for Spinoza’s political writings, leads us to the question: is there a general concept of freedom in Spinoza? Is political freedom related to ethical freedom? The reason why I would use the clearly anachronistic term “emancipation” referring to Spinoza’s concept of freedom is mostly because his concept is based on empowerment of the subject. As a general rule a free subject is an empowered

subject to Spinoza.<sup>1</sup> *Potentia* is a key concept when we think of human freedom even in the *Ethics*, although [as Antonio Negri has shown (Negri 1991, 2004)] it became well defined only in opposition to *potestas* in Spinoza's writings on political philosophy. In the fifth part of the *Ethics* the freedom that an intellectual way of life can achieve has clearly nothing to do with political freedom, nevertheless it is formulated in a general way; freedom is achieved by the power of human consciousness instituting the *imperium* over affects (E5Praefatio). That is, not by opposing *potentia* and *potestas* (in this case *potentia* and the *imperium* of the human *mens* over affections), but rather treating them as correlates. This formulation links the *Ethics* to Spinoza's political philosophy and at the same time differentiates it. Freedom – be it ethical or political – is empowerment, but ethical and political empowerment are two different processes.

## 2. The Autonomy of the Affective World

While it is evident from the 5<sup>th</sup> part of the *Ethics*, that through reason one can achieve freedom – meaning that *potentia rationis* has the possibility of detaching the human *mens* from the accidental external causes, thus instituting the *imperium* of the reason over passions – it is less evident in the case of the affective world. Affectivity for Spinoza constitutes a strictly parallel world to reason. It is based on imagination, the first form of knowledge (according to E2P40S2), and thus has a subordinate role in understanding, since it is impossible to reach adequate ideas by them. At the same time the world of affects has its own territory, different from and not subordinated to understanding. Consciousness is not always understanding, if we mean by understanding the representation – adequate or inadequate – of the essence of a thing in consciousness. Other than things or essences of things, we represent events occurring to the body or as Spinoza calls them: affections (*affectiones*). According to E3D3 he calls affect both the event and the representation of the event. The event moreover can be characterized intrinsically by a change of *potentia*: an event occurring when the acting power of a body increases or diminishes. Affects are representations of the events usually accompanied by the representation of some kind of

---

1 Moreover this philosophy is emancipatory for it is based on the equality of the subjects. To establish the reign of reason is to establish the state of the empowered subjects based on either the understanding “what is common to all things” (E5P4D and E2P38) or the unification of their affective bodies into one rationally instituted common body, the commonwealth. Although the unification of bodies does not involve an equality of rights it involves equal accessibility of the power. The one hand *potentia* and the natural rights are the same thing to Spinoza (PT2.2) and natural rights remain in effect even in the state. Thus the more power someone has to act, the more rights she has as well. On the other hand this difference is based on the equality of the subjects to gain power: an individual is under control of another so long it is under its power [*potestas*] (PT2.9), but it is always possible to rise to power. Since power to act [*potentia*] is not something we are born with, it is dependent on our will and ability to cooperate with similar individuals. Thus Spinoza's philosophy promotes both equality and empowerment, a clear example of an emancipatory engagement.

(accidental) cause. “By affect I understand affections of the Body *by which the Body’s power of acting is increased or diminished*, aided or restrained, and at the same time, the ideas of these affections” (Spinoza 1985b, 493. Emphasis added).

The affects are naturally ideas of consciousness. From the point of view of understanding, these representations are necessarily inadequate, since they do not represent any essence, only changing values of the “power of acting” (*potentia*). On the other hand imagination and the affective world it is based on is not a mere fiction. Although in the case of understanding, it is subordinate to reason, nevertheless it is a possibly autonomous form of being in the world. We are about to establish this claim on textual evidences.

The possible autonomy of the world of affections is proved by Spinoza in E3P15: “Any thing can be the accidental cause of Joy, Sadness, or Desire” (Spinoza 1985b, 503). Generally speaking Spinoza is opposed to the concept of “accidental cause”, stating that God is never an accidental cause and since “everything is in God”, nothing has an accidental cause. In the world of affects (since “Joy, Sadness, Or Desire” are the most basic affects that provide ground for any affect), this is not the case. To put it in another way, although Spinoza is generally regarded as the thinker of necessity, and while this is undoubtedly true that Nature has a strict causality, excluding even possibility, there is one major exception to this rule: the world of affections, where not necessity but accidental causes reign. Since the problems of this world comprise Spinoza’s political and theological thought, which makes up about half of his *oeuvre* (including most of the third and fourth part of the *Ethics*), it is clearly a simplification to call Spinoza the philosopher of necessity. This domain of “contingence” is not excluded from the philosophy; the claim of a reign of accidental causes by no means leads to the alleged conclusion that this world would not have any philosophically observable laws. Its laws are forms and variations of sympathy and antipathy (E3P15S) and they are legitimate subjects for rational analysis. At the same time they are not “rational” themselves (if one calls “rational” the second form of knowledge as in E2P40S2).

What we have to add here to avoid misunderstandings is the assertion that the possible autonomy of the world of accidental causes does not provide any possibility for free will at all. For example having an accidental cause for action does not makes this action “free” in the eyes of Spinoza. It is rather the opposite; freedom becomes problematic in a world that lacks necessity (see PT2.7). As we have seen freedom (as autonomy) is “free necessity”, that is when something “exists from the necessity of its nature alone”. Accidental causes are external causes that do not follow from the nature of the thing they are the cause of. This is why autonomy never can be *fully* established in the affective world, and this is why affectivity has to be integrated into a single affective cooperation, the commonwealth.

Consciousness is submerged into two worlds, one rational and one affective; but there is another difference between these two, and this could shed some light on the reason why I use the term “world” in Spinoza’s case. While it is possible for the understanding to leave the determinate perspective of the human body by way of “common notions” (E2P38–39), it is impossible for the passive affects (passions). Since passions are ideas about events of the body, and they are nothing more than the representations of the change of *potentia* of the body by these events (thus they are not notions, representations of an essence of

a thing, and they are unthinkable without reference to the body they are events of), they are always bound to the perspective of the body. The passionate body is a different thing than the physical body, the *ideatum* of the understanding of the *mens* (as formal essence represented by objective beings, ideas). The passionate body is represented in the *mens* by the way of passive affects, and it is in fact nothing more than a series of events, a change of power levels with regard to a peculiar point of view: the body. Passions, as opposed to ideas of understanding, are essentially representations determined by the body – there are no passions that are “common to, and peculiar to, the human Body and certain external bodies by which the human Body is usually affected, and is equally in the part and in the whole of each of them” (E2P39). The purely affective world is a bodily world.

Now there is a possibility for the understanding to turn the passionate world into an active world, the world of active affections. According to E5P4: “There is no affection of the Body of which we cannot form a clear and distinct concept” (Spinoza 1985b, 598). The demonstration of this proposition relies on the above mentioned common notions. Forming a clear and distinct concept of the passion (passive affect) means precisely that we obtain a concept out of a confused idea that originally was not a concept or notion. It becomes a representation of an essence instead of a representation of an event (rising or diminishing of *potentia*). Moreover this possibility – and this is a particularly delicate piece of the *Ethics* that we cannot trace here fully – does not change the thing itself; the representation of the affection of the body remains the same, only it becomes an idea of an idea, that is an affection represented from a different point of view. According to E5P3Dem, the idea we form of the affection “will only be distinguished by reason [*non nisi ratione distinguetur*] from the affect itself, insofar as it is related only to the Mind [consciousness] (by IIP21 and P21S)” (Spinoza 1985b, 598). That is to say between the passion and the active affection of the consciousness there is no other distinction than a rational one. Rational distinctions (as opposed to real distinctions) are distinctions not between different things but between different points of view. Passions and active affects (affects based on understanding) have the same affection of the body as their *ideatum*, but they look at it from a different point of view. One is an idea of an event, the other is an idea of an idea (E2P21 and E2P21S), meaning: an idea of an essence of an idea of an event (the reflexion on a passion and the understanding of its essence).

That is to say it is quite true that by understanding our passions we can empower ourselves and free ourselves from them (we can gain *imperium*, rulership over them), but at the same time we change a perspective and simply leave the body (the individuality of ourselves) behind. This would not be the empowerment of the body but its abandonment. This is why the emancipation of the body, the empowerment of the body (if it’s possible at all) must always be affective: the rational way of life leaves the perspective of the body behind, the affective way of life remains in it. These two points of view constitute two worlds (in a quasi-phenomenological meaning of this term), two systems of representations where the “*ordo et connectio*” of the ideas differ from each other. Spinoza calls the first one the point of view of eternity. By contrast we could refer to the latter as the point of view of the events. Rational and affective worlds differ from each other even though they represent the same thing.

Let us note that although the rational world also has affects (representations of changes in the *potentia*), these affects are subordinate to understanding and have no autonomy, and accordingly I will not count them in the affective world. The affective world is a world of passions, but “passion” for Spinoza is a relative term. It is different from active affections in an absolute way, but also differentiated in itself relatively. The passive affects are differentiated among themselves by degrees of power. E3P9 states “[b]oth insofar as the Mind [consciousness] has clear and distinct ideas, and insofar as it has confused ideas, it strives, for an indefinite duration, to persevere in its being and it is conscious of this striving it has” (Spinoza 1985b, 499). Therefore every affect has a different value for the consciousness, according to whether it is helping or hindering this effort. Moreover it is conscious of this value: in fact joy and sadness are the ideas (affects) that represent this value (E3P11S). The affective world is a world of changing values of the passions, where (imaginary) power of the given affect determines the actions the affective consciousness takes. It strives (as of E3P7) to maintain its essence (not just its existence) and its ability to produce, thus it strives to raise its power to act (freedom), even if it is impossible for the affective body to reach complete freedom. Therefore there are more active (better) passions and there are more passive (worse) passions.

After this quick characterization of the two worlds that constitutes human consciousness, the autonomy of the later remains to be argued. One could think that Spinoza as a rationalist (and indeed an extreme rationalist) should have treated the irrational affective world as irrelevant or reducible, but clearly this is not the case. First of all it has been demonstrated by Alexandre Matheron (Matheron 1988, 81 sqq), and later by Étienne Balibar (Balibar 1998, 76 sqq) that the third and fourth part of the *Ethics* present a parallel series of propositions for these two worlds, a process that concludes in E4P37 where two different demonstrations (one for the rational one for the affective way of life) prove the necessity of a commonwealth. The two worlds are – if not equally, at least proportionally – important for Spinoza even in the *Ethics*. Moreover in the *Political Treatise* he writes that whoever thinks that a rational way of life can be established for the majority of the people is dreaming: “So people who persuade themselves that a multitude, which may be divided over public affairs, can be induced to live only according to the prescription of reason, those people are dreaming of the golden age of the Poets” (PT 1.5, Spinoza 2016, 506). The political theory in fact wishes to demonstrate that an empowerment of the subject is possible even if the subject is not willing to follow the way of the reason. “So, a state whose well-being depends on someone’s good faith, and whose affairs can’t be properly looked after unless the people who handle them are willing to act in good faith, won’t be stable at all. For it to be able to last, its affairs must be so ordered that, whether the people who administer them are *led by reason or by an affect*, they can’t be induced to be disloyal or to act badly” (PT 1.6 My emphasis, Spinoza 2016, 506). The affective world has to have a rationally organized frame, the state, to establish the possibility of an integration of the bodies. But the integration itself does not follow the rules of the reason, but that of the affective world: sympathy and antipathy. It is driven by the *conatus*, the strive to increase the power (*potentia*) of the body, the empowerment of the body. Although the best way of elevating one’s own power

of acting is the way of the reason, most of the people (one could argue even a philosopher in her ordinary life) would never see that and remain in the affective world, striving to emancipate themselves in that world.

It is Spinoza's political philosophy where one can find the clues to reconstruct a never fully explicated other form of emancipation, that of the affective body. To reach autonomy through affections one must get rid of the determinations of the body in a similar way reason frees itself from them. However emancipation from determinations in this case does not involve understanding but an integration of the bodies of a *multitude* into one common body, the commonwealth. The nature of this affective integration is the subject of my inquiry. The affective integration differs both from its predecessors (Hobbes' leviathan) and later forms of integrative social unity of the enlightenment (like Rousseau's *volonté general*), and offers a form of social integration where – as we will see – differentiation plays just as important a role as unification.

### 3. The Integration of the Body

The aim of integration (meanwhile also the foundation) is the *individual* body. The commonwealth is in fact an individual body itself. Since it is an affective individual and not a physical one, it poses special problems, but the same general laws pertain to it as to any individual. What needs to be examined is how the multitude of bodies can be integrated into a single one. Naturally we have to take our point of departure in the individual body in general: the human body is just a particular case of the former. One crucial point to be stressed however, is that the integration of bodies in general, and the integration of human bodies into one single empowered community (the commonwealth, *civitas*) in particular, does not involve – in either case – the subordination of the single to the whole. Just as individual bodies uniting to be the common cause of a singular effect do not necessarily become a body with an essential form or *eidōs*, the citizens of a state do not suppress their own individual essences either when they constitute a state. The community of bodies (the commonwealth) is integrated not because its subjects undergo the process of unification by the one and only general will, but because they constitute the common body and unified soul by a process where every body (every citizen) contributes with what is its most peculiar propriety, the very thing that can be part of the community only by that particular body. Every single body universalizes its own particularity, and not the other way around.

The foundations of individuality of the body – that is, when one can call a body a single entity and not just an aggregate – is treated in the fourth to seventh lemmas of the second part of the *Ethics*. According to a general definition of the singular thing in the seventh definition of the second part: “By singular things I understand things that are finite and have a determinate existence. And if a number of Individuals so concur in one action that together they are all the cause of one effect, I consider them all, to that extent, as one singular thing (*unam rem singularem considero*)” (E2D7, Spinoza 1985b, 447). For Spinoza,

being is none other than being a cause, or, in other words, being is production. When a single effect or product has more than one being or thing as cause or producer, then these beings in essence constitute one single being or thing.

There is – on the other hand – a specified definition of the individual body (as opposed to the general singular thing) in the second definition of the above mentioned excursus in the second part after the thirteenth proposition.

When a number of bodies, whether of the same or of different size, are so constrained by other bodies that they lie upon one another, or if they so move, whether with the same degree or different degrees of speed, that they communicate their motions to each other in a certain fixed manner, we shall say that those bodies are united with one another and that they all together compose one body or Individual, which is distinguished from the others by this union of bodies (E2D after E2P13, Spinoza 1985b, 460).

Alexandre Matheron has analyzed this passage in detail and I am recapitulating here his results. The individual can be characterized by an equation where the principle of individuality is represented by an invariant (K) and the addition of the parts of the body ( $C_1, C_2 \dots C_n$ ) equals K.  $K=C_1+C_2+ \dots C_n$ . Since the individual body or body part is characterized by “the definite proportion of movement and rest” (Matheron 1988, 38), the relation of the parts is a relation of these proportions, thus we can arrive to the full equation as

$$K = \frac{m_1 v_1 + m_2 v_2 \dots + m_n v_n}{m_1 + m_2 \dots + m_n}$$

where m is the mass and v is the speed of the body. The individual remains the same as long as one of its parts alters the magnitude or vector of its velocity or its mass while maintaining the same proportional relations with the rest of the parts; so, if the remaining parts alter their velocities or masses accordingly, for example: if one of the parts is completely destroyed, the individual remains the same if the other parts of its body can alter their velocities or their masses in such a way that the equation remains the same. In the case of the human body, the individual remains the same as long as the loss of functionality of one of its organs can be substituted by other ones, for example: the loss of eyesight can be substituted by haptic and auditory sensations.

However it is also evident from the formula – and this will gain major importance for us later – that the number of parts can be multiplied as long as the equation stands. For example: one can divide  $C_1$  into an arbitrary number of smaller parts as long as their combined proportion of movements and masses equals the original  $C_1$ . In this case the integrity of the individual remains uninjured. From the viewpoint of the integrity it doesn't matter if the task of the given  $C_1$  part is fulfilled by only one undivided part or a composite part organized on the basis of a multitude of parts. The only requirement is that the proportion of the parts must remain the same.

Integration, as it is the upholding of individuality, is self-definition. The body as an individual seeks to uphold its integrity during the various affections it suffers, i.e. it seeks to maintain the above equation. It seeks to be its own cause (*scausa sui*) in the most complete way possible. Therefore – and for Matheron this is the principal consequence – the body does not suffer the events in a passive way; it is active. When one component of the equation changes in accord with an event, the body takes action in such a way that its other parts would uphold its identity. In the singular case when a body can define itself completely (in case of the infinite body, the *facies totius universi* is what always remains the same), “the passivity disappears completely” (Matheron 1988, 48). Naturally in the case of determined beings, such self-definition can never be fully accomplished; nonetheless the body strives to act in such a way that its actions, not its determinations, would define it. Self-definition means that a body makes integrant components of its self out of the events it suffers by way of presenting them as its actions and not its passions. We can call this activity of the body integrating the events into the story of its own active life *affirmation*. In other words affirmation is the resituating of affectivity, the rewriting of the passive affections in the context of the productive self, thus transforming them into active affections. Affective integration is the empowerment of affectivity itself.

Now, from the above analysis one can reach another conclusion as well. One that not only Matheron did not draw,<sup>2</sup> but even Spinoza did not fully elaborate. Still it seems to me inevitable if one seeks to answer the question of how the emancipation of bodies is possible in Spinoza’s philosophy of affection. Instead of physical bodies in the case of the affective world, we deal with affective bodies where integration – just like anything else – is based on accidental causes: sympathy and antipathy. Without being able to analyze the *Political Treaty* due to space considerations, I must draw some outlines here. The goal of the affective synthesis (that is, the emancipation of the body) is to institute a community of bodies where they can cooperate freely. According to Spinoza, in its natural state, a human body does not have complete freedom. Rather, the opposite: it is determined by its own boundaries (PT 2.15). Achieving freedom in any other way than this understanding is only possible by community, by establishing a commonwealth. Although Spinoza has a deep aversion against the theory of social contract, and his political philosophy is rather a criticism of that of Hobbes’ than an affirmation, it is still the political state where the human body can accomplish its ambition of gaining power. From this fact it is evident that the individuality of a body is less important from the point of view of its freedom than its ability to cooperate (E4P37, PT2.15).

Thus the inquiry on the above mentioned conclusion will not concern the affirmation of self-definition of bodies, but will the possibility of their adaptation. As long as we take the bodies in question in isolation, we can reasonably say that the intrinsic law of their

---

2 For Matheron the difference between (1) an integrated complex, (2) a non-integrated complex, (3) an integrated simple and a (4) non-integrated simple being is natural: they are (1) the human body, (2) the society, (3) a stone respectively (Matheron 1988, 58) (the fourth one is omitted but would be a pile of stones, for example). In my opinion complexity is a changing value, and a better society – as I shall demonstrate – is more complex (more differentiated in itself) than a flawed homogenous one.

identities renders it sufficient to conceive of their parts in the simplest way possible. The less differentiated the parts are in themselves, the easier to maintain the equation. However an individual body – unless it is the infinite modus of extension (*facies totius universi*) – never stands alone, never in itself. Its parts become altered (in the case of a living body they are interchanged by the ingestion of other bodies), they make contact with other bodies, they age and become dysfunctional, and so on. These events can either help or hinder the efforts of the individual (they can augment or diminish its *potentia*, its ability to remain the same individual). In general one can say that a determined body, in order to maintain its individuality has to adapt in many ways to other bodies.

Yet in this case it is far from being indifferent in what degree the parts and the whole are differentiated in themselves. Let's suppose for example that a composite body meets another body that hinders the functioning of one of its parts. The actual essence of the given composite body – that is its *conatus* – prescribes for the parts of the body to alter their functioning in such a way that they can substitute the hindered part. The remaining parts must alter their functioning, that is – in Spinoza's terminology – their movements or their masses. Yet it is impossible to imagine doing so if the part in question is able to move only one way, that is, to fulfill only one task. The remaining parts must be differentiated in themselves sufficiently to be able to alter their respective tasks. The more complex a composite body is, the more easily it can adapt to external events. This is to say that according to Spinoza, unification or homogenization might improve the performance of a system of bodies in itself (the system would be simpler and thus easier to control), but it declines its adaptivity. That has at least two consequences for our inquiry: 1. In the case of a singular human body the differentiation makes it more apt to cooperate with other bodies instead of being hostile against them. Without adequate differentiation the only possibility would be to avoid or destroy the other body. 2. In the case of an empowered community of bodies, i.e. an emancipated society, the persistence of differences is desirable because they help the society to adapt to the changes in external demands. A differentiated society is more powerful than a homogenous one as long as it can affirm the differences.

Although it is hard to find textual evidence for such an assertion in Spinoza's *Political Treaties*, one can find such evidence in a different place: a *scholium* of the *Ethics*, where Spinoza mentions the question of the difference of the human soul in regard to other souls. This latter question involves the differentiation between various degrees of perfection among the consciousnesses, thus it addresses our problem, although admittedly in an indirect way. Our conclusion, what we have drawn from the nature of the individual body, was that the more complex a body is, the more adaptive it is. In other words: more complex bodies are more perfect as well – as long as they can remain *one* body, as long as they can affirm themselves.

[1] However, we also cannot deny that ideas differ among themselves, as the objects themselves do, and that one is more excellent than the other, and contains more reality, just as the object of the one is more excellent than the object of

the other and contains more reality. And so to determine what is the difference between the human Mind and the others, and how it surpasses them, it is necessary for us, as we have said, to know the nature of its object, i.e., of the human Body. [1a] I cannot explain this here, nor is that necessary for the things I wish to demonstrate. [2] Nevertheless, I say this in general, that in proportion as a Body is more capable than others of doing many things at once, or being acted on in many ways at once, so its Mind is more capable than others of perceiving many things at once. [3] And in proportion as the actions of a body depend more on itself alone, and as other bodies concur with it less in acting, so its mind is more capable of understanding distinctly (E2P13S, Spinoza 1985b, 458).

This passage tells us the following three moments about the problem we face. (1) It is evidence (“we also cannot deny”) that the formal reality of ideas (this means also the consciousnesses [minds] since they are composite ideas) has degrees of perfection. In the then contemporary scholastic terminology: one idea has more reality than another. According to the Preface of the fourth part of the *Ethics*, degrees of perfection (in contrast to the scholastic concept) are not differences between essences but differences between levels of *potentia*. “But the main thing to note is that when I say that someone passes from a lesser to a greater perfection, or the opposite, I do not understand that he is changed from one essence, *or* form, to another. For example, a horse is destroyed as much if it is changed into a man as if it is changed into an insect. Rather, we conceive that his power of acting, insofar as it is understood through his nature, is increased or diminished” (Spinoza 1985b, 545). Perfection for Spinoza is difference in the power to act. For example two plans of a machine (two ideas) would differ from each other by their success: the one that represents a working machine has more reality than the one that represent a flawed one. But recognizing an error in the plan does not require building the machine first. Thus the formal reality of an idea can be judged in itself, in this case without building the machine and testing it. Moreover one can correct the error in the wrong plan and thus it can have more reality than before. Thus, perfection is, so to speak, not a metaphysical but a practical concept.

This degree of perfection in the ideas corresponds to a varying degree of reality of their *ideatum*. That is to say the perfection of the formal reality of an idea (the idea in itself) corresponds to its objective reality (the idea in relation to its object). This implies an affirmation of evidence we are familiar with from Descartes’ a posteriori argument for God’s existence. Naturally it neither implies the causal relation between the *ideatum* and the *idea* also present in Descartes’ argument (it is just as well conceivable according to Spinoza’s strict parallelism), nor the theory of analogy also present in Descartes (degrees of reality are degrees of *potentia*, not degrees of essences). Since the ideatum of the human consciousness (*mens*) is the human body, the perfection of the former is dependent on the perfection of the later. But (1a) Spinoza is not about to explain the nature of the human body, only makes (2) a general remark about their respective perfection. The perfection of a body (and correspondingly a consciousness) is determined by the adaptivity of the body:

“in proportion as a Body is more capable than others of doing many things at once, or being acted on in many ways at once, so its Mind is more capable than others of perceiving many things at once”. Simultaneous perceptions are only conceivable through simultaneous affectability. Simultaneous affections on the other hand suppose a composite body; the more cooperating different components a body has, the more simultaneous affections are possible. Thus point (2) asserts the principle of adaptivity we analyzed above. The more differentiated a body becomes the more perfect it is. That is to say the more power it has.

Lastly, (3) states the principle of affirmation. The principle of affirmation is the touchstone of perfection or degree of reality. The more a body can affirm itself the more autonomous it is and correspondingly its consciousness is more able to grasp its objects in a clear and distinct manner. While the principle of adaptivity extends to the perfection of the body, (and accordingly the field of knowledge), the principle of affirmation intensifies it. *Esprit de géométrie* and *esprit de finesse à la Spinoza*.

## 4. The Integrated Affective Body and the Diversity of the Commonwealth

All of this points to the conclusion that a human body – just like its integration into a more complex body, the state or the commonwealth – is not a clock. Not because it would not be a machine – most certainly it is – but because a clockwork is a different kind of machine than the human body. A clock does not have to adapt to anything, it is self-contained and self-sufficient. The two principles do not always play equally important roles in the process of integration. A clockwork has no use of the principle of differentiation or adaptivity. The more simple its composition the more effective it is. This is not the case of human bodies neither of the state. Social beings have to adapt to external demands, as their effectivity is largely dependent on their adaptivity.

Integration of the bodies (an affective integration) can be achieved in a number of different ways. The commonwealth can be based just as much on fear as on hope (PT 5.1, 3.9). It is another question though how this integration be achieved in the best possible way. To the question of priorities Spinoza’s answer is clear: the aim of the state is peace (PT 5.2), and the best possible state has to start with providing peace to its citizens. Clearly, it is the principle of affirmation or unification that provides internal peace for a state. A state must maintain its consistency by affirming its being. Dividedness is the ultimate danger for a state as it was very well exemplified for Spinoza in the events of 1672 – the fate of the de Witt brothers in the hands of a discontent *multitudo* was a hard lesson to learn. The first goal of a state thus should be to avoid dividedness by the way of a well established unity. “For certainly we should impute rebellions [*seditiones*], wars, and contempt for, or violation of, the laws not so much to the wickedness of the subjects as to the corruption of the state” (PT 5.2, Spinoza 2016, 529).

At the same time it is equally true, that not every form of “peace” is the same, and the question of priorities must not to be confused with the question of perfection. A perfect state cannot be achieved by the sole priority of the principle of affirmation. A perfect state is a state in accord with the laws of reason, but to be in accord with reason supposes autonomy.

In II, 1 1, we showed that a man is most his own master when he is led most by reason, and so (by III, 7) that a Commonwealth is most powerful, and most its own master, when it’s founded on and directed by reason. Moreover, since the best way of living, to preserve yourself as much as possible, is one guided by the prescription of reason, it follows that the best course is the one a man or a Commonwealth pursues, insofar as it’s most its own master (PT 5.1, Spinoza 2016, 528–529).

Since an affective integration can never be governed by reason, it can only be in accord with it, the “natural” order of things reversed in this case. By the law of nature, man is autonomous (“it’s most its own master”) when he or she is led by reason. But in case of a commonwealth it is autonomy and an irrational way of life that can be established. A commonwealth is an affective community: by way of a (rationally organized) state one can alter the context of affections in such a way that their cooperation would ensure a common life which is autonomous, therefore in accord with reason. Thus living in autonomy ensures living in accord with reason. But autonomy in the case of a complex body can only be achieved by extensive adaptivity. This is why living in peace does not only mean living in unity but also living in diversity.

Emancipation for an affective way of life (as opposed to the rational way of life) means empowerment by integration of the bodies. The success of the empowerment depends on a specific integration: integration through differentiation.

## References

- Balibar, Étienne. 1988. *Spinoza and Politics*. Translated by Peter Snowdon. London – New York: Verso.
- Matheron, Alexandre. 1988. *Individu et communauté chez Spinoza*. (Nouvelle édition). Paris: Minuit.
- Molina, Luis de. 1876. *Concordia liberi arbitrii, cum gratia donis, divina praescientia, providentia praedestinatione et reprobatione*. Parisii: P. Lethielleux.
- Negri, Antonio. 1991. *The Savage Anomaly. The Power of Spinoza’s Metaphysics and Politics*. Translated by Michael Hardt. Minneapolis, Oxford: University of Minnesota Press.
- Negri, Antonio. 2004. *Subversive Spinoza*. Edited by Timothy S. Murphy. Translated by Timothy S. Murphy, Michael Hardt, Ted Stolze and Charles T. Wolfe. Manchester – New York: Manchester University Press.
- Spinoza, Baruch de. 1925. *Opera*. I–IV. Edited by Carl Gebhardt, Heidelberg: Carl Winters.

- Spinoza, Benedictus de. 1985a. Treatise on the Emendation of the Intellect. In: Spinoza, Benedictus de: *The Collected Works of Spinoza*. Vol. I. Translated by Edwin Curley. Princeton, New Jersey: Princeton University Press.
- Spinoza, Benedictus de. 1985b. Ethics. In: Spinoza, Benedictus de: *The Collected Works of Spinoza*. Vol. I. Translated by Edwin Curley. Princeton, New Jersey: Princeton University Press. = E
- Spinoza, Benedictus de. 2016. Political Treatise. In: Spinoza, Benedictus de: *The Collected Works of Spinoza*. Vol. II. Translated by Edwin Curley. Princeton, New Jersey: Princeton University Press. = PT

# XIV

*Dávid Bartha*

## The Human Passions and the Purely Active Will of God: An Introduction to Berkeley's Theory of Emotion

### 1. Introduction

Despite the growing popularity of historical research into the philosophy of passions, to my knowledge not a single article or book chapter discusses Berkeley's philosophy of emotion systematically. The simplest but most important reason for not regarding Berkeley as an ideal or even suitable target of this sort of investigation is, I assume, that he is still often seen as a philosopher who only had interesting, if not outrageous, things to say about theoretical subjects such as epistemology, metaphysics or philosophy of religion, but was not interested in more practically-minded philosophy at all. I will refer to this view as the *dismissive view*.

However, as I will show in 2.1, the dismissive view is fundamentally false. In fact, Berkeley's theoretical project fits neatly into a religiously motivated moral framework which does not exclude the possibility of expressing philosophically relevant thoughts on passions and emotions.<sup>1</sup> Moreover, I will conclude in 2.2 by claiming that a Berkeleyan theory of emotion is not only possible, but because passions play such an important role in his philosophy, it cannot be simply dismissed as uninteresting or irrelevant. Rather, the properly grounded passions urging us to act can be seen as constituting the ultimate end of his philosophy.

Undeniably, Berkeley, unlike some of his contemporaries, wrote quite little directly on the topic of emotions. The planned second part of the *Principles* would probably have dealt with some related ethically motivated issues, but the manuscript was lost during Berkeley's

---

1 Following Berkeley, in this paper passions and emotions are taken not only as synonymous but also to cover a wide range of non-cognitive mental states, including natural inclinations, desires, etc. On Berkeley's account, sense perceptions, although not counting as emotions *as such*, are inseparable from the sensations of pain and pleasure, which in turn necessarily entail emotional responses.

travels in Italy. In the second part of my paper, I identify some foundational principles of Berkeley's theory of emotions and try to compensate for the lack of an extant systematic work on passions. I will analyze his few direct references to human affections, discussing first his hedonistic categorization of pleasures (in 3.1), and contextualize them in his more elaborate considerations about God's pure, unaffected and impassive nature (in 3.2 and 3.3).

In Berkeley's view God provides both a model and a contrast for us in our efforts to handle emotions. Since we are made in the image of God, we share some crucial similarities, especially in the activity of our minds. By removing as many imperfections and negative passions from our human nature as possible, and by exercising active benevolence or charity towards others, we might approximate the activity of God. The similarity, however, is inevitably limited. Since the purely active divine will is absolutely unconstrained, God is not affected by sensations or passions at all. In the human case, however, Berkeley's verdict on passions can be seen as positive not only because he thinks that there are natural passions we cannot, and inclinations we do not have to, extinguish, but also due to his oft-repeated claim that mental considerations – whether purely rational or religious – are absolutely inert, ineffectual and unable to motivate us without the occurrence of accompanying passions.

## 2. Rejecting the dismissive view

### 2.1. The possibility of a Berkeleyan theory of emotion

The dismissive view is nicely illustrated by Hicks' straightforward statement that Berkeley's moral philosophy is "not important" (Hicks 1932, 181, cited by Clark 1988, 233). Even if only implicitly, interpreters of Berkeley widely share this view as an implication of the standard interpretive attitude toward his later works. That his late dialogue, the *Alciphron* – a work which deals with various ethical issues – is traditionally considered to be merely an apologetic writing is a great example of the common tendency of ignoring the philosophical contribution of the works he wrote after the *Dialogues* and the *Principles*. Warnock, for example, mentions the early *Passive Obedience* as the *only* work in which Berkeley directly discussed moral philosophy. The exclusive focus on the meta-ethical views of this work, which Warnock describes as a "curious little pamphlet" (1990, 48), is an oversight which motivates his interpretation that Berkeley was not interested in the details of a substantive moral theory.

Warnock (1990) subscribes to a weaker version of the dismissive view. By claiming that once the dangers of the modern scientific worldview are avoided, along with its concomitant epistemological and metaphysical absurdities, Berkeley considered moral issues easily solvable, being deducible from obvious principles, and hence hardly worth serious

discussion.<sup>2</sup> A stronger version, in contrast, would have Berkeley deny the very possibility of substantive ethics.<sup>3</sup>

Contrary to what the dismissive view suggests, Berkeley expresses on various occasions and in various ways the view that the “true business of a philosopher” is essentially ethical.<sup>4</sup> Before the publication of the *Principles*, in a letter to Percival he praises Socrates as the role-model of philosophers, claiming:

Socrates spent his time in reasoning on the most noble and important subjects, the nature of the gods, the dignity and duration of the soul, and the duties of a rational creature. He was always exposing the vanity of Sophists, painting vice and virtue in their proper colours, deliberating on the public good, enflaming the most noble and ungenerous tempers with the love of great actions. In short his whole employment was the *turning men aside from vice, impertinence, and trifling speculations to the study of solid wisdom, temperance, justice, and piety, which is the true business of a philosopher* (Letter to Percival, 27<sup>th</sup> December 1709, *Works* v8.28, emphasis added).

The very first sentence of the preface to the *Three Dialogues* also expresses his demand that “the end of speculation be practice, or the improvement and regulation of our lives and actions” (*Three Dialogues*, *Works* v2.167).

It is hardly surprising that for the later bishop, ethics has a theological foundation; God being said to be one of the great principles of morality (see *Notebooks* 508). One way in which the ethical aim of his philosophy is connected to theological considerations is that ethics is grounded on proving the existence and providential nature of God. In the preface, he claims that in addition to laying down the principles of idealism,

[...] a plain demonstration of the immediate Providence of an all-seeing God, and the natural immortality of the soul, should seem the readiest preparation, as well

---

2 Warnock assumes that Berkeley took moral philosophy to be a deductive science, as did Locke for instance. Though having entertained the idea, Berkeley clearly rejected it. See *Notebooks* 660, 676, 690; *Principles* 10 and Olscamp 1970, 60–67.

3 This view might be inspired by misinterpreting passages in which Berkeley criticizes the *abstract* concepts of goodness or virtue, like this in the *Principles*: “What it is for a man to be happy, or an object good, every one may think he knows. But to frame an abstract idea of *happiness*, prescinded from all particular pleasure, or of *goodness*, from every thing that is good, this is what few can pretend to. So likewise, *a man may be just and virtuous, without having precise ideas of justice and virtue. The opinion that those and the like words stand for general notions abstracted from all particular persons and actions, seems to have rendered morality difficult, and the study thereof of less use to mankind.* And in effect, the doctrine of *abstraction* has not a little contributed towards spoiling the most useful parts of knowledge” (*Principles* 100, emphasis added). Even in this passage Berkeley rejects only the abstract – i.e. *a priori* – way of doing moral philosophy (see also *Principles* 143), not the non-cognitivist or empiricist view he endorses.

4 Of course, I am not the first to point this out. For great introductions to Berkeley’s moral philosophy, see Darwall 2005 or Clark 1988.

as the strongest motive, to the study and practice of virtue (*Three Dialogues*, Works v2.167).

It is important to note even at the outset that, for Berkeley, religion is not a set of inert principles, but serves as an effective ethical guide of our lives. His theoretical work, as he sees it, prepares the way for embracing the “salutary truth of the Gospel”, which constitutes the “highest perfection of human nature”, not only in terms of knowledge but also in terms of practice. As he puts it in the very last sentence of the *Principles*:

For after all, what deserves the first place in our studies, is the consideration of GOD, and our duty; which to promote, as it was the main drift and design of my labours, so shall I esteem them altogether useless and ineffectual, if by what I have said I cannot inspire my readers with a pious sense of the presence of God: and having shewn the falseness or vanity of those barren speculations, which make the chief employment of learned men, *the better dispose them to reverence and embrace the salutary truths of the GOSPEL, which to know and to practise is the highest perfection of human nature* (*Principles* 156, emphasis added).

Indeed, morality is so dependent on its theological basis that without natural religion it is even said to be impossible.<sup>5</sup> Berkeley argues in the *Alciphron* that morality rests on conscience, which in turn “always supposeth the being of a God” (1.12).

Of course, it should not mean at all that ethics *as such* has no important role in Berkeley’s system, rather it means that a *religiously* established ethics is the end point of his philosophy. Rebutting materialism – including what we call dualism today – is needed in order to combat atheism. Combating atheism is in turn necessary to achieve the ethical aim. This gradual approach might explain why he wrote so little on ethics and passions. One might speculate that, given the many criticisms his theoretical position (i.e. the denial of matter) had to face, he did not feel that he could move on to the next step of developing his ethical theory in detail. In a way, the *Alciphron* can be seen as resetting the project, trying to reach the ethical-religious aim without (explicit reference to) immaterialism.

Another version of the dismissive view denies not the ethical aspect of his philosophy in general, but directly the view that Berkeley could have had a theory of passion. It might come from the consideration that Berkeley was an immaterialist, and, allegedly, those who deny the existence of the material world, have no room for discussing passions, since they are clearly related to or dependent on the body and its different physical states. Descartes, for instance, regards the passions as passive phenomena explained through bodily mechanisms.

But this concern is based on an elementary misreading of Berkeley (and of idealism or immaterialism in general), namely on the claim that, for an immaterialist, there is nothing in the world which might be labelled as body (as opposed to the soul or the mind, strictly

---

<sup>5</sup> For more on this, see Airaksinen 2016, 217.

speaking). But Berkeley has no problem speaking about the body as a publicly observable bundle of perceptions that corresponds to the volitions and sensations of a particular mind according to rules established and supervised by God – which is arguably no more mysterious than how it happens on the interactionist account. My body as a bundle of ideas counts as physical in the sense that, unlike my ideas of imagination, it can be perceived – at least partially – by more of us, and is experienced passively, in accordance to laws independent of my will. This implies, among other things, that it makes sense to say in immaterialist terms that my soul is affected by my body’s various states. Indeed, Berkeley also links the existence of passions to the concept of body, when claiming, for instance, that God has no passions precisely because he has no body limiting his volitional power (see *Three Dialogues*, Works v2.241).

There might be another – probably more serious – reason for adopting the dismissive view with regard to Berkeley’s theory of emotion. It is based on Berkeley’s famous doctrine that we do not have and cannot have ideas of the soul and its various operations (see *Notebooks* 230, 576, 828–829; *Principles* 27, 135–136). It seems to suggest that we have no chance of making publicly shareable and meaningful claims about our passions; rather, we are only experiencing them without any cognitive content to communicate in a non-subjective manner. However, he never denied that we can have notions, i.e. some sort of *non-ideational knowledge* of the mind and its operations based on the immediate experience of our agency (see *Principles* 89, 142; *Three Dialogues*; Works v2.233; *Alciphron* 7.5). Whatever constitutes exactly this sort of knowledge, Berkeley makes it clear that it enables us to discuss meaningfully the mental states, including passions, even without perceiving them as mental objects in our minds. Moreover, Flage (2006) quite plausibly argues that Berkeley does not deny that we might have reflective *ideas* of our passions, he only rejects Locke’s “inner sense” model of self-reflection with regard to the essentially active nature or substance of the mind itself. It is precisely through this perceptual awareness that we can identify and differentiate the particular passions we have or states of mind we are in.

## 2.2. The significant place of emotions in Berkeley’s philosophy

What is shown so far, I hope, proves the possibility that Berkeley might have had a philosophy of emotion, but it is a further question whether he had one deserving our attention three hundred years later. Before discussing some of his particular views on passion, I want to investigate the evidence supporting that emotions played an important role in his philosophy.

Firstly, we have seen that Berkeley, from the very beginning, had a moral aim in mind while developing his theoretical philosophy – which of course is not a curiosity in his time at all. It is natural to think that he was also interested in the way we should deal with our passions as a crucial part of this ethical picture he seeks to establish in the end. It might be

relevant that Berkeley often alludes to passions in those passages in which he declares the moral aim of his philosophy, for instance in the letter to Percival quoted in 2.1.

Secondly, his notorious commitment to common sense also supports his interest in human instincts, presumably including natural appetites and passions. As Berkeley quite ironically remarks at the very beginning of the introduction to the *Principles*, too much speculative philosophy can be harmful as opposed to just following “the dictates of Nature”, suggesting that natural inclinations can and should be trusted:

[...] it is we see the illiterate bulk of mankind that walk the high-road of plain, common sense, and are *governed by the dictates of Nature*, for the most part easy and undisturbed. To them nothing that’s familiar appears unaccountable or difficult to comprehend. They complain not of any want of evidence in their senses, and are out of all danger of becoming *sceptics*. But no sooner do we *depart from sense and instinct* to follow the light of a superior principle, to reason, meditate, and reflect on the nature of things, but a thousand scruples spring up in our minds, concerning those things which before we seemed fully to comprehend (*Principles*, Introduction 1, emphasis added).

Thirdly, in Berkeley’s idealist view, God created the world in a peculiarly direct way, without the mediation of material substances, based on which I see no reason to think that natural appetite is not as essential a component or means of providence as any other perception through which God communicates with us. Moreover, natural desires are to be satisfied, because, according to “the wonted, indulgent methods of Providence,” “whatever appetites it may have implanted in the creatures, doth usually furnish them with such means as, if rightly made use of, will not fail to satisfy them” (*Principles*, Introduction 2). In fact, Berkeley goes so far as to say that the laws of pain and pleasure and the natural instincts or passions of animals are the most remarkable features of God’s creation – and the most telling marks of his providential nature.

[...] if we attentively consider the constant regularity, order, and concatenation of natural things, the surprising magnificence, beauty, and perfection of the larger, and the exquisite contrivance of the smaller parts of the creation, together with the exact harmony and correspondence of the whole, but *above all, the never enough admired laws of pain and pleasure, and the instincts or natural inclinations, appetites, and passions of animals*; I say if we consider all these things, and at the same time attend to the meaning and import of the attributes, one, eternal, infinitely wise, good, and perfect, we shall clearly perceive that they belong to the aforesaid spirit, *who works all in all, and by whom all things consist* (*Principles* 146, emphasis added).

Another way to support my contention that emotion has a pivotal role in God’s providence and in how we are supposed to decode his message, so to speak, about the life he wants us to conduct might come from Berkeley’s well-known criticism of the Lockean philosophy of

language. It was first put forward in the *Principles*, where Berkeley concludes by emphasizing the non-cognitive or emotive function of language (see *Principles*, Introduction 20). We find this view in the *Alciphron* as well:

Thus much, upon the whole, may be said of all signs: that they do not always suggest ideas signified to the mind: that when they suggest ideas, they are not general abstract ideas: that they have other uses besides barely standing for and exhibiting ideas, such as *raising proper emotions, producing certain dispositions or habits of mind, and directing our actions in pursuit of that happiness which is the ultimate end and design, the primary spring and motive, that sets rational agents at work* (*Alciphron* 7.14, emphasis added).

For our purposes, the importance of this claim is not only that language is essentially aimed at generating emotions – suggesting their crucial role in our lives as communicating beings – but also, given that the natural world is construed as a divine language (see, for instance, *New Theory of Vision* 147 or *Alciphron* 4.7–4.15), what is true of language is true of nature as well – with all its emotive features.<sup>6</sup> So, nature being under God’s direct supervision provides not only pieces of information to cognize, but – which is probably more important to our survival and happiness – gives us guidance in the form of sensations, always accompanied by corresponding emotions, passions and appetites.

It is less often discussed in the literature, but emotions play an important part in religious faith as well. In the seventh dialogue of the *Alciphron* Berkeley appeals to the significance of emotions with respect to the belief in Christian mysteries. Faith in such things as the Trinity or the Holy grace is said to be justified if it evokes some action, disposition or emotion in us. As I hinted at earlier, Berkeley defines faith as being “not an indolent perception, but an operative persuasion of mind, which ever worketh some suitable action, disposition, or emotion in those who have it” (*Alciphron* 7.10), or as “a vital operative principle, productive of charity and obedience” (*Alciphron* 7.9).

Accordingly, the effects of the Christian doctrines are not inert ideas or intellectual principles, but “love, hope, gratitude, and obedience, and thereby [religion] becomes a lively operative principle, influencing [one’s] life and actions” (*Alciphron* 7.8). Alluding directly to affections, the same thought is nicely expressed in the *Sermon on Religious Zeal*:

Religion must not be thought to consist in a lazy inactive contemplation of virtue and morality, of God and his attributes, of the rewards or punishments he has annexed to the good or evil actions of men. *Religion, I say, is no such speculative knowlege which rests merely in the understanding. She makes her residence in the*

---

6 As Berman (1993, 203) helpfully puts it, “[b]oth languages usefully direct our actions, attitudes, and emotions; they teach us ‘how to act in respect of Things distant from us, as well in Time as Place’ (sect. 7). Seeing a visual precipice before us is like hearing someone shouting out: ‘Don’t come any further; it’s dangerous!’”.

*heart, warms the affections and engages the will* (Sermon II: On Religious Zeal, *Works* v7.16. Emphasis added).

Moreover, just as language and faith, the working of reason and science is “of an active operative nature”.

[T]he true end of speech, reason, science, faith, assent, in all its different degrees, is not merely, or principally, or always, the imparting or acquiring of ideas, but *rather something of an active operative nature, tending to a conceived good*: which may sometimes be obtained, not only although the ideas marked are not offered to the mind, but even although there should be no possibility of offering or exhibiting any such idea to the mind (*Alciphron* 7.14. Emphasis added).

The analogy between language, faith and reason suggests that rational deliberation, just as language or a religious doctrine, is psychologically ineffective, unless supported by emotions that can motivate us to act. It seems that we cannot do without passions, since intellectual reasons are insufficient by themselves. Moreover, passions have control over *mere* intellectual considerations. In an essay titled *The Sanctions of Religion* Berkeley accuses the free-thinkers who advocate moral sense theory of being ignorant of the power of passions.

*Surely they must be destitute of passion themselves, and unacquainted with the force it hath on the minds of others*, who can imagine that the mere beauty of fortitude, temperance, and justice is sufficient to sustain the mind of man in a severe course of self-denial against all the temptations of present profit and sensuality (Essays in the Guardian V: The Sanctions of Religion, *Works* v7.199. Emphasis added).

In terms of moral motivation, the only solution is to balance the presently achievable, yet momentary, pleasures with other passions raised by considering the prospects of an after-life.

Are not men actuated by their passions, and are not hope and fear the most powerful of our passions? And are there any objects which can rouse and awaken our hopes and fears, like those prospects that warm and penetrate the heart of a Christian, but are not regarded by a Free-thinker? (Essays in the Guardian V: The Sanctions of Religion, *Works* v7.199.)

In the second part I will pursue these issues a little further, but at this point this should suffice for concluding that Berkeley’s philosophical enterprise not only opens the door to the investigation of the passions, but in fact they contribute to his philosophy in various important ways. In other words, it is not only possible, but fairly plausible, if not true, that he regarded passions as an important philosophical topic. Indeed, the theme of emotions is so deeply interwoven with some crucial tenets of his philosophy that interpreters cannot

ignore it on the grounds of being irrelevant to his *allegedly* theoretically-oriented project. Berkeley himself realized the significance of passions making it clear that – since reason is ineffective without passions and passions have power over the intellectual side of our lives – conducting a good life is impossible without finding the right balance of our emotions. Given that philosophy has practical goals for Berkeley, it is fair to say that properly balanced and controlled passions with the power to motivate our actions can be regarded as the principal aim of his philosophy.

### 3. Foundations of Berkeley's theory of emotions

In the second part of my paper I will begin by examining the hedonist classification of pleasures we find most clearly in the *Notebooks*, before trying to understand Berkeley's position on human passions by comparing them with God's purely unaffected and impassive nature.

#### 3.1. Hedonism and the classification of pleasures

As we have seen, pain and pleasure are central concepts of Berkeley's philosophy. Starting from the earliest layers of his *Notebooks*, he subscribes to the view that “[n]o agent can be conceiv'd indifferent as to pain or pleasure” (*Notebooks* 143). In Berkeley's view, perceptions are always followed by passions, because (i.) “in proportion to the Pleasure & pain Ideas are attended with desire aversion & other actions” (*Notebooks* 692), and because (ii.) all perceptions include pleasure and/or pain.

It seems there can be no perception, no Idea without Will, being there are no Ideas so indifferent but one had rather Have them than annihilation, or annihilation than them. or if there be such an equall Ballance there must be an equal mixture of pleasure & pain to Cause it. there being No Ideas perfectly void of all pain & uneasiness But wt are preferable to annihilation (*Notebooks* 833).

Objects “are pleasing or disagreeable, excite the passions of love, hatred, joy, grief, and so forth” (*Principles* 1) independently of our will, but according to God's “never enough admired laws of pain and pleasure” (*Principles* 146). As he expressed it in the *Notebooks*, “we do not properly speaking in a strict philosophical sense make objects more or less pleasant, but the laws of Nature do that” (*Notebooks* 144).

Since pleasure and pain are tied to the God-given laws of nature, it is not surprising that what God orders to be pleasurable is also profitable. That is why Berkeley “[.] allow[s] not

of the Distinction there is made twix't Profit & Pleasure" (*Notebooks* 541). In fact, Berkeley overtly advocates an egoistic and hedonistic *starting point* of ethics, acknowledging that he would "never blame a Man for acting upon Interest. he's a fool that acts on any other Principle. the not understanding these things has been of ill consequence in Morality" (*Notebooks* 542).<sup>7</sup> Expressed in a similar vein:

Sensual Pleasure qua Pleasure is Good & desirable by a Wise Man. but if it be Contemptible tis not qua pleasure but qua pain or Cause of pain. or (wch is the same thing) of loss of greater pleasure (*Notebooks* 773).

But Berkeley's hedonism – also examined by Clark (1988, 241–244) – is not at all introduced as being in contradiction with religion. Rather, it is said to be the great principle of morality (presumably in addition to God and free will, mentioned in *Notebooks* 508), leading us to the right comprehension of the Christian doctrine.

Sensual Pleasure is the Summum Bonum. This the Great Principle of Morality. This once rightly understood all the Doctrines even the severest of the [Gospels] may clearly be Demonstrated (*Notebooks* 769).

Just as he points out later in his *Essay on the Sanctions of Religion* (quoted in 2.2), in the *Notebooks* he also regards it as a great mistake if we ignore future pleasures (including those of the afterlife) in favour of present ones, or, in his own words, if "[i]n Valuing Good we reckon too much on ye present & our own" (*Notebooks* 851).

Berkeley proposes various ways of distinguishing between pleasures or desires. This last quote from the *Notebooks* implies that we can and have to distinguish between different types of pleasure: short-term and long-term ones as well as pleasures enjoyed by an individual and by more of us, i.e. between private and public pleasures.<sup>8</sup> In his *Essay on Pleasures*, Berkeley differentiates natural from fantastical or imaginary desires:

It is evident that a desire terminated in money is *fantastical*; so is the desire of outward distinctions, which bring no delight of sense, nor recommend us as useful to mankind; and the desire of things meerly because they are new or foreign. Men who are indisposed to a due exertion of their higher parts are driven to such pursuits as these from the restlessness of the mind, and the sensitive appetites being easily satisfied. It is, in some sort, *owing to the bounty of Providence that,*

---

7 This passage clearly implies, contrary to what Warnock's dismissive view holds, that these "ill consequences in morality" cannot be solved simply by rejecting some mistaken metaphysical views, such as materialism.

8 It is not so important for our purposes now, but he offers another way of differentiating pleasures: "There be two sorts of Pleasure the one is ordain'd as a spur or incitement to somewhat else & has a visible relation & subordination thereto, the other is not. Thus the pleasure of eating is of the former sort, of Musick is ye later sort. These may be used for recreation, those not but in order to their End" (*Notebooks* 852).

*disdaining a cheap and vulgar happiness, they frame to themselves imaginary goods, in which there is nothing can raise desire, but the difficulty of obtaining them.* Thus men become the contrivers of their own misery, as a punishment on themselves for *departing from the measures of nature*. Having by an habitual reflection on these truths made them familiar, the effect is, that I, among a number of persons who have debauched their natural taste, see things in a peculiar light, which I have arrived at, not by any uncommon force of genius or acquired knowledge, but only by unlearning the false notions instilled by custom and education (Essays in the Guardian IV: Pleasures, *Works* v7.194–195. Emphasis added).

A noteworthy thing about this passage is that Berkeley does not posit a supernatural sort of pleasure or an intellectual pleasure as a separate kind. As I will try to further support this contention later on, it seems to me that Berkeley finds the basis of “real” pleasures not in some purely intellectual activity, but in experiencing the direct presence and providence of God *in nature*. The highest sort of pleasure in *this* life available for all of us seems to be identified with or reduced to the “cheap and vulgar happiness” we can all find in the pleasures of the God-given experiences of nature. Of course, this natural pleasure is by no means in opposition with the considerations about our long-term interests. Indeed, they point towards the ultimate pleasures of providence to be experienced after death and to be taken into account in this life.

### 3.2. The image of God doctrine and the negative side of passions

It is pretty clear that for Berkeley theological considerations play several important roles in his theory of emotions. It is not only that God’s all-seeing nature and capability of rewarding or punishing us in the afterlife are essential for ethics or, more specifically, for conscience and moral motivation. It is not just that God communicates with us primarily through the laws of pain and pleasure producing sensations necessarily accompanied by emotional responses and creates us with trustworthy natural inclinations. God with the purely active nature of his will also sets an example of the way we should handle our emotions, and thereby God provides an ethical model to imitate.

They made us look up to heaven, and bethink our selves, that we act our several parts, in the eye of an all-seeing Spirit, who hath both power to reward, and wisdom to direct our actions: that *the perfection and end of our being, is to imitate our great creator*. And, because it is impossible for thoughts, truly great and generous, to inhabit a soul taken up, and filled with the present life, therefore they have taught us to overlook this short transient being, and extend our view to long after-scenes of eternity (Sermon X: On the Will of God, *Works* v7.136. Emphasis added).

Accordingly, Berkeley often appeals to the popular Christian doctrine that humans were created in the image of God. Moreover, he often links this thought to moral action, arguing that we can realize the image of God in our souls through acting charitably and generously.<sup>9</sup>

There is something so noble and excellent in charity, that it may be said in some sort, to exalt and transform us into a similitude with God himself, there being no one perfection or attribute of the Deity, more essential than the most diffused and active benevolence. And at the same time, that this Christian grace doth brighten up and restore the *image of God* in our souls, it doth also render us in the highest degree useful to our neighbours [...] (Sermon X: On the Will of God, *Works* 7.133. Emphasis added).

The image of God theme with its ethical relevance is expressed in many other texts (*Three Dialogues*, *Works* v2.231–232, can be read this way as well). In the *Sermon on the Mystery of Godliness* (*Works* v7.88–89), for instance the mind, “which is made in the image of God”, is said to naturally tend “to the knowledge and love of God” and “to all things excellent and praiseworthy”, leading “to order and decency, to temperance moderation and justice”. That morality and God are intimately connected is not a new idea in Berkeley’s thought, but here it is more clearly expressed than ever that there is a divine principle in us that should guide our actions and the improvement of our nature. If we do not act virtuously and “the carnal earthy part prevail[s]” in our souls, Berkeley warns us, this “image of God is blurred and defaced and the divine life extinguished”, rendering us “alive unto sin and dead unto righteousness”.

The crucial point of the similarity between God’s perfection and us lies in our capability of being active – in particular, in the “the most diffused and active benevolence” – instead of being the inert subject of earthly passions. From this perspective, passions are seen as deficiencies and imperfections of men, which cannot be attributed to God, and ideally should be avoided, if not extinguished, by us as well. Accordingly, sometimes Berkeley speaks as if passion is just the same as thoughtlessness or prejudice (see *Three Dialogues*, *Works* v2.243), or that the passions are “imperfections of the mind of man” (*Alciphron* 1.2). Moreover, some passions directly oppose our similarity to God insofar as the “*private* passions and motions of the soul do often obstruct the operation of that benevolent uniting instinct” (Essays in the Guardian XII: The Bond of Society, *Works* v7.227. Emphasis added) in which our likeness to God is said to lie most clearly.

In a rare intellectualistic mood, Berkeley says that because passions (antipathies in this case) induced by education can be mistaken for natural ones, we cannot base ethics and politics on passions at all, and “we must not be directed in respect of [virtue and morality] by any emotions in our blood and spirits, but by the dictates of sober and impartial reason” (*Passive Obedience* XXI). Even more radical expressions concerning the dangers of rationally uncontrolled passions can be found in the *Passive Obedience*:

---

<sup>9</sup> Charity is discussed more directly in another sermon, see *Works* v7.33–34.

Tenderness and benevolence of temper are often motives to the best and greatest actions; but we must not make them the sole rule of our actions: *they are passions rooted in our nature, and, like all other passions, must be restrained and kept under, otherwise they may possibly betray us into as great enormities as any other unbridled lust.* Nay, they are more dangerous than other passions, insomuch as they are more plausible, and apt to dazzle and corrupt the mind with the appearance of goodness and generosity (*Passive Obedience* XIII. Emphasis added).

Berkeley was adamant that if passions, though being part of human nature, are not regulated by religious and/or rational considerations – that is, “chained and fettered by the laws of nature and reason” (*Passive Obedience* XXXVIII) – they can terribly lead us astray. In fact, without the help of religion, reason would be easily vanquished by the stronger lower part of our nature, i.e. by our senses and passions, and consequently “man [would] become a slave to his passions, which as it is the most grievous and shameful slavery” (Essays in the Guardian IX: Happiness, *Works* v7.216, see *Alciphron* 2.13 for another passage Hume might have enjoyed).

### 3.3. The limits of the image of God model and the positive side of passions

Clearly, to act morally, we have to get rid first of the negative – unnatural, private and short-sighted – nature of passions and get as close to God’s active and benevolent perfection as possible. However, the image of God in us cannot be fully realized, as our bodily nature is obviously and inevitably imperfect and dependent. While the purely active divine will is utterly unlimited and unaffected by sensations or passions (and, according to my interpretation, not even determined by the contents of God’s understanding), humans are not to achieve an absolute freedom through ignoring their natural appetites or inclinations (nor their intellectual reason). On the contrary, interpreting passions and emotions as God’s instructions on how to conduct a good life, we are to let them affect our actions, always of course supervised by religious and rational deliberations.

Whatever we think about God’s will being determined by his intellect,<sup>10</sup> it is absolutely certain, and expressed many times by Berkeley, that God does not suffer pain or in fact have any sort of sensations or passions at all. In *Notebooks* 675 he writes that God’s

---

10 As I see it, Berkeley was a theological voluntarist who thought that God’s volitions and actions are free from any sort of intellectualistic constraint and determination. Furthermore, as we have just seen, he thinks that rational considerations are ineffective without passions. So even if God has merely intellectual contents independent of his volitions, they cannot affect his actions, since, as it will be clear soon, he cannot have the corresponding passions. On my interpretation, God’s volitions, unlike ours, are effective without passions and not dependent on ideas.

knowledge extends to all ideas – even those of pain – without his being passively affected by those sensations. In the *Dialogues* he tells more about divine knowledge:

That God knows or understands all things, and that *He knows among other things what pain is*, even every sort of painful sensation, and what it is for His creatures to suffer pain, I make no question. But that God, though He knows and sometimes causes painful sensations in us, *can Himself suffer pain, I positively deny*. We who are limited and dependent spirits, are liable to impressions of sense, the effects of an external agent, which being produced against our wills, are sometimes painful and uneasy. *But God, whom no external being can affect, who perceives nothing by sense as we do, whose will is absolute and independent, causing all things, and liable to be thwarted or resisted by nothing; it is evident, such a being as this can suffer nothing, nor be affected with any painful sensation, or indeed any sensation at all*. We are chained to a body, that is to say, our perceptions are connected with corporeal motions [...] (*Three Dialogues*, Works 240–241. Emphasis added).

The reason for God not having any sensation is not only that every sensation includes some sort of pain or pleasure, which, as we have seen, triggers emotional responses, but sensory perception itself entails dependence, limitation, passivity and hence imperfection.

*There is no sense nor sensory, nor anything like a sense or sensory, in God. Sense implies an impression from some other being, and denotes a dependence in the soul which hath it. Sense is a passion; and passions imply imperfection. God knoweth all things as pure mind or intellect; but nothing by sense, nor in nor through a sensory. Therefore to suppose a sensory of any kind – whether space or any other – in God, would be very wrong, and lead us into false conceptions of His nature* (*Siris* 289. Emphasis added).

God's purely active and independent power is not determined by any uneasiness (see *Notebooks* 610) or affected by anything, like a sensation, as there is nothing he lacks (and consequently could desire) and nothing that could be beyond the scope of his voluntary control. Any pleasure or delightful emotion, as states dependent on some external factors, count as passive imperfections just as much as the painful sensations and passions that clearly cannot be attributed to God (other than merely metaphorically, see *Alciphron* 4.21).

On the other hand, humans aim at happiness and pleasure, states which are not only up to them, so to speak, but depend on external factors; first and foremost on God and the laws of pain and pleasure he established. In *Notebooks* 640 Berkeley claims that there is a great difference between man and God, even bigger than between worms and humans. Berkeley himself emphasizes that much of the difference lies in our dependent nature and that partly due to the irregularity of passions, "our nature was debased and corrupted having lost that rectitude and perfection which it must be supposed to have had coming new

made out of the hands of its creator” (Sermon VI: On the Mystery of Godliness, *Works* v7.86).<sup>11</sup>

Unlike God, we are not to be free from passions altogether, as our wills cannot and should not be completely unaffected by external causes (or, for that matter, by internal, i.e. intellectual reasons). So, in respect of morality, God’s intellectual and emotional life does not provide humans with a model to follow, and consequently we should not judge our own pleasures and passions in the negative way Berkeley – and basically the whole Christian tradition – speaks about divine passions.<sup>12</sup> While there are negative – private, short-term, rationally unguided, etc. – passions and pleasures (discussed at the end of 3.2), of which Berkeley was relentlessly critical (just as he was of the alleged immoderate sensuality of the materialists and deists), there are also passions *humans* have to have in order to live a good creaturely life.

In his later works, such as the *Alciphron*, Berkeley did not give up his early hedonism. One understanding of hedonism is expressed by Alciphron:

[Man] considers *his appetites as natural guides directing to his proper good, his passions and senses as the natural true means of enjoying this good. Hence, he endeavours to keep his appetites in high relish, his passions and senses strong and lively*, and to provide the greatest quantity and variety of real objects suited to them, which he studieth to enjoy by all possible means, and in the highest perfection imaginable (*Alciphron* 1.9. Emphasis added).

Berkeley through his spokesmen does not entirely reject this view; he only makes some important qualifications to the way Alciphron presents it. He makes it clear that (preferring long-term, natural and public pleasures to short-term, unnatural and private ones, and contrary to the uncontrolled sensuality advocated by the free-thinkers) we need to realize that our souls’ natural inclinations are not only those “which show themselves upon his first entrance into the world; to wit, the senses, and such passions and appetites as are discovered upon the first application of their respective objects” (*Alciphron* 1.14). In Berkeley’s view, our shared natural capacities include also the notion of God as well as other inclinations which are not strictly and exclusively sensual or based on instantly realizable (short-term) and private pleasures. As Crito puts it, we must not “[consider] man only as an instrument of passion” and “[absolve] him from all ties of conscience and religion” (*Alciphron* 2.6). In this way we can achieve

---

11 Beyond the irregular passions it is also due to the fact that “the understanding of man was obscure, his will perverse” (*Works* v7.86).

12 I think it is important to emphasize that God is not necessarily the *model* of goodness in every area of human life (without denying that he is the *source* of it), because one might get this impression when reading statements like Airaksinen’s (2016, 217) that God “is the ideal model of and the measure of the good and the right, or virtue”.

an healthy constitution of soul, when the notions are right, the judgments true, the will regular, the *passions and appetites directed to their proper objects, and confined within due bounds* [...] And the man whose mind is so constituted, is he not properly called virtuous? (*Alciphron* 2.12. Emphasis added.)

With regard to those passions which are “confined within due bounds”, Berkeley was more positive than we might have expected on the traditional interpretation of him, or than many philosophers of that period were. It is clear for instance that he did not contrast passions *as such* with reason and/or morality – he considered only the unbalanced, unnatural and short-sighted emotions as being in contradiction with morality, religion and reason.<sup>13</sup>

One reason – which I cannot dwell on now (see footnote 2) – is that according to Berkeley there is no light of reason or science of ethics that could and should correct or extinguish emotions considered to be necessarily irrational, unreliable and detrimental, etc., rather he took passions to be a natural and significant part of the empirical observation of what God wills us to do with our lives. As I mentioned earlier, earthly pleasures merely need to be balanced by spiritual ones generated by considerations about religious matters, such as the prospects of the afterlife.

*That men have certain instinctive sensations or passions from nature, which make them amiable and useful to each other, I am clearly convinced. Such are a fellow feeling with the distressed, a tenderness for our offspring, an affection towards our friends, our neighbours, and our country, an indignation against things base, cruel, or unjust. These passions are implanted in the human soul, with several other fears and appetites, aversions and desires, some of which are strongest and uppermost in one mind, others in another. [...] duty and virtue are in a fairer way of being practised, if men are led by reason and judgment, balancing low and sensual pleasures with those of a higher kind, comparing present losses with future gains, and the uneasiness and disgust of every vice with the delightful practice of the opposite virtue, and the pleasing reflexions and hopes which attend it? Or can there be a stronger motive to virtue than the shewing that, considered in all lights, it is every man's true interest?* (*Alciphron* 3.5. Emphasis added.)

Of course, Berkeley never asserted that passions should not be in accord with rational deliberation. Passions nonetheless seem not to be in irreconcilable contradiction with reasoning, which indeed leads us to passions which can be stronger than the unreflective affections. One crucial aim of reasoning is to free our passions from biases.

*Light and heat are both found in a religious mind duly disposed. Light in due order goes first. It is dangerous to begin with heat, that is with the affections. To ballance earthly affections by spiritual affections is right. But our affections shou'd grow from*

---

13 On my interpretation, even in those passages I quoted at the end of 2.2 Berkeley contrasts reason only with the negatively understood – private, unnatural, short-sighted – passions.

*inquiry and deliberation else there is danger of our being superstitious or Enthusiasts.* An affection conceived towards a particular Church, upon reading some spiritual authors of that Communion which might have left a bias in the mind is I apprehend to be suspected. Most men act with a byas. God knows how far my education may have byassed me against the Church of Rome, or how far a love of retreat and a fine climate may byas me towards it. It is our duty to strive to divest our selves of all byas whatsoever (Letter to Sir John James, *Works* v7.147. Emphasis added).

Berkeley makes it clear that light (i.e. reason) and heat (i.e. passion) should work together in the balancing process. This passage of course emphasizes the light, which “in due order goes first”, but does not deny the relevance of heat; light is first in terms of the temporal order, not in terms of explanation. Berkeley thought that without passion and affection there is no psychologically effective motivation for us to act at all – mere light, i.e. intellectual deliberation, is inert. The continuation of the already quoted passage from the *Sermon on Religious Zeal* nicely illustrates the collaborative relationship between reason and passion:

[...] Small are the advantages we derive from the dawning of the Sun of righteousness tho we shoud discover by it's light the beauty of Holiness, and the deformity and wretchedness of sin. if withall, the *heat thereof be not sufficient to stir our passions, to work in us strong aversion from the one and ardent desires and thirst after the other, if it does not kindle in our hearts the flames of Divine love*, if it serves not to quicken our endeavour after christian perfection and inspire us with a jealousy for the honour of God and the prosperity of his Church. In a word if we are not affected with a religious zeal. *But as it is highly needful that all the motions, and passions of the soul shoud be under the regulation and influence of Reason, whose office it is to see they are placed on proper objects, that they spring from worthy motives and are contained within a just degree.* So is this in a peculiar manner necessary with regard to religious zeal the impulses whereof are so strong and powerful. *Let yr zeal be according to knowlege* (Sermon II: On Religious Zeal, *Works* 7.16. Emphasis added).

However important it is that “all the motions, and passions of the soul should be under the regulation and influence of Reason”, if the “heat thereof [is] not sufficient to stir our passion”, nothing will have been achieved at all. But luckily (or probably due to God’s wise providence) we find beauty not only in “perusing the volume of Nature” (*Principles* 109), but religion – as realized in virtuous actions – itself provides a lot of joy and can produce great emotions in us. The rightly conducted actions and love towards God and our fellow creatures result in unsurpassable “peace of mind”, “inward comfort” and “conscientious joy”.

[...] For my part, I believe, if matters were fairly stated, *that rational satisfaction, that peace of mind, that inward comfort, and conscientious joy, which a good Christian finds in good actions, would not be found to fall short of all the ecstasy, rapture,*

*and enthusiasm supposed to be the effect of that high and undescribed principle. In earnest, can any ecstasy be higher, any rapture more affecting, than that which springs from the love of God and man, from a conscience void of offence, and an inward discharge of duty, with the secret delight, trust, and hope that attend it? (Alciphron 3.7. Emphasis added.)*

To stress it again, for Berkeley the love of God is not a merely intellectual sort of love or such that amounts to an abstract “metaphysical knowledge” – that is, not anything like the intellectual love of Descartes’s or Spinoza’s abstract God (see *Notebooks* 845) – but consists in the love of our “Creator, Redeemer, and Sanctifier” and universal charity towards all men as well as nature. The “love of God and man, the practising every virtue, the living reasonably while we are here upon earth, proportioning our esteem to the value of things, and so using this world as not to abuse it [...] is what Christianity requires” (*Alciphron* 5.5), and what provides the greatest pleasure one might hope for, at least in this life. We find this view advocated in the *Sermon on the Mystery of Godliness* (*Works* 7.88–89, referred to above), but, I think, most expressively formulated in the following sermon:

For St. *John* tells us, that *whosoever sinneth, hath not seen Christ, nor known Him* (*1 John* iii. 6). And again, *He that loveth not knoweth not God* (*1 John* iv. 8). *To know God as we ought, we must love him; and love him so as withal to love our Brethren, his Creatures and his Children.* I say, that Knowledge of God and Christ, which is *Life eternal implies universal Charity*, with all the Duties ingrafted thereon, or ensuing from thence, that is to say, *the Love of God and Man.* And our Lord expressly saith, *He that hath My commandments, and keepeth them, he it is that loveth Me* (*John* xiv. 21). From all which it is evident, that *this saving Knowledge of God is inseparable from the Knowledge and Practice of his Will*; the explicit Declaration whereof, and of the Means to perform it, are contained in the Gospel, that Divine Instrument of Grace and Mercy to the Sons of Men. *The Metaphysical Knowledge of God, considered in his absolute Nature or Essence, is one Thing, and to know him as he stands related to us as Creator, Redeemer, and Sanctifier is another.* The former Kind of Knowledge (whatever it amounts to) hath been, and may be, in Gentiles as well as Christians, but not the *latter, which is Life eternal* (*Sermon IX: Anniversary Sermon Before the Society for the Propagation of the Gospel, Works* v7.115. Emphasis added).

## 4. Conclusion

Although Berkeley was determined to criticize the crudely sensual pleasures distancing us from God as well as other people, he nonetheless acknowledged that without passions human life is inconceivable. Not just because certain natural inclinations inevitably consti-

tute our limited human nature, but that only emotions have the power necessary for motivating us to act. As I have tried to support textually, his view is that no rational deliberation can be an effective guide in our lives without raising appropriate passions in us. In light of this, Berkeley belonged to the tradition which regarded at least some passions – and not only the intellectual joy of contemplating ideas, but “genuine” emotions – as crucial to our moral and bodily well-being.

In Berkeley’s theologically grounded view this would not be possible without God’s providential activity in nature, which serves as an important ethical norm to follow insofar as we are to act benevolently to others and balance out the negative effects of the private, short-sighted, unnatural passions. Another essential component of Berkeley’s theory of passion is, however, to show that in a few important respects we have to be different from God, most crucially by realizing in our lives all those natural inclinations and appetites God has imprinted in our souls, and by acting in accordance to the pains, pleasures and the accompanying passions he raises in us.

## References

- Airaksinen, Timo. 2016. Idealistic Ethics and Berkeley’s Good God. In: Cowan, Steven B. – Spiegel, James S. (eds.): *Idealism and Christian Philosophy: Idealism and Christianity*. Vol. I. New York: Bloomsbury Publishing. 217–235.
- Berkeley, George. 1948–1957. *The Works of George Berkeley*. Edited by Luce, A. A. and Jessop, T. E. 9 vols. London: Thomas Nelson. (Abbreviated as *Works*)
- Berman, David. 1993. Cognitive theology and emotive mysteries in Berkeley’s *Alciphron*. In: Berman, David (ed.): *Alciphron in Focus*. London: Routledge. 200–213.
- Clark, Stephen. 1988. God-Appointed Berkeley and the General Good. In: Foster, John – Robinson, Howard (eds.): *Essays on Berkeley: A Tercentennial Celebration*. Oxford: Clarendon Press. 233–253.
- Darwall, Stephen. 2005. Berkeley’s Moral and Political Philosophy. In: Winkler, Kenneth (ed.): *The Cambridge Companion to Berkeley*. Cambridge: Cambridge University Press. 311–338.
- Flage, Daniel. 2006. Berkeley’s Ideas of Reflection. *The Berkeley Newsletter* 17: 7–13.
- Hicks, G.D. 1932. *Berkeley*. London: Ernest Benn.
- Olscamp, Paul J. 1970. *The Moral Philosophy of George Berkeley*. The Hague: Martinus Nijhoff.
- Warnock, G. 1990. Berkeley’s moral philosophy. *Journal of Medical Ethics* 16 (1): 48–50.

# XV

*Dan O'Brien*

## Hume, Sympathy and Belief

Hume's official account of the acquisition of testimonial belief is taken from his discussion of miracles in section 10 of the first *Enquiry*. Our 'assurance' that someone speaks the truth 'is derived from no other principle than our observation of the veracity of human testimony, and of the usual conformity of facts to the reports of witnesses' (EHU 10.5). Testimonial belief is thus grounded in inductive or causal reasoning. Hume's account is evidentialist: the testimonial beliefs we acquire are those that are supported by empirical evidence concerning their likely truth. In contemporary discussions of the epistemology of testimony Hume's position is also characterized as reductionist. The justification possessed by testimonially-acquired beliefs is reducible to the epistemic properties of the more basic faculties of perception, memory and inductive inference. Lackey (2008, 142) takes Hume to be 'the most well-known proponent of this view'.<sup>1</sup>

For Hume, though, there are various ways we can acquire beliefs: through causal inference, through the effects of contiguity and resemblance, through the eloquence and literary skill of poets and writers of fiction (T 1.3.10.7–11), through madness (T 1.3.10.9), from astonishment (at the apparent occurrence of a miracle) (T 1.3.10.4), through the kind of repetition involved in indoctrination or 'education', and through sympathy.<sup>2</sup> It is upon sympathetic mechanisms that this paper focuses and their role in belief transmission.<sup>3</sup>

---

1 Coady 1992; Fricker 2002; Schmitt 1987; McMyler 2011, 23–24 and Shieber 1999, amongst others, also interpret Hume in this way.

2 See Falkenstein (1997, 32–42) for a detailed inventory of the various mechanisms that Hume sees as underlying the acquisition of belief. Falkenstein, though, does not discuss sympathy.

3 In places, Hume claims that belief can only originate in causal reasoning: 'But as we find by experience, that belief arises only from causation, and that we can draw no inference from one object to another, except they be connected by this relation' (T 1.3.9.2; also see T 1.3.2.2; EHU 4.4). Causal reasoning is our 'single empirical "inference ticket"' (Allison 2010, 88). Loeb (2002, 76–7), however, argues that Hume is here only claiming that causal inference is required for knowledge [or 'reasonable belief' (Loeb

Given that beliefs can be acquired via sympathy, there is a further question concerning whether such beliefs are ever justified or reasonable. Most interpreters of Hume assume that they are not. This is so, for example, where there is the sympathetic spread of religious belief (to which we shall return in §2). However, in order to judge whether belief acquired in this way is always problematic we need to know more about the mechanisms involved. Hume does not make these clear, and so the primary aim of this paper is to develop a plausible account, by Hume's lights, of just how sympathy is involved in belief acquisition. In §1 I sketch the mechanisms involved in the sympathetic acquisition of emotion. §2 turns to the relation Hume sees between sympathy and belief, and §3 considers various explanations of the mechanisms involved. I argue that these mechanisms are not best seen as underlying sympathetic belief acquisition and so, in §4, I present my own suggested mechanism. §5 is a preliminary foray into the normative issues concerning sympathetic belief in which I look at different ways that such belief can be corrected.

## 1. The Mechanisms of Sympathy

'Sympathy', for Hume, does not refer to our general concern for the welfare of others – what we might call pity or compassion. 'Sympathy' should be taken in a technical sense: it refers to associative processes by which we come to share the emotions and beliefs of others. 'Sympathy',<sup>4</sup> for Hume, is closer to what we would today call 'empathy'.<sup>5</sup>

Two metaphors are used to illustrate these phenomena: those of mirrors and stringed musical instruments: '[T]he minds of men are mirrors to one another, not only because they reflect each other's emotions, but also because those rays of passions, sentiments and opinions may be often reverberated, and may decay by insensible degrees' (T 2.2.6.21). And, in judging virtue – which we shall see below involves sympathy (§4) – one 'must move some universal principle of the human frame, and touch a string to which all mankind

---

2012, 326)] whereas mere belief can be acquired via various routes. Noonan (1999, 96–99) notes that Hume makes the claim that causal reasoning is the only relation that 'informs us of the existences and objects, which we do not see or feel' (T 1.3.2.3) in order to contrast such reasoning with the relations of *identity* and *situation in space and time* that cannot by themselves inform us of such things. It is also clear that Hume does not think causal reasoning grounds *all* belief since belief in the external world, for example, does not have such an origin (T 1.4.2).

- 4 In Hume's *Enquiries* (EHU; EPM) the associative mechanisms of sympathy are not mentioned. There is debate as to whether his picture remains the same – with the 'principle of humanity' shorthand for such mechanisms – or whether his associationism is discarded. I will not consider this question here. See Debes (2007a; 2007b).
- 5 For a useful discussion of the various kinds of empathy and the relation between Hume's notion and those of interest to contemporary psychology, cognitive science and philosophy of mind, see Coplan–Goldie 2014, ix–xlvii.

have an accord and symphony' (EPM 9.6; also see T 2.3.9.12; 3.3.1.7). There is an ancient tradition in which musical resonance is seen as indicative of an occult property 'operating throughout nature and binding the universe in all its parts' (Gerbino 2015, 102).<sup>6</sup> Hume, however, has psychological accounts of how sympathy plays a role in empathizing with the emotions of others: the mimicry account and the idea-mediated account. I shall first sketch these and then explore how they might be involved in the acquisition of testimonial belief.

When observing the behaviour of others we can have a '*presensation*' (T 2.2.1.9) of what another is feeling, and Waldow plausibly claims that a *presensation* should be seen as a Humean impression. On such an account we *directly* acquire the emotions of others (Baier and Waldow 2008, 71). Inference is not involved, nor are the principles of association constitutive of the Humean imagination; emotions, rather, are propagated or caught from others. Herdt describes this as 'emotional infection' (1997, 42). Hume talks of how 'a cheerful countenance infuses a sensible complacency and serenity into my mind' (T 2.1.11.2); how '[o]thers enter into the same humour, and catch the sentiment, by a contagion or natural sympathy' (EPM 7.2); and, in a letter to Adam Smith, of how the behaviour and countenance of a depressed person can throw a 'Damp on Company' (LDH 1, 313). ('Infection' and 'contagion' are a little misleading since the emotionally infectious subject need not actually have the emotion in question. It is observable behaviour that throws a damp on company and this could be feigned or misread.)<sup>7</sup> This account is sometimes described as involving 'mirroring'. 'Mimicry', however, is a better term since 'mirroring' brings to mind Hume's claim that the minds of men are mirrors to one another, reflecting passions, sentiments and opinions. However, different kinds of mirroring are at work here, some involving mere mimicry and others idea-mediated. It is therefore better to use 'mirroring' as a more general description of what goes on in both kinds of cases.

Hume, then, also has an idea-mediated account of sympathy. My awareness of your 'countenance and conversation' (T 2.1.11.2) enables me to infer that their causal origin lies in you feeling a passion such as sorrow: via the principles of association the impression I have of your behaviour causes in me the idea of sorrow. Hume is generally taken to hold that our inferences here are based on analogy: 'Reflection on your own past experiences tells you what feelings are regularly associated with what bodily expressions, actions, or verbal behaviours. When you see similar expressions and behaviours in others, you infer that they are caused by similar passions' (Brown 2009, 233). An idea acquired in this way is subsequently 'converted into an impression, and acquires such a degree of force and vivacity, as to become the very passion itself, and produce an equal emotion, as any original affection' (T 2.1.11.3; also see T 3.3.1.7). There is, first, idea acquisition, and second, the enlivening or vivification of the acquired idea.

---

6 See Schleisser (2015) for a history of the concept of sympathy.

7 See Ardal 1966, 16.

Postema notes that Hume discusses a range of cases that ‘are best seen to fall along a spectrum from very primitive responses to responses that depend on sophisticated, intentionally engaged exercises of imaginative projection’ (2005, 258). At one end there are cases of sympathy that are ‘immediate, involuntary, and unreflective’ (ibid.); at the other, those where ‘the sympathizer is engaged cognitively and imaginatively, and the process seems to be in substantial part under the sympathizer’s voluntary control’ (ibid.): when, for example, we feel pity for someone who is about to have an operation when we see the prepared surgical instruments and the faces of onlookers (T 3.3.1.7). Here I have to imagine what kinds of procedures will be performed on the patient and whether, say, fear or relief would be the predominant emotion felt by them.

Smith talks of cases in which we might be said to step into the shoes of another:

By the imagination we place ourselves in his situation, we conceive ourselves enduring all the same torments, we enter as it were into his body, and become in some measure the same person with him, and thence form some idea of his sensations, and even feel something which, though weaker in degree, is not altogether unlike them (1976, 9).

Hume’s account does not involve such imaginative identification with another person. We do not think *ourselves* into their situation. Sympathy involves, ‘rather, a kind of forgetting of the self, as one is taken over by the sentiments of another person and in a way briefly becomes that person in the process’ (Harris 2015, 110). Only, though, *in a way*, since ‘[n]o force of imagination can convert us into another person, and make us fancy, that we, being that person, reap benefit from those valuable qualities which belong to him’ (EPM 6.3). Postema (2005, 259) sees Hume as having a *projective* account in which one can ‘transcend the immediate, momentary experiences of the object of sympathy and ... sympathize with the suffering individual as a temporally extended and socially located person’, as opposed to Smith’s *introjective* account in which one places oneself in the position of another. The projection involved is, as it were, thin: enough for us to feel what others would feel – and for this to guide our responses to their plight – but not enough for us to take ourselves to be in their situation. Through sympathy we are taken ‘out of ourselves’ (T 3.3.1.11); ‘our own person is not the object of any passion, nor is there any thing, that fixes our attention on ourselves’ (T 2.2.2.17).<sup>8</sup>

The associative mechanisms involved in Humean sympathy are parallel in certain ways with those involved in Hume’s account of the generation of belief via causal reasoning. Kemp Smith (1941) speculates that Hume’s account of belief was modelled on his account of sympathy and that the latter was the first to be developed. In both, an idea is enlivened; its vivacity is increased so as to become ‘the emotion itself’ or a belief, and in both accounts

---

8 For the contrast with Smith see Herdt 1997, 143–56. Taylor (2015, 42n16) notes that Capaldi (1976), Mercer (1972), and Hirschmann (2000) do not draw such a distinction with Smith, claiming that Hume does take sympathy to involve imagining oneself in the place of another.

this extra vivacity has its source in an impression. For beliefs the source of vivacity is a '*present impression*' (T 1.3.7.5) of sensation or reflection; for our sympathetic response to the emotions of others, the source is the impression of the self. It is the principles of association that modulate the vivacity of ideas and impressions, transferring vivacity between them; in the case of emotions, vivacity is transferred from the impression of the self to the ideas we have of the emotions of others.

Two features of Hume's account of sympathy warrant further discussion. They are the roles played, respectively, by the impression of the self and the principles of association. With respect to the former, there would appear to be tension given Hume's apparent scepticism concerning the self at T 1.4.6. Hume, however, is only skeptical with regards to our having the impression, and thus idea, of a 'simple and continu'd' self (T 1.4.6.3), a self with 'perfect identity and simplicity' (T 1.4.6.1), a Cartesian self. We can have impressions and ideas of our social, embodied selves: 'the true idea of the human mind' is the idea of 'a system of different perceptions or different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other' (T 1.4.6.19).<sup>9</sup> I see myself as a 'particular sentient and thinking person of flesh and bones, appearing and evolving in space and time, and thriving only in a certain kind of natural and social environment' (Alanen 2014, 13). It is this kind of self of which Hume is aware when he claims that '[t]he idea of ourselves is always intimately present to us' (T 2.2.4.7) and that we are 'intimately conscious of ourselves' (T 2.2.2.15). Such a self changes over time, but only gradually – as do animals, plants, ships and republics – and the imagination, via relations of resemblance and causation, leads us to ascribe identity to such a succession, even though it is only 'imperfect' (T 1.4.6.8–9). (More shall be said in the next section about such a notion of the self and its relation to pride and to sympathetically acquired beliefs.)

Vivacity is transferred from the impression of the self by the principles of association. All three principles play a role: the principles, that is, of contiguity, causation and resemblance. We sympathize more with those suffering in our own time or country, than with those in the past or in distant lands, and 'relations of blood, being a species of causation, may sometimes contribute to the same effect' (T 2.1.11.6). Resemblance, though, plays the key role. We sympathize more with those who resemble us. There is general resemblance: 'nature has preserv'd a great resemblance among human creatures, ... and this resemblance must very much contribute to make us enter into the sentiments of others, and embrace them with facility and pleasure' (T 2.1.11.5). And there are more specific resemblances with those who share our language, interests, culture, age, sex, character or circumstances. Sympathy therefore involves positive feedback. Resemblance leads, via sympathy, to the sharing of emotions and beliefs, and this has the consequence that we come more closely

---

9 For discussion of this distinction between a Cartesian and a social self, see Alanen (2014), who takes the latter as analogous to a commonwealth or republic; Greco (2015), who argues that Hume has a narrative notion of the self; Herdt 1997, 41; Ardal 1966, 43–45 and Baier 1991, 130–131.

to resemble those with whom we sympathize – sympathy, in turn, encouraged by closer resemblance.<sup>10</sup>

One must distinguish here between two modes of operation of the principles of association. First, these principles attract, as it were, ideas into the mind. The principle of association of *ideas* by the relation of resemblance leads the mind to associate ideas of things that resemble each other. My idea of a lake may lead me to think of the ocean. This kind of association is not characteristic of sympathy. The important resemblance is not between ideas, but between thinkers – between, that is, oneself and the person with whom one sympathizes. The principles of association regulate the flow of vivacity from impressions to ideas, as if controlling valves. If a relation is distant or tenuous then little vivacity is allowed to flow; if, however, there are close resemblances between oneself and another, then vivacity flows easily and the idea one acquires from another is enlivened. Debes (2007b, 317) calls this ‘self-association’, and this should be contrasted with the operation of the principle of association of ideas (by the relation of resemblance).

## 2. Sympathy and Belief

It has been noted by various authors that for Hume one can sympathize not just with the emotions of another, but with their thoughts, opinions, ‘inclinations’ and ‘turn of thinking’ (T 2.1.11.2).<sup>11</sup> As Farr (1978, 291) puts it: ‘[c]learly sympathy involves the communication of both thought(s) and feeling(s); of both cognitive and affective elements’. Hume claims:

So close and intimate is the correspondence of human souls, that no sooner any person approaches me, than he diffuses on me all *his opinions*, and draws along *my judgment* in a greater or lesser degree (T 3.3.2.2; emphasis added).

Here, he does not talk explicitly of the *sympathetic* acquisition of *belief*; but ‘diffuses’ suggests that one passively comes to adopt the opinions of another – thus, by sympathy – and elsewhere he talks of how ‘we enter so deep into the opinions ... of others’ (T 2.1.11.7) and of ‘embracing’ opinions (T 2.1.11.2) to the point where:

men of the greatest judgment and understanding ... find it very difficult to follow their own reason or inclination, in opposition to that of their friends and daily companions (T 2.1.11.2; also see T 2.3.6.8).

---

<sup>10</sup> See Baier 1991, 70.

<sup>11</sup> See Brown 2008, 232; Herdt 1997, 43; Mercer 1972, 34; Ardal 1966, 46; Baier and Waldow 2008, 82; Bricke 1996, 133; Schmidt 2003, 179, 181–182; Taylor 2015a, 32–70; Taylor 2015b, 194–200; Butler 1975, 15–16 and Postema 2005, 257.

Such deeply-held opinions would, it seem, amount to belief.<sup>12</sup>

Hume focuses on various kinds of sympathetically acquired beliefs. Here I shall briefly consider three such types.<sup>13</sup> First, there are those collective, moral and social beliefs constitutive of our national character – the ‘great uniformity we may observe in the humours and turn of thinking of those of the same nation’ (T 2.1.11.2). As put by Butler: ‘thoughts so deeply embedded in the body of accepted opinion that to question them is suspect, to disestablish them well nigh impossible’ (1975, 15).

The human mind is of a very imitative nature; nor is it possible for any set of men to converse often together, without acquiring a similitude of manners, and communicating<sup>14</sup> to each other their vices as well as virtues. The propensity to company and society is strong in all rational creatures; and the same disposition, which gives us this propensity, makes us enter deeply into each other’s sentiments, and causes like passions and inclinations to run, as it were, by contagion, through the whole club or knot of companions (Hume, E 202; ‘Of National Characters’).

Second, at various places in the *History of England* Hume suggestively talks of religion as a ‘contagion’, with narrow circles of religious enthusiasts becoming factions that endanger the stability of society.<sup>15</sup> In *The Natural History of Religion* he also talks of the ‘irresistible contagion of opinion’ (NHR 15,13). Given that beliefs comprising national character can be acquired sympathetically, it’s plausible to take contagion here to suggest the sympathetic spread of religious beliefs.

Third, beliefs concerning what others think of us are acquired via sympathy. Taylor (2015a, 30) takes Hume to have a ‘social theoretical account of human nature’ in which ‘our passionate education includes ... a process of learning appropriate ways of expressing the passions, as well as who can do so and under what circumstances’ (ibid. 31). We feel pride, for example, through sympathizing with the opinions of others concerning our

---

12 Also see T 1.3.7.5: ‘An opinion, therefore, or belief may be most accurately defin’d, A LIVELY IDEA RELATED TO OR ASSOCIATED WITH A PRESENT IMPRESSION’.

13 There are other relations between sympathy and belief that I do not have space to consider here. Herdt (1997, 191) discusses Booth’s claim that Hume’s irony ‘builds a bond of sympathy between author and reader’. When reading, for example, of the clergy’s hatred of ‘long-pointed shoes’ (H 1.241–242) the private laughs we indulge in amount to a guffaw across the centuries, with such comedic sympathy possibly affecting the beliefs we acquire concerning the respect due to the clergy and religion. Harris (2015, 348–349) also discusses how in the *History of England* ‘Hume’s reader was encouraged, coerced even, into a sympathetic emotional engagement with the victims of history. ... Perhaps Hume believed that success as a historian could not be the result of intellectual appreciation alone. Possibly he thought that the reader needed to have his emotions stirred as well’.

14 By ‘communication’ Hume does not mean verbal communication. Communication refers to sympathetic sharing of passions and emotions: ‘No quality of human nature is more remarkable, both in itself and in its consequences, than this propensity we have to sympathize with others, and to receive by communication their inclinations and sentiments’ (T 2.1.11.2).

15 See, for example, H 1.333, 4.223, 5.258 and Kemp Smith 1941, 378.

successes and place in society, and through ‘the process of being educated in the ways of a given community’ (ibid. 55). In this way I may come to feel pride in my profession, wealth and family, learn how to express this pride appropriately, and, in turn, come to see myself as the person that I am (T 2.1.5.5–6). Taylor (2015a, 39) notes that ‘the cultural competence required for recognizing the value of things is a process of learning and habituating ourselves to the social meanings, values, and norms of our community’, and this is acquired through education and sympathy: ‘sympathy is the means of communicating and sustaining the different valuations placed on the various qualities that produce pride’ (ibid. 40). There is a reciprocal relation between our own opinion of ourselves and the opinions that others have of us. Sources of pride, such as ‘virtue, beauty and riches; have little influence, when not seconded by the opinions and sentiments of others’ (T 2.1.11.1); yet, ‘[t]he praises of others never give us much pleasure, unless they concur with our own opinion, and extol us for those qualities, in which we chiefly excel’ (T 2.1.11.13).

The passions also play a crucial role in self-consciousness. The object of pride and humility ‘is self, or that succession of related ideas and impressions, of which we have an intimate memory and consciousness’ (T 2.1.2.2). The claim is not merely that we have an independent impression of the self that at times pleases us and leads us to feel pride. As Greco puts it: ‘the self and pride or humility present themselves simultaneously, in a sort of reciprocal construction’ (2015, 15): ‘nature has given to the organs of the human mind, a certain disposition fitted to produce a peculiar impression or emotion, which we call *pride*: To this emotion she has assign’d a certain idea, *viz.* that of self, which it never fails to produce’ (T 2.1.5.6). You only become self-conscious as you become aware of yourself as a possessor of virtuous and vicious character traits – those that are useful or agreeable to oneself or others: a motley bag of moral, intellectual and bodily traits such as compassion, ‘courage of mind’ (H 2.63) and broad shoulders (T 3.3.5.3). And, as discussed above, we become aware of our virtues and vices from sympathizing with what others say about us. As Harris (2015, 115) nicely puts it: ‘Hume’s conception of human nature is intensely, almost claustrophobically, social’ – our very conception of ourselves dependent on the reactions of others to us and our sympathetic absorption of their opinions.<sup>16</sup>

We have looked at three kinds of belief that are acquired sympathetically: those constitutive of national character, religious beliefs, and those that lead to feelings of pride. Sympathy, though, should be seen as having a wider influence on belief. Ardal (1966, 47) suggests ‘there is no reason to suppose that Hume does not mean his words to be taken to refer to people’s opinions as to matters of fact as well as to their moral opinions’. Hume says that children ‘implicitly embrace every opinion propos’d to them’ (T 2.1.11.2) and the suggestion here is that this is via sympathy. In the next section I shall explore various suggestions for the mechanisms involved in such belief acquisition.

---

16 See Greco 2012, 207–210.

### 3. Mechanisms of Sympathetic Belief Acquisition

The first question that arises is whether sympathetic belief involves sympathy as mimicry or sympathetic mechanisms mediated by ideas, or both. It was noted that Hume saw certain religious beliefs as contagious, and such a description may suggest the mimicry account of sympathy. I think, though, that Hume is here emphasising the sometimes virulent nature of sympathetic belief transmission, and not the direct character of the mechanism involved. Sympathetic belief acquisition is better seen as involving idea-mediated sympathy. This is more plausible given what it is like to acquire beliefs: we do not just find ourselves with beliefs, as we sometimes do with moods and emotions caught from others; we first hear an idea, entertain it and then perhaps come to find we believe it.

I will first consider idea-mediated mechanisms suggested by Vitz (2015) and Welbourne (2002). It will turn out, though, that the mechanisms suggested are not best characterized as sympathetic, or only in an oblique way. Such sympathy, for Hume, involves a two-stage process of idea acquisition and vivacity transfer from a present impression of the self, mediated by resemblance, and, as we shall see, the mechanisms suggested by Vitz and Welbourne are not of this kind.

Vitz considers various ways in which sympathy could lead to belief. His first kind of case involves a child coming to believe that she and her mother are in danger, through observing the fear on her mother's face. Vitz does not elaborate on the mechanism involved here, although it's plausible to assume a two-stage process at work: first, the mother's emotion is caught, via sympathy (either through mimicry or mediated by the idea of fear); second, fear leads the daughter to infer that they are in danger. This second stage could involve causal reasoning grounded in previous instances of where fear had been felt in dangerous circumstances. There is a sense in which such a case could be described as involving sympathetic belief, given the role of sympathy in the acquisition of emotion from which the belief is then acquired. However, the belief-acquisition itself is not sympathetic; it is, rather, grounded in causal reasoning.

Second, there are cases Vitz describes as involving 'Sympathy and Custom'. This is when repetitive experience leads to the enlivening of ideas – when they 'kindle in the common blaze' (EPM 9.9). On moving to a new town you acquire the idea that the local team is better than its rivals, and if you hear this enough you also come to believe that it's true. For Hume, repetition plays an enlivening role: 'Suppose, that a mere idea alone ... shou'd frequently make its appearance in the mind, this idea must by degrees acquire a facility and force; and both by its firm hold and easy introduction distinguish itself from any new and unusual idea' (T 1.3.9.16). This phenomenon explains how 'we may feel sickness and pain from the mere force of imagination, and make a malady real by often thinking of it' (T 2.1.11.7), and how 'liars, by the frequent repetition of their lies, come at last to remember them' (T 1.3.9.19). Taylor (2015a, 195) also aligns sympathetically acquired

beliefs with the enlivening effect of repetition, describing such cases as those involving ‘sympathetic absorption’ of repeated judgment.<sup>17</sup> Again, though, such cases should not be seen as necessarily involving sympathetic belief. Repetition may supplement the effect of sympathy, further enlivening ideas that one has acquired sympathetically, but repetition itself does not amount to sympathy because there is no role to play for the impression of the self and resemblance.

The subversive role of repetition extends to what Hume calls ‘education’, that is, indoctrination, which should be ‘disclaim’d by philosophy, as a fallacious ground of assent to any opinion’ (T 1.3.10.1). Vitz (2015, 206n15) says beliefs acquired in this way ‘probably are in many cases’ examples of contagious or sympathetic beliefs, but that ‘such beliefs could be acquired without the operations of sympathy and, hence, are not necessarily instances of contagious beliefs’. He seems, then, to be drawing a distinction between mere repetition and sympathetic belief acquisition, but it is not obvious on what this is based. In the next section I develop an account that grounds this distinction.

Third, Vitz turns to cases of ‘Sympathy, Custom and Reason’. This is where the effects of repetitive experience are supplemented by inductive reasoning; where we come to believe the reports of those who have been reliable in the past – those who have ‘authority’. Again, there is no role to play here for the impression of the self and resemblance, and Hume also explicitly distinguishes between beliefs acquired via sympathy and those grounded in authority.

[N]othing is more natural than for us to embrace the opinions of others in this particular; both from *sympathy*, which renders all their sentiments intimately present to us; and from *reasoning*, which makes us regard their judgment, as a kind of argument for what they affirm. These two principles of authority and sympathy influence almost all our opinion (T 2.1.11.9).

Welbourne (2002, 421–422) suggests a distinct account of testimonial belief acquisition which could be seen as involving sympathy (although he does not explicitly make this claim). He notes the importance of resemblance – not, though, between self and other, but between the content of testimonial reports and the ideas that they produce in the hearer’s mind. This resemblance between the impression of someone’s testimony – testimony ‘consider’d as an image’ (T 1.3.9.12) – and the idea that is subsequently acquired, enlivens the latter via the principle of association of resemblance, thus facilitating the acquisition of testimonial belief. The content of your testimony concerning Rome resembles the idea of Rome that I acquire from you, and this idea is then enlivened in virtue of this resemblance and becomes a belief concerning Rome. Such resemblance could have a supplementary effect on vivacity transfer, but, I claim, this is not the kind of resemblance relevant to sympathy. The resemblance in question is between impressions of utterances and ideas, not between the speaker and hearer involved in a testimonial exchange.

---

<sup>17</sup> Also see Ardal 1966, 47–48.

It concerns the principle of association of ideas by the relation of resemblance, rather than self-association (§1); as discussed above, though, it is the latter role for resemblance that characterizes sympathetic relations.

## 4. Testimonial Belief and Self-Association

In the sympathetic acquisition of emotions there is, first, the acquisition of the idea of the emotion felt by another, and then there is the transfer of vivacity from a present impression of the self. A parallel two-stage mechanism could result in the acquisition of belief. In this section I shall suggest such a mechanism of belief transmission.

First, from your linguistic behaviour I acquire the idea that, for example, Rome is an ancient city. Hume does not say much about how this transfer works; merely, that since ‘such a particular idea is commonly annex’d to such a particular word, nothing is requir’d but the hearing of that word to produce the correspondent idea’ (T 1.3.7.14). No deeper associative mechanism is suggested. One possible explanation is that, as with the acquisition of emotions, there is inference by analogy at work. I say ‘Rome’ when the idea of Rome is present before my mind and so I infer that you also have this idea when you say this word. Further, my idea of Rome may not be derived from an impression of sensation of actually being in Rome; rather, ‘I form an idea of *Rome*, which I neither see nor remember; but which is connected with such impressions as I remember to have receiv’d from the *conversation* and books of travelers and historians’ (T 1.3.9.4).

Belief, for Hume, is a lively idea related to a present impression, and in the context of sympathetic belief the relevant impression is that of the self. The second stage of sympathetic belief acquisition therefore involves the enlivening of ideas through self-association. And here, as with emotion, resemblance will play an important role. The more I resemble the speaker in general and specific ways, the more vivacity is transferred to the ideas acquired from them. We therefore have a lively idea, the source of vivacity being a present impression of the self – and this, for Hume, amounts to a belief.

The interrelation between the mechanisms involved in the production of the idea of the self and in sympathetic belief acquisition should again be noted. A friend admires my garden, noting that my dahlias are a most unusual colour. The resemblance between my friend and I – with respect to our interests, age, nationality, upbringing, etc. – enlivens the idea I acquire from him regarding my flowers. The source of this vivacity is the impression I have of myself, an impression that is itself produced by just such acquired beliefs concerning aspects of my life of which I should be proud, in this case, my garden. Beliefs acquired sympathetically do not always reflect upon oneself, but all sympathetic beliefs ride piggy-back on this underlying mechanism involving pride, since, without this there would not be an impression of the self to fuel sympathetic belief acquisition in general.

The mechanism I have suggested could appear *ad hoc*. Hume’s mechanistic account of vivacity transfer allows ideas to be enlivened by a variety of routes: by, for example, repetition,

inductive inference, resemblance between ideas and impressions, and between individuals. Hume does not provide comprehensive details of this account and so interpreters are free to tweak it in different ways in order to account for the phenomenology of our mental life and the flux of vivacity between ideas, beliefs, sensory impressions and emotions. Let us, though, consider the rationale behind Hume's account of vivacity transfer. The vivacity of an idea is a measure of its believability. The principle of association by the relation of causation increases the vivacity of the ideas of the usual effects of certain causes, therefore leading me to believe in their occurrence (T 1.3.6). Vivacity is transferred from one's impression of the cause – say, a fire flickering in the hearth – to the associated idea of heat that has customarily followed such impressions. Here, then, the vivacity of such an idea is keyed to our ongoing experience of the world; regularities in past experience determining the credit that should be placed in present ideas. The source of vivacity in the suggested mechanism for sympathetic belief acquisition, though, is not our experience of the world, but rather the self. The ideas I acquire from you are enlivened – are given credit – in proportion to how closely you resemble me. I trust the 'testimony' of my own senses, and I place a proportion of this trust in what I take to be the deliverances of the senses of those who resemble me in general and specific ways. I place more trust in family and friends, for example, than I do in strangers. The claim is not that this is justified, but that it is a plausible interpretation of the mechanism behind Hume's account of sympathetic belief acquisition.<sup>18</sup> There may be no rational basis for this, just as there may be no rational basis for feeling sad when those around you do so. It is not the case, though, that we sympathize with just any person who resembles us, or that we acquire any beliefs from them. The last section of the paper explores why this is so.

## 5. The Correction of Sympathy

Hume describes various ways in which sympathetic responses can be 'corrected'. These include the move from partial to corrected sympathy, the interplay of reason and sympathy in moral judgment, and the application of general rules. I shall consider how these could relate to sympathetic belief.<sup>19</sup> As said, Hume does not provide explicit details of how sympathetic belief can be corrected, but my aim here is to provide plausible mechanisms, by Hume's lights, of how this might proceed.

First, let us look at how our natural partiality is corrected in the context of moral judgment. Those aspects of an agent's character towards which we feel approval are virtues and those towards which we feel disapproval are vices. Sympathy plays a crucial role here. Our responses to the actions of others can be biased. I may feel 'delight' at Warren's arrogant

---

18 The relation between the suggested role of self-trust in Hume's account of sympathetic belief acquisition and his skeptical arguments concerning our epistemic capacities warrants further discussion.

19 It may also be instructive in this regard to consider Hume's 'Of the Standard of Taste' (E 226–249) and the normative role of sympathy with respect to aesthetic appreciation. See Baceski 2013 and Herdt 1997, ch. 6.

showboating because I am dazzled by his charm. Judgments concerning his character, though, should be independent of my own weaknesses, interests, mood or circumstances. They should not be made solely from my own point of view, but from ‘some common point of view’ (T 3.3.1.30). Hume does not demand that we adopt a view from nowhere, or an ‘angelic equi-sympathetic engagement with all of humanity’ (Sayre–McCord 1994, 203). Rather, we should ‘confine our view to that narrow circle, in which any person moves, in order to form a judgment of his moral character’ (T 3.3.3.2). I consider those who have an ‘immediate connexion or intercourse’ (T 3.3.3.2) with Warren. I may feel delight at his intellectual arrogance, but others disapprove and in sympathising with their disapproval I can come to judge Warren’s arrogance as a character flaw. Thus, via sympathy I ‘receive by communication’ (T 2.1.11.2) and ‘enter into’ and ‘embrace’ (T 2.1.11.5) the sentiments of others. To judge virtue we must follow an imaginative ‘progress of sentiments’ (T 3.2.2.25) in order to attain the common point of view from where the usefulness and agreeability of the effects of character traits can be assessed. Our initial sympathetic responses are therefore corrected.

A similar mechanism could perhaps be in play in the case of sympathetic belief, one that enables our sympathetic responses to be refined. Hume sees the shift from partial to corrected sympathy as being analogous to the correction of perceptual judgments, where, for example, we judge that a plate is circular even though it looks oval, or that the moon is a large body even though it appears small. We can correct for our partiality with respect to morality as we do for distance and perspective by adopting the common point of view. We are also partial when it comes to testimonial belief: the opinions of those close to us weigh more upon us (‘close to us’, that is, in terms of relations of causation (e.g. familial), contiguity and resemblance). A system analogous to the correction of sympathy with respect to morality could involve us coming to sympathize more readily with those who are likely to have true beliefs.

Humean reason could play a role here, and to see how this is so I will first consider the role of reason with respect to morality. For Hume, moral distinctions are not based on, nor discoverable through, reason or argument; they are, rather, grounded in our natural, sympathetic responses to the actions of our fellows. Reason, however, is not completely inert with respect to morality. Probabilistic or causal reasoning is involved in gauging the likely effects of character traits on others (EPM App. 1.2). Moral judgments can be corrected by ‘argument and reflection’: just as with ‘many orders of beauty ... it is requisite to employ much reasoning, in order to feel the proper sentiment’, so too with ‘moral beauty ... [this] demands the assistance of our intellectual faculties, in order to give it a suitable influence on the human mind’ (EPM 1.9). Reasoning is therefore involved in various ways in moral judgement, either enabling the necessary moral sentiments to be felt or reminding us what should be felt if disturbing influences were not present and we could adopt the common point of view.

Similarly, reason could play an analogous role with respect to sympathetic belief: it could, for example, provide information relevant to deciding with whom we should sympathize. Inductive inference could suggest that certain individuals are more trustworthy than

others, that certain types of people – perhaps those of a particular profession – are not to be trusted, or that one needs to take care in certain kinds of circumstances where sympathetic belief has led us astray in the past. Here, Hume's thoughts on belief in miracles and the contagious nature of certain religious beliefs are relevant. There are three distinct ways that reason could correct sympathetic belief: it could override our sympathetic responses, focus them, or enable us to acknowledge what we should believe, even though we do not.

First, sympathetic belief can be overridden by causal reasoning. Hume conceives of reason in terms of our natural associative processes regulated by general rules. At times we make rash generalizations and apply misguided rules. Having never met a witty Irishman (Hume's example) I may infer that no Irishman is witty. Such prejudices, however, can be corrected by general rules of an epistemically superior kind. These are second order judgements about the reliability of the various forms of reasoning in which we engage – what Kemp Smith calls 'wider and more reliable forms of custom' (1941, 95). Such a generalization concerning the Irish would lead to epistemic tension when one meets a witty Irishman or when one considers, for example, that Irish authors have written humorous books. One thus holds back from this kind of generalization, perhaps demanding more evidence, and thus regulating one's biases. General rules of the first kind – a species of what Hume calls 'unphilosophical probability' (T 1.3.13.1) – therefore stand in need of correction by general rules of the second kind, or more philosophical ways of thinking. In this way, beliefs acquired sympathetically could be corrected. One could, for example, correct for prejudices that have been absorbed from those around you, perhaps those constitutive of your national character.

Sceptical reasoning can also play a role here, since, '[f]or Hume, an encounter with sceptical arguments diminishes the vivacity of all of our ideas, but certain beliefs (those originating from causes that we consider to be legitimate) are better able to recover from the blow' (Falkenstein 1997, 31). As skepticism dims or extinguishes the products of the various mechanisms of belief acquisition, including that involving sympathy, the force or vivacity derived from the principle of association by the relation of causation can shine through. Further, we should not just think here in terms of the kinds of skeptical considerations raised by Hume in the *Treatise*. Children are not usually introduced to doubts concerning the existence of the external world or their own (simple and continu'd) selves, but education in the arts and sciences can involve recognition of the limits of our knowledge of the natural world and of our mental life. Education, for example, can alert us to how we are easily swayed in our beliefs by eloquent writing or how we too easily sympathize with certain kinds of people; due skepticism, then, can reduce the vivacity of ideas acquired from these sources, allowing beliefs acquired dispassionately from causal reasoning and the application of general rules to have more influence.<sup>20</sup>

Second, instead of overriding beliefs that have been acquired sympathetically, reason could help us sympathize with the right people, with, that is, those who are likely to have true beliefs. One way it could do this is by making salient certain resemblances between

---

<sup>20</sup> See Falkenstein 1997, 53–62 and Norton 1994, 131–132.

ourselves and others of which we may not have been aware, thus facilitating vivacity transfer between the impression of the self and the ideas we acquire from them; conversely, it could highlight significant differences between us, thus damping down sympathy.

Third, reasoning could lead us to acknowledge that we believe certain things that we should not. General rules are required for times when we do not feel the ‘correct’ moral sentiments – when, for example, someone’s charm blinds me to their arrogance, but I nevertheless describe them as thus; we are ‘taught this method of correcting our sentiments, or at least, of correcting our language’ (T 3.3.1.16). Similarly, with respect to belief: the application of reason and general rules could lead to linguistic expression of what we *should* believe even if we are sympathetically led astray by the beliefs of others.

I began this paper by sketching the evidentialist interpretation of Hume on testimony in which thinkers are trusted according to evidence concerning whether they have been reliable in the past. This is an individualistic picture, with the acquisition of belief wholly dependent on the epistemic capacities of the believer themselves. As Traiger (1994, 241–242) puts it: ‘Beliefs are formed and corrected by one’s private stock of perceptions. On this interpretation, the social contexts are just fluff in Hume’s account; they provide interesting but eliminable examples’. This is a picture to which I am opposed (as is Traiger, although for different reasons). For Hume, thinkers are naturally tuned in to the beliefs of others via sympathy. This is clear in his texts. He does not, though, spell out how this is to be cashed out in terms of his associationist psychology. I have provided a plausible reading of how this is to be done, and suggested ways in which such belief could be regulated, ways that could also introduce normative considerations into this widening picture of Hume’s anatomy of belief, with holistic interplay between evidentialist factors and sympathy.

## References

- Alanen, Lilli. 2014. Personal Identity, Passions, and “The True Idea of the Human Mind”. *Hume Studies* 40 (1): 3–28.
- Allison, Henry. 2010. *Custom and Reason in Hume: A Kantian Reading of the First Book of the Treatise*. Oxford: Oxford University Press.
- Ardal, Pall. 1966. *Passion and Value in Hume’s Treatise*. Edinburgh: Edinburgh University Press.
- Baceski, Tina. 2013. Hume on Art Critics, Wise Men, and the Virtues of Taste. *Hume Studies* 39 (2): 233–256.
- Baier, Annette. 1991. *A Progress of Sentiments: Reflections on Hume’s Treatise*. Cambridge, Mass.: Harvard University Press.
- Baier, Annette – Waldow, Anik. 2008. A Conversation between Annette Baier and Anik Waldow about Hume’s Account of Sympathy. *Hume Studies* 34 (1): 61–87.
- Bricke, John. 1996. *Mind and Morality: An Examination of Hume’s Moral Psychology*. Oxford: Oxford University Press.
- Brown, Charlotte. 2008. Hume on Moral Rationalism, Sentimentalism, and Sympathy. In: Radcliffe, Elizabeth (ed.): *A Companion to Hume*. Oxford: Wiley-Blackwell.

- Butler, Ronald J. 1975. Tea and Sympathy. *Proceedings of the Aristotelian Society* 49: 1–20.
- Capaldi, Nicholas. 1976. Hume's Theory of the Passions. In: Livingston, Donald – King, James T. (eds.): *Hume: A Re-Evaluation*. New York: Fordham University Press.
- Coplan, Amy – Goldie, Peter. (eds.) 2014. *Empathy: Philosophical and Psychological Perspectives*. Oxford: Oxford University Press.
- Debes, Remy. 2007a. Humanity, Sympathy and the Puzzle of Hume's Second Enquiry. *British Journal of the History of Philosophy* 15: 27–57.
- Debes, Remy. 2007b. Has Anything Changed? Hume's Theory of Association and Sympathy after the Treatise. *British Journal of the History of Philosophy* 15: 313–38.
- Falkenstein, Lorne. 1997. Naturalism, Normativity, and Scepticism in Hume's Account of Belief. *Hume Studies* 23 (1): 29–72.
- Farr, James. 1978. Hume, Hermeneutics, and History: A "Sympathetic Account". *History and Theory* 17 (3): 285–310.
- Fricker, Elizabeth. 2002. Trusting Others in the Sciences: A Priori or Empirical warrant? *Studies in History and Philosophy of Science* 33: 373–83.
- Gelfert, Axel. 2010. Hume on Testimony Revisited. *Logical Analysis and History of Philosophy* 13: 60–75.
- Gerbino, Giuseppe. 2015. Music and Sympathy. In: Schliesser, Eric (ed.): *Sympathy: A History*. Oxford: Oxford University Press.
- Greco, Lorenzo. 2012. The Force of Sympathy in the Ethics of David Hume. In: Greco, Lorenzo – Vaccari, Alessio (eds.): *Hume Readings*. Rome: Edizioni di Storia e Letteratura.
- Greco, Lorenzo. 2015. The Self as Narrative in Hume. *Journal of the History of Philosophy* 53 (4): 699–722.
- Harris, James. 2015. *Hume: An Intellectual Biography*. Cambridge: Cambridge University Press.
- Herdt, Jennifer. 1997. *Religion and Faction in Hume's Moral Philosophy*. Cambridge: Cambridge University Press.
- Hirschmann, Nancy. 2000. Sympathy, Empathy, and Obligation: A Feminist Rereading. In: Jacobson, Anne J. (ed.): *Feminist Interpretations of David Hume*. Pennsylvania: Pennsylvania State University Press.
- Hume, David. 1739–1740. *A Treatise of Human Nature*. Ed. David F. and Mary Norton. Oxford: Oxford University Press (2000) (T).
- Hume, David. 1772. *Enquiry concerning Human Understanding*. Ed. Tom Beauchamp. Oxford: Oxford University Press (2000) (EHU).
- Hume, David. 1777a. *An Inquiry concerning the Principles of Morals*. Ed. Tom Beauchamp. Oxford: Oxford University Press (1998) (EPM).
- Hume, David. 1777b. *Essays, Moral, Political and Literary*. Ed. Eugene F. Miller. Indianapolis: Liberty Fund (1987) (E).
- Hume, David. 1777c. The Natural History of Religion. In: Hume, D.: *Dialogues and Natural History of Religion*. Ed. John C.A. Gaskin. Oxford: Oxford University Press (1993) (NHR).
- Hume, David. 1778. *The History of England*. 6 vols. Indianapolis: Liberty Fund Press (1983) (H).
- Hume, David. 1932. *Letters of David Hume*. Ed. John Y.T. Greig. Oxford: Oxford University Press.
- Kemp Smith, Norman. 1941. *The Philosophy of David Hume*. London: MacMillan.

- Lackey, Jennifer. 2008. *Learning from Words: Testimony as a Source of Knowledge*. Oxford: Oxford University Press.
- Loeb, Louis. 2002. *Stability and Justification in Hume's Treatise*. Oxford: Oxford University Press.
- Loeb, Louis. 2012. Epistemological Commitment in Hume's Treatise. In: Garber, Daniel – Rutherford, Donald (eds.): *Oxford Studies in Early Modern Philosophy*. Vol. VI. Oxford: Oxford University Press. 309–347.
- McMyler, Benjamin. 2011. *Testimony, Trust, and Authority*. Oxford: Oxford University Press.
- Mercer, Philip. 1972. *Sympathy and Ethics: A Study of the Relationship between Sympathy and Morality with Special Reference to Hume's Treatise*. Oxford: Clarendon Press.
- Noonan, Harold. 1999. *Hume on Knowledge*. London: Routledge.
- Norton, David. F. 1994. How a Sceptic May Live Scepticism. In: Macintosh, John J. – Meynell, Hugo A. (eds.): *Faith, Scepticism and Personal Identity*. Calgary: University of Calgary Press.
- Postema, Gerald. 2005. "Cemented with Diseased Qualities": Sympathy and Comparison in Hume's Moral Psychology. *Hume Studies* 31 (2): 249–298.
- Sayre-McCord, Geoffrey. 1995. On Why Hume's "General Point of View" Isn't Ideal – and Shouldn't Be. *Social Philosophy and Policy* 11 (1): 202–228.
- Schleisser, Eric. (ed.) 2015. *Sympathy: A History*. Oxford: Oxford University Press.
- Schmidt, Claudia. 2003. *David Hume: Reason in History*. Pennsylvania: Penn State Press.
- Schmitt, Frederick. 1987. Justification, Sociality, and Autonomy. *Synthese* 73: 43–85.
- Shieber, Joseph. 1999. Reid on Cartesianism with Regard to Testimony: A Non-Reductionist Reappraisal. *Reid Studies* 2 (2): 59–69.
- Smith, Adam. 1976. *A Theory of the Moral Sentiments*. Ed. David D. Raphael and Alexander L. Macfie. Oxford: Oxford University Press.
- Taylor, Jacqueline. 2015a. *Reflecting Subjects: Passions, Sympathy, and Society in Hume's Philosophy*. Oxford: Oxford University Press.
- Taylor, Jacqueline. 2015b. Sympathy, Self, and Others. In: Ainslie, Donald C. – Butler, Annemarie (eds.): *The Cambridge Companion to Hume's Treatise*. Cambridge: Cambridge University Press.
- Traiger, Saul. 1994. Beyond Our senses: Recasting Book 1, Part 3 of Hume's Treatise. *Hume Studies* 20 (2): 241–59.
- Vitz, Rico. 2015. Contagion, Community, and Virtue in Hume's Epistemology. In: Matheson, Jon – Vitz, Rico (eds.): *The Ethics of Belief*. Oxford: Oxford University Press.
- Welbourne, Michael. 2002. Is Hume Really a Reductivist? *Studies in History and Philosophy of Science* 33: 407–423.<sup>21</sup>

---

21 Thanks to audiences in Antwerp, Groningen, Budapest and Oxford for their helpful feedback and encouragement, and to two referees for this volume.

# XVI

*Hans D. Muller*

## Hume on Sympathy, Pity and Impartiality

### **Introduction: descriptive account in the service of a prescriptive project**

A persistent and consistent challenge facing any sentimentalist ethics is the charge that our emotions are too prone to bias to be the proper basis of an adequately normative moral system. This objection can also be framed in terms of a failure to appropriately accommodate the principle of fairness or to meet the standard of impartiality. When considering David Hume's specific version of sentimentalism, there is a related concern that it can be difficult to draw the line between the descriptive and the prescriptive elements in Hume's writings about moral philosophy. Hume himself seems to say that his work is firmly on the descriptive side of that divide when he invokes the contrasting metaphors of the anatomist and the painter. While the former's job is to accurately depict the workings of the moral sentiments, the latter's task is to prescribe how one *ought* to feel, think, and act. Casting himself in the role of the anatomist, Hume counsels: "The anatomist ought never to emulate the painter; nor in his accurate dissections and portraitures of the smaller parts of the human body, pretend to give his figures any graceful or engaging attitude or expression". And he goes on to note that what the anatomist depicts may even be quite unpleasant: "There is even something hideous, or at least minute in the views of things, which he presents; and 'tis necessary the objects shou'd be set more at a distance, and be more cover'd up from sight, to make them engaging to the eye and imagination" (THN, 3.3.6.6: 395).

Of course the *locus classicus* of the critique of sentimentalism on the basis of its purported inability to overcome bias is the rationalist ethical system of Immanuel Kant. For Kant, sentiments are "mere inclinations" and, as such, are far too influenced by such

factors as whether we like the person with whom we are interacting and whether we happen to be in either a good or a sour mood when doing so. But it isn't just sentimentalism's rationalist opponents who lay down the charge of bias. Even philosophers who consider themselves committed Humeans do so. In a fascinating historical study, Jane McIntyre argues that Hume's original contribution to the philosophical literature on the passions is his rejection of the traditional hierarchy of the objects of our passions, according to which it is good to love God and not so good to feel sexual desire toward one's neighbor's spouse, for example, and to replace that hierarchy with "an alternative standard... derived from what is natural and usual in human nature" (2006: 212). McIntyre gives two specific examples with which she suggests that Hume's enthusiasm for the anatomist's descriptive accuracy leads him to overlook fairly obviously morally relevant details: "A certain degree of selfishness is excused because it is common (T 3.3.1.17; SBN 583); lack of affection for one's own children is blamed because it is uncommon (T 3.2.1.5; SBN 478)" (2006, 212).

But it is important to note Hume's contention that the purpose of learning all the lessons of anatomy is to deploy them in the service of the painter:

An anatomist, however, is admirably fitted to give advice to the painter, and 'tis even impracticable to excel in the latter art, without the assistance of the former. We must have an exact knowledge of the parts, their situation and connexion, before we can design with any elegance or correctness. And thus the most abstract speculations concerning human nature, however cold and unentertaining, become subservient to *practical morality*; and may render the latter science more correct in its precepts, and more pervasive in its exhortations (THN, 3.3.6.6: 395).

The passage above is from the *Treatise of Human Nature* (THN) and Hume returned to the metaphor of the cooperating anatomist and painter in the *Essay Concerning Human Understanding* (EHU), which appeared nearly a decade later and, some would argue, represents his mature philosophy:

The anatomist presents to the eye the most hideous and disagreeable objects; but his science is useful to the painter in delineating even a VENUS or an HELEN. While the latter employs all the richest colors of his art, and gives his figures the most graceful and engaging airs; he must still carry his attention to the inward structure of the human body, the position of the muscles, the fabric of the bones, and the use and figure of every part or organ. Accuracy is, in every case, advantageous to beauty, and just reasoning to delicate sentiment. In vain would we exalt the one by deprecating the other (EHU, 1.8: 90).

The goal of this essay is to assess how well one aspect of this transition from descriptive philosophical psychology to prescriptive ethics can be made to work in Hume's system.

This will be accomplished in three stages. First we will give an overview of the manner in which Hume's theory of the passions explains the systematic ways that emotions can dispose us toward bias, focusing on the emotion Hume calls "benevolence" in particular. Next we will show how Hume's account of sympathy provides a psychological mechanism for agents to take up the concerns of others as their own. In the third part of the essay we will examine the specific case of how sympathy can induce one to feel pity, which Hume regards as an other-regarding analog of benevolence towards a stranger.

But before beginning that three-stage project, it will be useful to pause to address a potential objection to my approach to the issue of impartiality. There is a traditional interpretation of Hume according to which he is an advocate of an Ideal Observer theory.<sup>1</sup> In an Ideal Observer theory, the correct moral judgment is the one that would be made by a spectator who is fully informed and completely devoid of both prejudice and self-interest. I will not be focusing on the Ideal Observer theory in this essay for two reasons. First, the structure of my argument requires that we keep separate two related goals of Hume's moral theory. Hume was interested first and foremost in *explaining* our moral practice and he also sought to *justify* that practice. In this essay, I seek to tease out one important connection between those two Humean projects. But drawing a strict equivalence between the notions of "the correct moral judgment" and "the judgment made by an Ideal Observer" conflates explanation and justification to such an extent that the connections between those two goals become obscured.

The second reason I will not be focusing on the Ideal Observer theory is that I think it is a mistake to interpret Hume as an advocate of it. Addressing that interpretative debate fully would take us beyond the scope of this paper, but the basic idea is that Hume's naturalism and his empiricism commit him advocating a moral psychology that is plausible for beings with cognitive, conative and affective capacities that human beings actually have. Humean moral agents do not have access to a perspective that includes full information. And Hume's commitment to sentimentalism is rooted in his observation that our sympathetic responses have the *actual* effect of causing us to take the well-being or misfortune of others as a matter of our own concern.<sup>2</sup>

For both of these reasons, I will not be focusing on the Ideal Observer theory but will instead explore the role of sympathy in Hume's moral psychology generally and its interactions with the passions of benevolence and pity specifically. Now it is time to begin the first of the three parts of my project.

---

1 Examples of scholars who have interpreted Hume in this way are Rawls 1971, 185–186; Firth 1952, 336–341; Harrison 1976, 114 and Glossop 1967, 527–536.

2 For a fuller discussion of the reasons not to read Hume as an advocate of the Ideal Observer theory, see Sayre–McCord 1994, 202–228.

# PART 1: The ways in which the passions can be prone to bias and the case of benevolence

As noted earlier, McIntyre has located a main source of bias in Hume's account of the passions in its descriptive or naturalistic character. She tells us that what is common is excused and what is rare is criticized. Hume certainly does give an extensive catalog of many instances where we are positively disposed towards certain people for reasons that hardly seem likely to bear up to moral or rational scrutiny. Perhaps all would agree that we are more likely to see the virtues than the faults of our friends, but Hume observes that this kind of bias can spread to our assessments of the relatives of our friends as well: "Nothing is more natural than to bear a kindness to one brother on account of our friendship with another, without any further examination of his character. A quarrel with one person gives us hatred for the whole family. Instances of this kind are every where to be met" (THN, 2.2.2.18: 221). Another example is the way that national conflicts mold our affections in certain ways: "When our own nation is at war with any other, we detest them under the character of cruel, perfidious, unjust and violent: But always esteem ourselves and our allies equitable, moderate, and merciful" (THN, 2.2.3.2: 225).

Hume argues that these reactions are evidence that people implicitly accept the quasi-utilitarian principle that "any person acquires our kindness, or is expos'd to our ill-will, in proportion to the pleasure or uneasiness we receive from him, and that the passions keep pace exactly with the sensations in all their changes and variations" (THN, 2.2.3.2: 225). I say that this is "quasi-utilitarian" because it is centered on how the actions of other people impact *me*, whereas the most prominent formulations of utilitarianism embrace the principle of non-egoism: according to which I should regard my own welfare only in light of its contribution to the aggregate of everybody impacted by the action or actions under consideration. Hume's discussion of people's tendency to think in this *egoistically* consequentialist way is followed by his examination of our "love of relations" which he presents as a kind of organizing principle of the biases of our affective reactions.

In Book Two of THN Hume provides us with a systematic examination of the passions, together with a psychological principle to help understand our emotional biases. The principle is that we feel affection for people in proportion to the extent that they stand in one of six relations to us. Those relations are consanguinity (blood relations), countrymen, neighbors, acquaintance, resemblance, and those of the same trade or profession (THN, 2.2.4.2: 228).

The particular passion that Hume sees as most susceptible to this class of biases is benevolence. In fact, one plausible characterization of Humean benevolence is that it is the default way we tend to feel towards our friends and family.<sup>3</sup> Benevolence is a direct passion. A direct passion is one for which the *cause* and the *intentional object* are one and the

---

3 Gerald Postema glosses it this way: "On Hume's view, benevolence is a spontaneous, and in that sense 'original', concern for the well-being of others, typically those to whom one is antecedently attached" (2005, 269).

same. Some other direct passions are the following three sets of contrasting pairs: desire and aversion, grief and joy, hope and fear (THN, 2.3.1.1: 257). So, for example, the snarling, unleashed dog I encounter on my morning run is both the cause of my fear and it is the thing that my fear is about (i.e. the dog is the *intentional object* of my fear). In an analogous manner, my daughter's distress at having fallen and scraped her knee leads to a sympathetic reaction in me such that her distress is both the cause and the object of my benevolence.

This talk of intentional objects may strike some Hume scholars as both anachronistic and an interpretation. And it is correct that, for Hume, any given instance of an emotion regarded on its own is simply a bare feeling with no *intrinsic* or necessarily object-directed character (THN 2.3.3.5: 266). But when considered as a part of one's overall mental economy, Hume believes it is a matter of contingent but law-like fact that an emotion will always be caused by either an idea or impression, and in turn, it will cause a subsequent idea. It is this placement within a causal network of ideas that gives the emotions an *extrinsic* intentionality. That is what Hume has in mind when he writes of the objects of emotions, as in "the *object* of love and hatred is some other person" (T 2.2.1.2: 214).<sup>4</sup>

Alongside his characterization of benevolence as a direct passion, Hume also tells us that it is a *compound* passion, in the sense that it is a mixture of two simpler passions. Hume uses the term "love" for one of the two simpler passions, while the other is a desire to increase the happiness of the person.

One of the key moral functions of benevolence is to moderate the influences of other passions that might otherwise lead an agent to be overly self-regarding: "Courage and ambition, when not regulated by benevolence, are fit only to make a tyrant and public robber" (THN, 3.3.3.3: 385). But given that benevolence is a sentiment that is typically directed toward family and friends, Hume has to answer the charge that it is a poor tool to help us advance the general good, as opposed to the more limited good of our inner circle. Hume's response to this challenge is to claim that it is "the peculiar merit of benevolence," that "even its weaknesses are virtuous and amiable" (THN, 3.3.3.6: 386). He argues that examples of benevolence are infectious and if enough people are kind to those near and dear to them, others will follow their example and the overall result will benefit the collective. Part of the evidence he gives in support of this claim is that we do not resent it when we see other people looking after their friends and family but, rather, we find it quite pleasing:

'Tis remarkable, that nothing touches a man of humanity more than any instance of delicacy in love or friendship, where a person is attentive to the smallest concerns of his friend, and is willing to sacrifice to them the most considerable interest of his own. Such delicacies have little influence on society; because they make us regard the greatest trifles: But they are the most engaging, the more minute the concern is, and proof of the highest merit of any one, who is capable of them.

---

<sup>4</sup> For further commentary on this point see Holden 2010, 55; Árdal 1966, 17–40; McIntyre 2006, 199–215, 209–210 and Alanen 2006, 178–198, 185–186.

The passions are so contagious, that they pass with the greatest facility from one person to another, and produce correspondent movements in all human breasts. Where friendship appears in very signal instances, my heart catches the same passion, and is warm'd by those warm sentiments, that display themselves before me (THN, 3.3.3.5: 385–386).

This raises the obvious question of why Hume thinks that your benevolent treatment of a member of your inner circle has the power to motivate me to treat those around me benevolently as well. We will take that issue up directly in the next section of this essay but before moving on to that I want to return to the metaphors of the anatomist and the painter and to Hume's suggestion that the work of the former has the capacity to underwrite the work of the latter.

In the passages on benevolence in THN, Hume is acting as an anatomist of the passions when he notes that we feel this emotion nearly exclusively toward those with whom we have some social or familial attachment. He is also doing anatomical work, as it were, when he notes that we find it pleasing when we encounter displays of benevolence that do not advance our self-interest. That is, this is under the remit of the anatomist because it is based on straightforward observation of how our passions operate. But it is fair to say that in this case the division of labor is not so clear-cut. In other words, our recognition that we are pleased by displays of benevolence among others points to the fact that this is an emotional interaction that has positive moral worth. And this should encourage us to also engage in such affective displays. This is the sort of synergy Hume has in mind when he wrote that, “the most abstract speculations concerning human nature, however cold and unentertaining, become subservient to *practical morality*; and may render the latter science more correct in its precepts, and more correct in his exhortations” (THN, 3.3.6.6: 395).

Now it is time to turn to the question of how Hume makes the case that displays of benevolence lead to benevolent feeling and actions in otherwise disinterested parties. And that takes us to the account of sympathy that is, in many ways, at the center of the theory of the passions in THN.

## **PART 2: SYMPATHY: from “I know you are in pain” to *I feel your pain***

When looking for a way for the agent to get beyond her own feelings and concerns, the psychological mechanism of sympathy is important. One of the ways it is important is that it provides a way to convert ideas into impressions. One of the key structural features of Hume's theory of mind is that impressions and ideas constitute the only two categories of mental states. A few preliminary remarks about them and the relation between them are in order.

Impressions are paradigmatically the direct products of sensory experience.<sup>5</sup> Ideas are copies of impressions and differ from them by being less intensely experienced. To take a simple example, when I look at a tree my perception of that object is a *tree-impression*, as it were. When I remember having seen that tree, or someone says “tree” to me, or I read a sentence about a tree, one of the thoughts I entertain is a *tree-idea*. Although it may not be immediately obvious that seeing a tree is a more intense or vivid experience than thinking of a tree, there are other examples in which this contrast is more salient. Hume’s favorite examples tend to be ones that involve the emotions:

A man, in a fit of anger, is actuated in a very different manner from one who only thinks of that emotion. If you tell me that any person is in love, I easily understand your meaning, and form a just conception of his situation; but never can mistake that conception for the real disorders and agitations of the passion (EHU, 2.2: 96).

It is significant that the examples provided by the emotions are the ones in which distinguishing between impressions and ideas in terms of the force and vivacity of the former and the relative calmness of the latter makes the most intuitive sense. For it is in the realm of emotions that sympathy plays a role of providing a pathway between those two categories of psychological states.

Sympathy is the psychological mechanism that can convert an idea into an impression. This mechanism plays an important role in interpersonal interactions because of the way that we come to be aware of the emotional states of others. Hume tells us that the fact that another person is experiencing an emotion, “is at first known to us only by its effects, and by those external signs in the countenance and conversation, which convey an idea of it” (THN, 2.1.11.3: 206). To take an example, when I observe you getting red in the face and begin to shout at the person next to you, the idea of anger becomes present to my mind. The distinctive function of sympathy is its capacity to transform such an idea into an impression. As mentioned earlier, on Hume’s theory of mind ideas and impressions differ only in degree, not in kind, with impressions simply being more “vivacious” and “lively” than ideas (THN, 1.1.1.1: 7). This brings us back to the mechanism of sympathy.

Once I have the type of perceptual experience outlined above, which brings the idea of your emotion to my mind, if that idea is “infus’d by sympathy”, then the idea of it, “is presently converted into an impression, and acquires such a degree of force and vivacity, as to become the very passion itself, and produce an equal emotion, as any original affection” (THN, 2.1.11.3: 206). An important aspect of the conversion of the idea into an impression, with all its characteristic force and vivacity, is that it becomes a *felt* experience: i.e. an *emotion* proper.

And our capacity to be affected by the emotional states of those around us can have the effect of changing our sentiment and mood in quite impactful ways: “No quality of human

---

5 I say “paradigmatically” because we will see shortly that there is at least one other route besides sensory experiences for impressions to be brought into consciousness, and that is the imagination. But the *original* source of all impressions is indeed sense-perception.

nature is more remarkable, both in itself and in its consequences, than that propensity we have to sympathize with others, and to receive by communication their inclinations and sentiments, however different from, or even contrary to our own" (THN, 2.1.11.2: 206). Since, on Hume's theory of the passions, all emotion-types are associated in a systematic way with either pain or pleasure, respectively, this "remarkable" tendency of the passions to be communicated as an idea and then converted into a felt impression will, in each case, be either a pleasant or an unpleasant experience: "A chearful countenance infuses a sensible complacency and serenity into the mind; as an angry or sorrowful one throws a sudden damp upon me" (THN, 2.1.11.2: 206).

It is crucial to note that, for Hume, "sympathy" is not a synonym for "pity" as it sometimes is in common parlance. In fact, on Hume's account, sympathy is not an emotion at all. It is a psychological mechanism that can perform the function of converting an idea of an emotion into the felt emotion itself (the impression, to use Hume's term). So sympathizing with someone is a matter of experiencing the same emotion-type as that person is experiencing as a matter of fellow-feeling. Using the adverbial expression allows perhaps the most precise way of putting the point: to sympathize, in Hume's sense, is to grieve or rejoice or be angry *sympathetically*. So on Hume's system, sympathy provides one straightforward way to be other-regarding. Through this psychological mechanism, there is a literal sense in which your joy can become my joy and your grief can become my grief. Your experiencing those emotions can become the cause, if mediated by sympathy in the right way, of my experiencing those emotions.

It is almost impossible to exaggerate the importance and the uniqueness of the impact of sympathy on our mental and emotional lives. In pretty much *every* other circumstance, the order of causation goes from-impressions-to-ideas. A typical scenario would be one in which I have some sensory contact with an object – I pick up a rock, say, feeling the roughness of its exterior and gauging its heft, and seeing its multiple hues – with those experiences entering my mind as impressions which, after a short amount of time, lose their immediacy and "vivacity" and fade into the less intense copies that Hume calls ideas. That is the order of things in nearly every situation. Every situation except those times when another person's outward behavior and expressions elicit the thought ("idea") of an emotion in me and then the mechanism of sympathy converts that emotion into a felt impression. And, remarkably, according to Hume my experience of that sympathetic emotion will very likely be more intense than a typical experience of such an emotion of my own: "Hatred, resentment, esteem, love, courage, mirth and melancholy; all these passions I feel more in communication than from my own natural temper and disposition" (THN, 2.1.11.2: 206).

Keeping that aspect of Hume's account firmly in view is the key to understanding his claim that my displays of benevolent affection toward members of my inner circle can have the capacity to motivate you to similarly benevolent behavior toward your friends, family, colleagues and neighbors. The Humean conception of the human condition is intensely social. He thinks that we are disposed to adopt the emotional tenor of those around us in a quite strong way.

With the preceding discussions about Humean benevolence and sympathy in place, we are ready to embark on the third section of this essay. In that section we will see how sympathy interacts with the imagination to produce the emotion of pity, which has the function of imitating the effects of benevolence. But in the case of pity, the agent's concern will extend beyond her inner circle to engage with the trials and tribulations of strangers. So it is with that emotion that some hope of reaching toward the norm of impartiality may be found.

## **PART 3: How sympathy allows us to feel benevolence toward strangers via pity**

In parts one and two of this essay we noted two important features of Humean benevolence: (a) its objects are almost exclusively people with whom we have some previous emotional attachment and (b) its potential positive effects on the wider community are grounded in the capacity of sympathy to spread benevolence through public displays of affection for loved ones leading to that sentiment being taken up by others. The combination of these two claims will likely leave the reader thinking that benevolence provides pretty thin fare for those of us who are interested in the prospects of Hume's sentimentalist ethics meeting the norm of impartiality or supporting the principle of fairness. For more help on that front, we now turn to pity.

As is typical in Hume's theory of the passions, he introduces pity as half of a pair with its negative counterpart, malice. And each of these is characterized as a "counterfeit" of one half of a pair of emotions that are typically directed at people with whom we are already familiar: that pair being benevolence (or love) and hate. The reason that pity and malice count as counterfeits is that they replicate the effects of love and hate, but do so in a way that is directed toward strangers:

But tho' the desire of the happiness or misery of others, according to the love or hatred we bear them, be an arbitrary and original instinct implanted in our nature, we find it may be counterfeited on many occasions, and may arise from secondary principles. *Pity* is a concern for, and *malice* a joy in the misery of others, without any friendship or enmity to occasion this concern or joy. We pity even strangers, and such as are perfectly indifferent to us... But if we examine these affections of pity and malice we shall find them to be secondary ones, arising from original affections, which are vary'd by some particular turn of thought or the imagination (THN, 2.2.7.1: 238).

So pity has the capacity to induce you to feel the same way toward a stranger as you would normally feel toward a family member or close friend: that is the sense in which pity is the

counterfeit of benevolence. The process begins with an “original affection” which is varied “by some particular turn of thought or imagination” so that you consider the stranger in the same light as you consider those to whom you are already bound by kinship or acquaintance.

To understand how this works, we need to compare and contrast the sentiments of benevolence and pity.<sup>6</sup> We have already noted that benevolence is often triggered by sympathetic discomfort with the distress of a person known to us. One of the basic principles of the way sympathy functions within the theory of the passions is that it creates what Hume calls a “reflex image” of the emotion that is observed by the agent. Another important feature of the account of the passions itself is the binary taxonomy of pleasant emotions, on the one hand, and painful or uncomfortable emotions, on the other hand. Putting the reflex image point together with the binary taxonomy point, it is clear that any occasion in which I feel benevolence toward you will involve some discomfort on my part because my sympathetic reaction to your distress will be experienced by me as involving some pain or at least discomfort. For Hume, one of the distinguishing features of benevolence is the fact that my love for the person who is in distress has the function of mitigating the painful experience of sympathizing with the other person’s anguished state. This facet of Hume’s account comes into particularly stark relief in a letter he wrote to Adam Smith in which he criticizes the account of sympathy in Smith’s *Theory of Moral Sentiments*. In that letter, Hume accuses Smith of treating all cases of sympathizing with cases of sympathizing with a friend.<sup>7</sup> In the course of setting out his objection, Hume is more explicit and exact about his view of the correct account of this aspect of sympathy than he is anywhere in THN, so I will quote from that piece of correspondence at length:

I wish you had more particularly and fully prov’d that all kinds of Sympathy are necessarily Agreeable. This is the Hinge of your System, and yet you only mention the Matter cursorily in p. 20. Now it would appear that there is a disagreeable Sympathy, as well as an agreeable: And indeed, as the Sympathetic Passion is a reflex Image of the principal, it must partake of its Qualities, and be painful where that is so. Indeed, *when we converse with a man with whom we can entirely sympathize*, that is, where there is a warm and intimate Friendship, the cordial openness of such a Commerce overpowers the Pain of a disagreeable Sympathy, and renders the whole Movement agreeable. But in ordinary Cases, this cannot have place. An ill-humor’d Fellow; a man tir’d and disgusted with every thing, always *ennuié*; sickly, complaining, embarrass’d; such a one throws an evident Damp on Company, which I suppose wou’d be accounted for by Sympathy; and yet is disagreeable (Hume 1932, 313, Hume’s letter to Smith of 28 July 1759).

---

6 In using this strategy for getting clearer on Hume’s account of pity, I am following Postema 2005, 269–273.

7 For a fuller discussion of this objection of Hume’s, as well as a survey of potential responses available to an advocate of Smith’s account, see Muller 2016, 228–230.

There is a fair amount of nuance in Hume's way of teasing out the differences between the respective experiences of sympathizing with a stranger and sympathizing with a friend. One of his ideas is that the presence of the agreeable sentiment of love (or benevolence) in the latter case can compensate for the pain of sympathizing with whatever unpleasant emotion one's friend is experiencing. Hume has in mind here a model of aggregating all the pleasure and pain in a given experience to arrive at a net total. In the contrasting case of sympathizing with the discomfort and pain of a stranger, Hume thinks the overall experience will likely be, on balance, disagreeable because in such a case there is no compensatory agreeable experience of love to offset the sympathetic pain.

In light of all this, one might wonder how pity could function in a way that motivates one to be positively disposed towards another person. Given that Hume thinks instances of sympathetic passion are going to be either painful or pleasurable *and* that any situation which might trigger pity is necessarily going to be one involving one or more unpleasant emotions, why would I not simply resent or even despise you for causing this unpleasant sensation in me? Hume is aware that his account faces this difficulty and his response comes in the form of a "maxim" which he says is, "necessary to the explication of the phænomena of pity and malice" (THN, 2.2.9.11: 247). That maxim and his explanation of it are as follows:

*That 'tis not the present sensation or momentary pain or pleasure, which determines the character of any passion, but the general bent or tendency of it from beginning to end. For this reason, pity or a sympathy with pain produces love, and that because it interests us in the fortunes of others, good and bad, and gives us a secondary sensation corresponding to the primary; in which it has the same influence of love and benevolence* (THN, 2.2.9.11: 247–248).

To understand how pity for a stranger ends up having the same influence of love and benevolence which, once again, are passions reserved for those with whom we are already familiar, it will be helpful to consider why we have the attitude toward our friends and family that we do. Hume's answer to this question is that our familiarity both with their past experiences and their future plans means that we engage with them not simply as momentary loci of current pleasures or pains, but as temporally extended entities whose longer-term interests are being impacted by what is happening in the here and now. An example will help make this clearer.

When my sister tells me that the job interview she had this afternoon did not go well because she inadvertently offended one of the interviewers by making a negative comment about the current American president, my reaction to this news is multifaceted. My first thoughts and feelings will be in response to the sad expression on her face and the dejected tone in her voice. As noted in the second section of this essay, Hume's analysis of this process is that the "external signs in the countenance and conversation... convey an idea," of the emotion and then, "this idea is presently converted into an impression, and acquires such a degree of force and vivacity, as to become the very passion itself" (THN, 2.1.11.3: 206).

In other words, I experience an episode of sympathetic fellow-feeling with her disappointment and sadness. But that is not the end of the process.

Since I am familiar with her hopes and plans, I also recall how thoroughly and studiously she prepared for the interview and how much she had hoped to get this new position because it is in a city that she and her partner would very much like to reside in. In short, I see her current disappointment as one small part of a much more extended narrative that engages a whole set of her interests and concerns. As I result, I feel a motivation first to console her in an effort to relieve her disappointment and also to offer any advice I can think of or provide any other assistance that might help her improve her situation moving forward. So that is an example of how a sympathetic response to a disagreeable emotion of a family member works in concert with the benevolence that I am predisposed to feel for her to provide a motivation for me to help her.

In the case of pity, Hume's suggestion is that my sympathetic response to a stranger's misery engages with my imagination in such a way that I come to see his current distress as embedded in a temporally extended set of preparations, concerns and plans. His explanation of how this happens appears in one of the most extraordinary passages in THN:

...however we may look forward to the future in sympathizing with any person, the extending of our sympathy depends in a great measure upon our sense of his present condition. 'Tis a great effort of imagination, to form such lively ideas even of the present sentiments of others as to feel these very sentiments; but 'tis impossible we cou'd extend this sympathy to the future, without being aided by some circumstance in the present, which strikes upon us in a lively manner. When the present misery of another has any strong influence upon me, the vivacity of the conception is not confin'd merely to its immediate object, but diffuses its influence over all the related ideas, and gives me a lively notion of all the circumstances of that person, whether past, present, or future; possible, probable or certain. By means of this lively notion I am interested in them; take part in them; and feel a sympathetic motion in my breast, conformable to whatever I imagine in his (THN, 2.2.9.14: 248).

So sympathy engenders in the pitying subject a pattern of sustained engagement with the concerns of the suffering person that imitates the way that benevolence functions. The key difference is that, in the case of pity, the engagement occurs *in the imagination*; whereas in the case of benevolence the engagement comes from such things as a shared history, lived experience, and first-hand knowledge of the other person's hopes and plans.

A question naturally arises here: why do we feel pity for strangers on some occasions, while at other times we do not? Answering this question would seem to be a key to our assessment of the capacity of this version of sentimentalism to meet the norm of impartiality. Hume identifies two variables as being particularly salient. The first factor is the level of intensity of the original sympathetic response: "When a sympathy with uneasiness is weak, it produces hatred or contempt... [but] when strong, it produces love or tenderness" (THN,

2.2.9.14: 248), and that, notably, would seem to be something over which we have little voluntary control. The second variable is whether or not the sympathizing subject engages in “a great effort of the imagination,” thereby generating an extremely vivid conception of the other person’s future suffering (THN, 2.2.9.14: 248). The amount of effort we put into our attempt to imagine our way into the other person’s shoes, as it were, *does* seem to be something over which we have a fair amount of voluntary control. So it is plausible to read Hume in these passages as playing the role of the painter in addition to his more usual role of the anatomist. To the answer, “How should I use my passions to be more impartial?”, the Humean answer would appear to be, “Put in more imaginative work.” Which is, after all, something like a secular version of the command to *love thy neighbor as thyself* or, to be more precise, *consider the concerns of the stranger as you would those of the members of your inner circle*.

## Conclusion

Hume tells us that his writings on the moral emotions are analogous to the work of an anatomist in the sense that he is trying to give an accurate description of how the emotions function and interact to produce behavior. In light of that, it is tempting to regard THN and EHU as works of descriptive psychology. But Hume also says that in order to understand how to harness the power of the emotions to get oneself, and perhaps others, to behave morally one needs first to understand how the emotions actually work. That is the point of his metaphor of the anatomist’s work serving to make the work of the painter, “more correct in its precepts, and more pervasive in its exhortations” (THN, 3.3.6.6: 395).

In this essay I have argued that Hume’s account of two psychological faculties, sympathy and the imagination, provides a thorough explanation of how a moral agent can come to feel pity, which is an impartial moral attitude. This is a significant because it can be used as one part of a response to the traditional rationalist objection that sentimentalism is vulnerable because it is based on psychological features that are too prone to bias. And that project can be seen as bearing a thematic relation to Hume’s metaphor of the sentimentalist moral theorist serving the roles of *both* the anatomist and the painter. In other words, Hume’s work in these areas counts as both psychology and moral theory. For the sake of brevity and clarity, I have not engaged directly with Hume’s utilitarian and hedonistic value theory and thus have not tried to explain how he establishes an independent moral standard. Leaving the complexities of those issues aside has allowed us to get clearer on what resources exist in Hume’s system to accomplish the following two tasks: (1) describe and explain the psychological mechanisms that are responsible for the ways in which our emotions are subject to bias, and (2) offer a prescription for harnessing the power of our emotions toward engendering behavior that moves us toward meeting the standard of impartiality and the principle of fairness. The fact that those two tasks are often seen as in fundamental tension is a testament to Hume’s recognition of the complexity

and difficulty of this subject matter and should commend continued careful reading of his work to philosophers now and in the future.

## Works Cited

- Alanen, Lilli. 2006. The Powers and Mechanisms of the Passions. In: Traiger, S. (ed.): *The Blackwell Guide to Hume's Treatise*. Malden, MA: Blackwell Publishing. 179–198.
- Árdal, Páll. 1966. *Passion and Value in Hume's Treatise*. Edinburgh: Edinburgh University Press.
- Firth, Roderick. 1952. Ethical Absolutism and the Ideal Observer. *Philosophy and Phenomenological Research* 12 (3): 317–345.
- Glossop, Ronald. 1967. The Nature of Hume's Ethics. *Philosophy and Phenomenological Research* 27 (4): 527–536.
- Harrison, Jonathan. 1976. *Hume's Moral Epistemology*. Oxford: Oxford University Press.
- Holden, Thomas. 2010. *Spectres of False Divinity: Hume's Moral Atheism*. Oxford: Oxford University Press.
- Hume, David. 1932. *The Letters of David Hume*. Edited by J. Y. T. Grieg. Oxford: Clarendon Press.
- Hume, David. 1999. *An Enquiry Concerning Human Understanding*. Edited by Tom L. Beauchamp. Oxford: Oxford University Press.
- Hume, David. 2000. *A Treatise of Human Nature*. Edited by D. F. Norton and M. J. Norton. Oxford: Oxford University Press.
- McIntyre, Jane. 2006. Hume's "New And Extraordinary" Account of the Passions. In: Traiger, S. (ed.): *The Blackwell Guide to Hume's Treatise*. Malden, MA: Blackwell Publishing. 199–215.
- Muller, Hans. 2016. Sympathy for Whom? Smith's Reply to Hume. *Journal of the American Philosophical Association* 2 (2): 212–232.
- Postema, Gerald. 2005. "Cemented with Diseased Qualities": Sympathy and Comparison in Hume's Moral Psychology. *Hume Studies* 31 (2): 249–298.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Belknap Press.
- Sayre-McCord, Geoffrey. 1994. On Why Hume's "General Point of View" Isn't Ideal – And Shouldn't Be. *Social Philosophy and Policy* 11 (1): 202–228.

## List of In-Text Abbreviations

EHU: *An Enquiry Concerning Human Understanding*

THN: *A Treatise of Human Nature*\*

\* One exception: since the citations of Jane MacIntyre's chapter are direct quotes, in those I follow her custom of using "T" to denote *A Treatise of Human Nature*.

# XVII

*Csaba Olay*

## Alienation in Rousseau

### Introduction

Jean-Jacques Rousseau formulated one of, if not the fiercest, critiques of modern civilization and society in general. He is often claimed to have accurately described or at least anticipated what later on was named alienation<sup>1</sup>, even if he himself made no use of this term. In this paper I examine what can be regarded as a precursor of the concept of alienation in Rousseau's work. In order to highlight theoretical differences and presuppositions, a brief discussion of the Marxian theory of alienation will be needed. This will be the background for my argument that what we find in Rousseau can only be labeled "alienation" in a looser sense, since the semantic core of alienation, i.e. something's becoming strange or foreign to someone, does not exactly correspond to Rousseau's basic problem. Rousseau works with an idea that might be called "alienation", but has a significantly different conceptual structure than alienation in Marx and in the tradition relying on him.

To show this, after some remarks on the use of the term "alienation", Rousseau's *Second Discourse* on inequality will serve as a point of reference in interpreting the concept. In this text Rousseau claims essentially that (1) human beings under conditions of civilization, while seemingly free, are in fact enslaved by their mutual social relations, and (2) that this slavery to one another was brought about by a socio-cultural development leading to a loss of authenticity. Alienation in a broad sense could be and has been understood in Rousseau as an analysis of "social pathologies" in the sense of this development of modern society. However, alienation in this sense has the structure of possession and subsequent disappropriation of man's original constitution. If we take a closer look at Rousseau in the light of a more specific concept of based on the theory of Marx, it can be pointed out

---

1 Zehnpfennig 2013, 179; Jaeggi 2014; Struma 2001, 161. Christoph Henning thinks that key motifs of Rousseau's complex work could be arranged around the center of "the major topic of alienation" (Henning 2015, 35–36).

that there is a general structure of alienation that might be described with the possession – disappropriation – reappropriation formula. In this paper I shall claim that Rousseau has a simplified version of alienation in the form of hypothetical possession – disappropriation. Additionally, I shall highlight that each of these theories has difficulties with defining an Archimedean point of non-alienated conditions.

## Rousseau's critique of modern society and civilization

Since there is no explicit use of “alienation” (*aliénation*) in the specific philosophical sense in Rousseau's writings<sup>2</sup>, to talk about alienation in his work requires finding a conceptual proxy for the term. The only candidate for this is the corrupted state or constitution of modern civilized humans, more precisely, modern man's distance from his original state. The basic tenet of Rousseau's theory, then, could be reformulated roughly as follows: modern civilized man has become alienated from man's original nature. With this rudimentary definition of “alienation” in mind, we have to work out the elements composing the semantic structure of the term: something's becoming strange to someone which previously belonged to it. Let us try to find with this minimal conception of alienation how Rousseau describes modern society and the state of modern, civilized man.

Rousseau's two early discourses expound perhaps the most radical critique of human culture and civilization. Even Plato's notorious critique of poetry is far more restricted in its scope. In the *Republic* Socrates rejected, in his famous analysis of poetry as *mimesis mimeseōs*, not the whole of human culture and civilization, but a part of it. His objections concern, thus, a limited area of culture, i.e., the fine arts. Rousseau radically goes beyond that in regarding the entire field of sciences, arts, and even morals as corrupting. Keeping in mind the proposal that the basis of this criticism is the alienation of persons under conditions of civilization, we obviously need to specify the elements entailed in the above rudimentary definition of alienation. Let us follow in some details this well-known line of thought.

Rousseau's basic arguments against arts and sciences are elaborated in the two *Discourses* published in 1750 and 1755. Whereas the *First Discourse* works out principal objections against sciences, fine arts, and morals in general, the *Second Discourse* tries to show how the development of human society creates heavily distorting conditions. The *First Discourse* seeks not only to show that sciences and fine arts are luxurious and superfluous, but even that they are morally perverting. Sciences and fine arts, as well as scientific and artistic activities are luxurious, in so far as they presuppose conditions beyond the struggle to survive and the concomitant leisure time (*primum vivere, deinde philosophare*) – a precon-

---

2 Rousseau makes ordinary use of the term in sense of giving over, renouncing, e.g. in the chapter defining what social contract is (*Social Contract*. I. 6).

dition that was already clearly formulated by Aristotle<sup>3</sup>. A further reason the sciences and fine arts are perverting is that they make people more dependent on one another and make them seek recognition<sup>4</sup>. As a consequence, the cultivation of the sciences and fine arts, Rousseau argues, inevitably leads to the weakening of morals and human character. This principal objection to the sciences, the arts, and morals claims at the same time that they contribute to the maintenance of socially-constructed false appearances. The scale scope of these appearances make this predicament even worse, since they are total.

Somewhat surprisingly, there are extraordinary personalities, e.g., Socrates, who are able to neutralize the negative effect of society upon them, and thus to step out of it. Since the more explicit development of these objections can be found only in the *Second Discourse*, where Socrates' case is not developed, we do not learn how this self-liberation in his case could be possible. It remains, however, disturbing, for it seems to imply the possibility of individual resistance to society's negative influences without explaining the structure of this opposition.

It is the predicament of modern man to live in appearances which reproduce day-by-day his situation of mutual slavery. In an often-quoted passage of the *First Discourse* Rousseau writes:

While the Government and the Laws see to the safety and the well-being of men assembled, the Sciences, Letters, and Arts, less despotic and perhaps more powerful, spread garlands of flowers over the iron chains with which they are laden, throttle in them the sentiment of that original freedom for which they seemed to born, make them love their slavery, and fashion them into what is called civilized Peoples. Need raised up Thrones; the Sciences and Arts have made them strong (Rousseau 1997, 6).

Two basic propositions can be discerned here: there is (1) an original freedom which (2) is transformed into an unnoticed slavery made to appear appealing to civilized man. To claim this, Rousseau needs to show there is such a thing as an original freedom, and the hypothesis of the original state serves exactly to justify its existence that freedom. Under closer inspection, a fundamental assumption of Rousseau reveals itself in terms of this passage. Existing as a free being before, man becomes a slave in modern civilization. Amazingly, this is a slavery which remains unnoticed. This is underlined by the formula that pleasant intellectual achievements "throttle" the feeling of freedom. On a descriptive level, this point does justice to the fact that even modern civilized man does not immediately experience any kind of repression.

---

3 It is clearly a consequence of Aristotle's regarding *theoria* to be an activity which does not change the world in contrast to *poiesis* and *praxis*. To contemplate the universe is a non-invasive treatment that presupposes the survival of the contemplator.

4 As Günter Figal remarks, there is a contradiction in refusing the utility of science, on the one hand, and the project of a scientific contribution to the least developed knowledge, to that of self-knowledge, on the other hand (Figal 1991, 101).

In an important footnote Rousseau gives a clue to a better understanding of the slavery: there is a wide range acquired of needs the satisfaction of which makes us dependent on others. These acquired needs, the longing for superfluous things, are tight and artificial constraints on modern individuals, or, as Rousseau expresses it rhetorically: “what yoke could be imposed upon men who need nothing?”<sup>5</sup>. The idea of “acquired needs” as key to modern man’s slavery makes sense obviously only under the assumption that we have a clear set of basic and natural needs. To describe these basic and natural needs, Rousseau uses the hypothesis of the original state of savage man developed only later in the *Second Discourse*.

The ideological character of the sciences and fine arts suggested by the passage is, however, not understandable at all. If Rousseau thinks that the sciences and fine arts could sweeten slavery, then they must be pleasant in themselves. But in this case it remains unclear how they could sweeten slavery, since Rousseau seems to suggest a kind of exchange, a sort of compensation: the slavery understood as mutual interdependence is perfectly counterbalanced by the pleasant activities of science and fine art. However, it is Rousseau’s own use of the ideas of counterbalancing or compensation that gives the lie to his claims about modern enslavement. Slavery seems to be an overstatement which basically could not be compensated by such pleasures. Instead, we could simply accept that division of labour enables activities like sciences and fine arts which were otherwise impossible.

Rousseau’s negative description of modernity can be given two basically different interpretations. Either we say that Rousseau’s fundamental point is mankind’s loss of its essence once for all, or we say that having lost its essence the task of mankind would be to regain something like its essence. The first reading is a sort of history of decline, whereas the second is harder to reconcile Rousseau’s requirements concerning natural and free human life. I exclude from the very outset the biographical reading of Rousseau which would suggest that the rejection of Parisian societies with their norms and practices societies (he comes from far, speaks a somewhat archaic language) forms the basis of his contempt of modern world. Let us see in more detail whether the description of the development of human society in the *Second Discourse* gives a better understanding of what could be called alienation in the present context.

## Rousseau’s “alienation”

In contrast to the *First*, the *Second Discourse* explains the point of reference that is the basis for Rousseau’s evaluation of modern society. Also, Rousseau develops here the point of inequality which turns out to be the basis of the critical assessment of sciences and fine

---

5 Ibid. The parallel passage in the *Second Discourse* also claims, although more carefully, the slavery in question: “man, who had previously been free and independent, is now so to speak subjugated by a multitude of new needs to the whole of Nature, and especially to those of his kind, whose slave he in a sense becomes even by becoming their master” (Rousseau 1997, 170).

arts. The Preface to the opera *Narcisse* underlines that inequality is both a precondition (inequality of leisure) and a major goal (prestige and distinction) of scientific and artistic activity: “A taste for letters always heralds the beginning of corruption on a people... For, in an entire nation, this taste can only rise from two sources, both of them bad, and both of them perpetuated and increased by study, namely idleness and a craving for distinction” (Rousseau 1997, 97)<sup>6</sup>. The refusal of the sciences and fine arts is rooted in the strong pre-supposition that the only goal, or at least the primary goal, of these activities is prestige, i.e., to distinguish oneself from others<sup>7</sup>. In Rousseau’s argumentation there are two major steps: first, he gives a picture of savage man, and secondly, he describes the complicated process by which modern society is constituted.

The picture of savage man serves to represent the original situation of man that had been abandoned step-by-step in the process of civilization. It is here that Rousseau’s line of thought becomes ambiguous. On the one hand he is no doubt aware of the special difficulties implied in genuinely depicting natural human beings<sup>8</sup>. In the Preface to the *Second Discourse*, he asks:

how will man ever succeed in seeing himself as Nature formed him, through all the changes which the succession of times and of things must have wrought in his original constitution, and to disentangle what he owes to his own stock from what circumstances and his progress have added to or changed in his primitive state? like the statue of Glaucus which time, sea, and storms had so disfigured that it less resembled a God than a ferocious Beast (Rousseau 1997, 124).

As the passage clearly suggests, it is the “original constitution” of human beings that Rousseau tries to grasp. He thinks it necessary to proceed on the assumption that “one must seek the first origin of the differences that distinguish men who, by common consent, are naturally as equal among themselves as were the animals of every species” (ibid). Rousseau’s solution of the problem is a sort of ideal standard articulated through a thought experiment. The key passage runs as follows:

For it is no light undertaking to disentangle what is original from what is artificial in man’s present Nature, and to know accurately a state which no longer exists, and which perhaps never did exist, which probably never will exist, and about which it is nevertheless necessary to have exact Notions in order accurately to judge of our present state (Rousseau 1997, 125).

---

6 It is worth mentioning that Rousseau’s view is similar to that of Freud’s sublimation thesis in tracing back scientific and artistic activity to motivations that the actors would not admit and would not be able to recognize. Intellectual and cultural activities are, ultimately, cravings for distinction for Rousseau and sublimated sexual satisfaction for Freud.

7 See Fetscher 1975, 20.

8 Some do not reflect on the status of the original condition of humans, but simply regard it as unproblematic.

Thus, the description of savage man proves to be a theoretical device in order to distinguish the natural from the artificial<sup>9</sup>. The formulation “man’s present Nature” (*dans la Nature actuelle de l’homme*) shows that the idea is not that man moves away from his supra-historical nature, rather that “human nature” seems to be historically changeable. The flexibility of human nature does not really harmonise with Rousseau’s critical approach, since the historically and geographically changing nature of man renders the idea of an original human nature elusive. The letter to D’Alembert explicitly refers to this historic-geographic relativity.

In spite of the explicitly hypothetical character of the original constitution, Rousseau talks in numerous passages about earlier or geographically more distant cultural and social formations as if they contained more elements of the original human nature, or at least they were nearer to that. Distorting pathologies of modern society are highlighted often by contrasts of this kind. The distance from the original state of man is not homogeneous in space and time, because Rousseau depicts natural folks often as less corrupted. In addition, there are differences between continents and within Europe too.

Despite the plasticity of human constitution, Rousseau identifies already in the *Second Discourse* invariant determinations of human nature, first of all perfectibility, love of the self (*l’amour de soi*) and compassion (*pitié*). These elements he considers to be part of human nature everywhere and every time, although they might be weakened as can be seen in the case of self-interest (*l’amour propre*). Universal features of human nature, however, cannot be observed in individuals directly, they show themselves in human reactions and conduct. A weak point in Rousseau’s conception is that persons are not able to justify these reactions, and similarly they can identify false needs without being able to demonstrate their falseness (Struma 2001, 73–74).

For present purposes it is not necessary to follow the complicated declining process of the evolution of human society.<sup>10</sup> Under closer inspection we find in Rousseau’s argumentation a description of an original, natural position containing an image of man which is hard to relate to what we think to be human life. No moral qualities, no reflection and consequently only a reduced form of freedom, almost no sense of time, no self and self-awareness, and a low-intensity attitude towards others – empathy put aside which cannot really be explained<sup>11</sup> – we have here a set of properties and abilities that on the one hand serve as a basis of critique of civilized human life, but on the other hand are not able to

---

9 The hypothetical character appears at the beginning of the *Second Discourse*, too: “Let us therefore begin by setting aside all the facts, for they do not affect the question. The Inquiries that may be pursued regarding this Subject ought not be taken for historical truths, but only of hypothetical and conditional reasonings; better suited to elucidate the Nature of things than to show their genuine origin” (Rousseau 1997, 132).

10 Some remark that this process is confusing: both contingent and causally determining factors play significant role in it (e.g. Struma).

11 Barbara Zehnpfennig observes that empathy is simply against the logic of the natural state, since human beings live isolated and their contact with others is marginal (Zehnpfennig 2013, 180).

offer a plausible concept of human life in general.<sup>12</sup> The savage and the civilized man differ in their innermost intentions, inclinations, and desires:

The first breathes nothing but repose and freedom, he wants only to live and to remain idle, and even the Stoic's *ataraxia* does not approximate his profound indifference to everything else. By contrast, the Citizen, forever active, sweats, scurries, constantly agonizes in search of ever more strenuous occupations: he works to the death, even rushes toward it in order to be in a position to live, or renounces life in order to acquire immortality (Rousseau 1997, 186–187).

It is far from being clear what the point of freedom could be in “living and remaining idle”. Furthermore, Rousseau simply does not consider a kind of Aristotelean approach to happiness, according to which happiness would be a by-product of activities.

The discourse on inequality goes beyond a critical diagnosis of human culture to a conception of positive freedom in the sense of self-sufficiency (autarchy) and self-determination. These implicit ideals are introduced through the critique of false needs and of conditions of recognition. Hence the theory of true needs and of positive freedom are interconnected: the satisfaction of true needs represent an antecedent form of positive freedom (see Struma 2001, 88–89).

In his concluding remarks, Rousseau highlights the contrast between the original constitution and civilized man: “the Savage lives within himself; sociable man, always outside himself, is capable of living only in the opinion of others and, so to speak, derives the sentiment of his own existence solely from their judgment”. And he hints at, without entering into details, “how such a disposition engenders so much indifference to good and evil together with such fine discourses on morality; how everything being reduced to appearances, everything becomes factitious and play-acting: honour, friendship, virtue, and often even vices” (Rousseau 1997, 187).

Rousseau's conception obviously has several weak points, since it cannot establish a distinction between two classes of feelings or states of mind that could be traced back to the distinction of self-sufficient love of the self of the savage man and dependent, egoistic *amour propre* of modern man. A second and more fundamental weak point is the general refusal of any kind of comparison and competition in human life which seems to exaggerate their disadvantages. Rousseau makes comparison responsible for the mutual dependence of human beings, and identifies it with slavery, since it plays an important role in creating false needs; what he establishes, however, is that in stating or articulating a need or a feature we need others, since we rely on their agreement as a kind of guarantee of normal judgment. It is mixing up two functions, if the intersubjective reliability becomes identified with the relationship of mutual dependence.

---

12 Rousseau thinks his picture of the savage man to be anti-Hobbesean, but it is far not so different from Hobbes as he assumes. Rousseau claimed in the *Second Discourse* that theories of an original contract in a situation before any society made the mistake of projecting modern man distorted by society into a position before society. See on this point Kersting's comments on *Contrat social* (Kersting 2002, 20).

Furthermore, society is in Rousseau's view something cannot but have destructive effects. To show that this view, to say the least, is reductive, it suffices to remind that division of labour is presupposed in any higher intellectual achievement that could not be accomplished without satisfying biological needs with the help of others. The narrow-mindedly negative estimation of these achievements is the price Rousseau is obviously ready to pay in order to have a perspective to criticize comfort, luxuriousness, and abundance. The ideal of frugality underlying Rousseau's critique enables him to refuse negative social tendencies in human history, but it cannot, on the other hand, allow higher intellectual achievements the value of which is hard to doubt.

The unreality and implausibility of the original condition of human beings builds the major difficulty in talking about alienation in Rousseau, since it simply makes it impossible for conceptual reasons to overcome the alienated situation in the sense of returning to the original condition or reconstructing it. Non-alienated human life in Rousseau's sense would be not human at all. It does not make any difference, if we discern aspects of alienation in Rousseau. There are, in fact, two alienation-claims in his work: first, that man gets alienated from nature, and second that man gets alienated from his- or herself. From this angle, diremption is the basic problem of stepping out of the original state: diremption with itself and diremption with nature – two in one in Rousseau's version. The obvious non-human character of the original state motivates some interpreters (e.g. Marks 2005) to build their readings on a "middle" state of mankind beyond the natural condition but before modern society. As we shall see, Marx's version of alienation has not this kind of basic unreality, even if it can be doubted whether non-alienated life in Marxian terms, i.e. Communism, could be realized at all. Before turning to Marx, it is worth to have a look at the description of a specific fine art, of theatre where Rousseau follows a similar logic.

## Alienated fine arts

Rousseau radically criticizes the arts, but in a rather unspecific meaning of the term, referring not only to the fine arts. He means by "Art" in the sense of Ancient-Greek *techne* every kind of organized human activity or ability which makes civilized life easier and more convenient. Fine arts are regarded as the ultimate cultural products, and within the framework of the fundamental refusal of alienated human culture in general there is no place to grasp the particular meaning of aesthetic experience.

More specifically, we find a critique of theatre in form of an anti-aesthetics in the letter to D'Alembert, who remarked in his Encyclopaedia-article on Geneva the absence of a theatre. In defence of his city, Rousseau tries to justify its principles, and so he is confronted with a favourite idea of the Enlightenment on the pedagogic value of theatre. Rousseau applies his sharp criticism of luxuriousness based on the values of naturalness and utility on theatre. He arrives by this move – directed originally against Voltaire – to a kind of anti-aesthetics.

Theatre is a frivolous and superficial institution, since it is an institution of luxuriousness and admiration which seeks but entertainment. In order to please the public, theatre cannot criticize the conditions it depicts, but must take people as they are, thereby reinforcing and recognizing what they are with their prejudices, corrupted needs, bad taste, and vices. In the final analysis, theatre is not able to improve the existing conditions, it is simply good for the good and bad for bad persons. Further major point of critique concerns aesthetic distance, which in Rousseau's view negates any practical consequences of fine art. His contempt for art notwithstanding, Rousseau attributes some positive role to the aesthetic. In the *Letter to D'Alambert* he praises pleine-air folk-festivities as events of innocence and purity.

Concluding this section, it has become clear that a rather wide conception of alienation lies in Rousseau's harsh disdain of human culture. The semantic core of this alienation is the loss of original capacities and natural instinctiveness. What Rousseau did not demonstrate, except in a very hypothetical manner, is the identification of human nature itself with original capacities and instinctiveness. To put it another way, the hypothetical character of savage man fails to show why culture and education cannot constitute, at least partly, the human essence. A detailed discussion of the social contract as a proposal for non-alienated human conditions is beyond the scope of this paper. But it is clear that even this reading of the social contract is bound to the assumption that it cannot mean a return to the original naturalness – since the skills and competences established in the process of human society alone make it impossible. New society on the basis of social contract is not a reconstruction of human nature which had been corrupted by the development of human society, but instead a proposal of a new order of the deeply changes human nature basically changed

From our vantage point, Rousseau's description of human civilization provokes a series of questions. First, one might ask whether we can experience or access the loss he had described. Under closer inspection, this turns out to be the problem of the status of the original and natural state and position. Secondly, and it is less important in present context, it is not exactly clear when the process of decline begun. Thirdly, it could be suggested that the social contract represented for Rousseau a kind of regain of the original state, with the advantage of having obtained several abilities and virtues resulting from the civilization process. It is instructive at this point that two readings of Rousseau's claim about the decline of human society can be distinguished. Whereas in the first reading socialization inevitably means the decline of human existence, the second one proposes to identify the points where the process of socialization goes awry in wrong directions. It is obviously only the second reading that is compatible with the proposal of *The Social Contract* to establish a really free political body (Zintl 2009). But even if we agree on the second reading, it might be doubted whether the new social order, provided it is possible to establish it, delivers a non-alienated state of mankind. In other words, people may become free after having made the social contract in Rousseau's sense, but they remain deeply different from to the savage man as described in the *Second Discourse*.

To sum up, talking about alienation in Rousseau is complicated by the fact that the point of departure of the alienation process remains unspecified. The basic point of my

argument has concerned this initial situation in Rousseau's analysis which cannot be regarded as a human situation at all. One could object to the presentation given above that Rousseau's problem exactly fits the scheme of alienation as described at the beginning of this paper: loss of something previously possessed. Rousseau's claim is, according to this objection, that human nature developed on the wrong way, and it is precisely human nature that disappears in the course of human history. The objection can be answered by considering the hypothetical status of human nature. If human nature is hypothetical, viz. a thought experiment as a methodological device, then it is not alienation in a narrower sense that could have taken place. For this reason, I propose to label Rousseau's description as a conception of alienation only in a broader sense. However, it does not mean a solution to the rest of the problem – i.e. what is the precise argumentative significance of human nature in Rousseau? It is open to debate whether his critique of the Enlightenment one-sided appreciation of rationality can be formulated in a less radical way so that it does not fall into a similarly one-sided overestimation of feeling and sentiment.

Let us now take a brief glance at Marx' theory of alienation.

## **Alienation in Marx**

As shown above, Rousseau described problematic features of modern man which could be thought of as alienation, but only in a broader sense. In contrast, a more specific conception of alienation can be explicitly found in Karl Marx's thought. The following sketch does not try to give a detailed account of alienation in Marx, but highlights important conceptual differences between his and Rousseau's theory.

In Marx's work there is a shift from alienation in the early Paris Manuscripts (*Ökonomisch-philosophische Manuskripte*) to reification and objectification in the later work (*Kritik der politischen Ökonomie*). It is debated whether this means a break in the treatment of or even an abandoning of the issue, or rather implies the presence of the topic in the whole work.<sup>13</sup> Be it as it may, normative basis of alienation in the early Marx is the concept of man's self-realization in the working process: "labor is the self-realizing

---

13 On various positions see Kolakowski 1978, 263ff. It lies beyond the scope of this paper to clarify the reasons for the shift from the early manuscripts' description of alienation to the later works. Kolakowski makes a case for the continuity thesis by claiming that the Paris Manuscripts "are in effect the first draft of the book that Marx went on writing all his life, and of which *Capital* is the final version" (Kolakowski 1978, 132–133). Tilman Reitz's proposal interestingly differentiates between what philosophy definitely cannot offer (revolutionary changes) and its actual functions (to support ideological agreement with the existing order). Kübler remarks that we do not find any justification of the refusal of capitalism in the later work, only in the *Manuscripts* (Kübler 2013). Zehnpfennig claims that there is no strict separation of the alienation-theorem and the later critique of capitalism: "Seine im *Kapital* entwickelte Kapitalismuskritik und seine Revolutionstheorie lassen sich im Grunde gar nicht verstehen, wenn es nicht die in der Entfremdungstheorie beschriebenen Defizite wären, die durch die Revolution behoben werden sollen" (Zehnpfennig 2013, 185).

human activity”<sup>14</sup>. The self-realization, in turn, takes place in a double movement of a prior objectification and a following re-appropriation.

As is well-known, the early Marx claimed that in capitalism labour cannot be but alienated.<sup>15</sup> He talks about alienation of the worker in four different senses: he is alienated a) from the product of his work, b) from the process of his working, c) from species-being (*Gattungswesen*) – i.e. man is not exercising activities proper to true human nature and capacities –, and finally d) from others. Considering the interdependence of these aspects, the essential point, in my view, can be found in the second form, since the first alienation is a consequence of the “alienation within the activity of work itself” (“in der Tätigkeit der Arbeit selbst”) which is a kind of self-alienation (*Selbstentfremdung*) of the worker. Talking about self-alienation means that the working activity is “independent” (*äußerlich*) from the worker, it does not belong to his essence, it is forced labour (*körperlich und geistig ruinöse Zwangsarbeit*), so that it is the exact opposite of work that should be self-realization in the sense of “free psychic and intellectual energy”<sup>16</sup>. One should nevertheless bear in mind that this definition of labour or work could easily be challenged – a problem which cannot be discussed here in detail.<sup>17</sup>

Marx’s conception of work as a contrary to counter-conception of self-realization involves contains the characterization of work as “abstract”. Marx follows in this context Adam Smith’ description of the poverty of workers, and considers his identification of work with pain as a naturalization of alienated work<sup>18</sup>. Marx regarded property as something that should be explained, not simply accepted, as Smith and Locke did. Whereas he explicitly acknowledges categories and “laws” of national economy, he refuses this discipline as being an ahistorical perspective without offering a basic principle for the explanation of property.<sup>19</sup>

---

14 “Die Selbstverwirklichung des Menschen bedeutet hier im ersten Schritt die Entäußerung und Vergegenständlichung seiner Fähigkeiten und Anlagen und im zweiten Schritt die Wiederaneignung des Entäußerten, durch die zugleich der arbeitende Mensch sich bildet und seine Bestimmung, d.h. sein Gattungswesen, realisiert” (Lohmann 1991, 23).

15 Ingo Elbe emphasizes that Marx analyzed alienated labour, not alienation in general (Elbe 2014, 7–8). He also refers to Marx’s remark against an „imagined original position” (*erdichteter Urzustand*) which he criticised in Adam Smith. The direction of this critique, however, saying that Smith presupposes what should be explained reminds the similar critique of Rousseau on Hobbes.

16 Der Arbeiter fühlt sich “nicht wohl, sondern unglücklich [...], [...] fühlt sich daher erst außer der Arbeit bei sich und bei der Arbeit außer sich” (Marx 1968, 514).

17 It is not here to discuss an alternative conception to this in the work of Mihály Csíkszentmihályi. His deeply Aristotelean conception of “flow” develops the basic point that each activity, even monotone and mechanic ones, might be the source of a pleasant contentedness.

18 As to Marx’s relationship to Rousseau, one could accentuate the refusal of the construction of an original non-alienated position as Elbe does. He does not, however, take into account that Marx’s critique fails to realize that Rousseau’s reference to the original position is, as indicated above, a thought experience, not a factual description of history.

19 “Die Nationalökonomie geht vom Faktum des Privateigentums aus. Sie erklärt uns dasselbe nicht” (Marx 1968, 510).

Without getting lost in the complexities of Marx's conception of alienation, it can be claimed that he thinks the transformation of alienated work into a non-alienated situation possible. Kolakowski regards a series of "critiques" of Marx – including among others the *Paris Manuscripts* and *Capital* itself – as more and more elaborated versions of the same basic idea which he formulates as follows:

We live in an age in which the dehumanization of man, that is to say the alienation between him and his own works, is growing to a climax which must end in a revolutionary upheaval; this will originate from the particular interest of the class which has suffered the most from dehumanization, but its effect will be to restore humanity to all mankind (Kolakowski 1978, 262).

The main purpose of the process of history, in Marx's view, is nothing but a situation without private property, i.e. Communism. At this point, it is enough to emphasize that even if the realization of Communism might be regarded as problematic from a practical point of view, in Marx's eyes it would undoubtedly mean a non-alienated state. With this we have a basically different feature of alienation in Marx as compared to Rousseau, since the former thinks a sequence of possession – disappropriation – reappropriation possible, whereas the latter has but a short version in the form of hypothetical possession – disappropriation.

## Conclusion

To conclude, the paper has shown that Rousseau's ideas might be labeled as alienation only in a broader sense. The main reason for this is that the semantic core of alienation, i.e. something's becoming strange or foreign to someone, does not exactly correspond to Rousseau's problem. As a first step, a sketch of Rousseau's critique of modern society has been given with special attention to pathologic features and their sources. Secondly, the essay has examined the basis of his critique, and focused on the hypothetical character of his account of man's original constitution. It has been demonstrated that man's original constitution is both unreal and implausible, and non-alienated life in Rousseau's sense is no human life at all. Finally, Rousseau's theory has been opposed to the more specific conception of Marx on alienating effects in society. Nonetheless, the paper has also pointed out that despite conceptual differences both Rousseau and Marx had difficulties with giving an account of non-alienated conditions. Talking about alienation seems to face again and again the problem of finding an Archimedean point of non-alienation.

# Bibliography

- Elbe, Ingo. 2014. Entfremdete und abstrakte Arbeit. Marx' Ökonomisch-philosophische Manuskripte im Vergleich zu seiner späteren Kritik der politischen Ökonomie. In: Elbe, Ingo – Hogh, Philip – Zunke, Christine (Hrsg.): *Oldenburger Jahrbuch für Philosophie 2012*. Oldenburg: BIS-Verlag. 7–69. ([http://www.rote-ruhr-uni.com/cms/IMG/pdf/Entfremdete\\_Arbeit.pdf](http://www.rote-ruhr-uni.com/cms/IMG/pdf/Entfremdete_Arbeit.pdf); downloaded 2016.08.16.)
- Fetscher, Iring. 1975. *Rousseaus politische Philosophie*. Frankfurt/M: Suhrkamp.
- Figal, Günter. 1991. Philosophische Zeitkritik im Selbstverständnis der Modernität. Rousseaus *Erster Discours* und Nietzsches zweite *Unzeitgemäße Betrachtung*. In: Figal, Günter – Roif-Peter Sieferle (Hg.): *Selbstverständnisse der Moderne. Formationen der Philosophie, Politik, Theologie und Ökonomie*. Stuttgart: Metzler, 100–132.
- Henning, Christoph. 2015. *Theorien der Entfremdung*. Hamburg: Junius.
- Iorio, Marco. 2013. Veränderung, Verdinglichung, Entfremdung. Über Marxens verhegelt-verhaelte Ontologie. In: Jaeggi, Rahel – Loick, Daniel (Hrsg.): *Karl Marx – Perspektiven der Gesellschaftskritik*. Berlin: Akademie Verlag.
- Jaeggi, Rahel. 2014. *Alienation*. New York: Columbia University Press.
- Kersting, Wolfgang. 2002. *Jean-Jacques Rousseaus „Gesellschaftsvertrag“*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Kolakowski, Leszek. 1978. *Main Currents of Marxism: Its Rise, Growth and Dissolution*. Vol. 1: The Founders. Oxford: Clarendon Press.
- Kübler, Lukas. 2013. Marx' Theorie der Entfremdung. In: Jaeggi, Rahel – Loick, Daniel (Hrsg.): *Karl Marx – Perspektiven der Gesellschaftskritik*. Berlin: Akademie Verlag. 47–66.
- Lohmann, Georg. 1991. *Indifferenz und Gesellschaft. Eine kritische Auseinandersetzung mit Marx*. Frankfurt am Main: Suhrkamp.
- Marks, Jonathan. 2005. *Perfection and Disharmony in the Thought of Jean-Jacques Rousseau*. Cambridge: Cambridge University Press.
- Marx, Karl. 1968. Ökonomisch-philosophische Manuskripte aus dem Jahre 1844. In: *Karl Marx – Friedrich Engels Werke*. Vol. 40. Berlin: Dietz Verlag. 465–588.
- Rousseau, Jean-Jacques. 1997. *The Discourses and other early political writings*. Cambridge: Cambridge University Press.
- Struma, Dieter. 2001. *Jean-Jacques Rousseau*. München: C. H. Beck.
- Zehnpfennig, Barbara. 2013. Rousseau und Marx: das Ende der Entfremdung. In: Hidalgo, Oliver (Hg.): *Der lange Schatten des Contrat social. Demokratie und Volkssouveränität bei Jean-Jacques Rousseau*. Wiesbaden: Springer VS. 177–209.
- Zintl, Reinhard. 2009. Rousseau. In: Becker, Michael – Schmidt, Johannes – Zintl, Reinhard: *Politische Philosophie*. Paderborn: Schöningh. 64–76.

# Contributors

**Zsolt Bagi** is senior lecturer in philosophy at the University of Pécs. He has obtained his PhD in 2010, he is working in a number of fields: aesthetics, political philosophy and 17<sup>th</sup> century rationalism. He is author of three books on political aesthetics published in Hungarian, including *The Phenomenology of the Language of Literature*. (Budapest, Balassi, 2006.) He is currently interested in philosophy of locality, presenting universalism without globalism. His other project involves studying the present possibilities of aesthetic emancipation by a double approach: participatory through sensual empowerment and constructive through dissensual art.

**Dávid Bartha** began his studies for a PhD at the Central European University in 2014. His main research area is the philosophy of George Berkeley, paying special attention to his relation to Malebranche. He published articles on Berkeley's philosophy of mind and theory of self-knowledge as well as on his views on archetypes. His dissertation deals with the question of theological voluntarism in Berkeley.

**Gábor Boros** is professor of philosophy in the Institute of Philosophy of the Faculty of Humanities of Eötvös Loránd University (ELTE), Budapest. He is the director of research of the NRDIO-funded project on "Self-interpretation, emotions, narrativity" (K120375). He translated several works of both 17<sup>th</sup> century authors and contemporary theoreticians. He published extensively on Descartes, Spinoza, Leibniz and the philosophy of emotions, particularly of love. His most recent achievement is a tripartite study on the concept of love in the 17<sup>th</sup> and 18<sup>th</sup> century in the *Fudan Journal of Humanities and Social Sciences*. Forthcoming with Brill is a comparative investigation on the concepts of love in China and in Europe together with Paolo Santangelo and Roberto Gigliucci.

**Dan O'Brien** is Reader in Philosophy and Subject Co-Ordinator for Philosophy at Oxford Brookes University. His research interests include the work of David Hume, the epistemology of testimony and trust, and conceptions of the divine mind. He is the founder and organizer of the Oxford Hume Forum and the annual Brookes International Hume Workshop. He has led funded research clusters on intellectual humility (2015) and special divine action (2016). His publications on Hume include the *Bloomsbury Companion to Hume* (2015), the *Continuum Reader's Guide to Hume's Enquiry* (2007), and *Hume's Critique of Religion* (Springer 2013). His most recent publication is 'Hume on Education' (*Pacific Philosophical Quarterly* 2017), and he is currently writing a monograph on Hume on testimony. A 2<sup>nd</sup> edition of his *Introduction to the Theory of Knowledge* (Polity 2006, 2017) has just been published.

**Filip Buyse** is a visiting researcher of the History Faculty at the University of Oxford, a visiting fellow of the Centre for the History and Philosophy of Physics at St Cross College in Oxford and fellow at the Vossius Center of the University of Amsterdam. He received his Master of Science and MA in Philosophy from the University of Louvain before completing an interuniversity DEA in the Philosophy of Science at the ULB, the UCL and the ULg. In 2014, he received his PhD in Philosophy from the University Paris 1 Panthéon-Sorbonne. His main interest includes epistemology, metaphysics and, the history and philosophy of early science. Dr. Filip Buyse edited a special issue of the *Intellectual History Review* on Galileo and Spinoza and is the editor of two special issues of the *Society and Politics*: one on Letters by Early Modern Philosophers and one on the Pendulum Clock in Seventeenth Century Philosophy.

**Christopher Davidson** is an assistant professor at Ball State University in the department of Philosophy and Religious Studies. He was awarded his PhD at Villanova University in 2014. His dissertation analyzes Foucault's last books on Ancient ethics and also extends those works' methods to Spinoza. The theme of philosophy as a way of life unites his work on Spinoza, Foucault, and Stoicism. He is the author of "Spinoza as an Exemplar of Foucault's Spirituality and Technologies of the Self" and "Foucault on Askesis in Epictetus: Freedom through Determination," and a co-editor of *New Philosophies of Sex and Love: Thinking Through Desire* (London: Rowman & Littlefield International, 2016).

**Jan Forsman** is a PhD student in Philosophy at the University of Tampere. His work concentrates mainly on early modern philosophy but he is also interested in ancient, medieval and renaissance philosophy, metaphysics, epistemology and philosophical skepticism. He is currently working on a thesis on Descartes's skepticism in the *Meditations*, scheduled to be finished in Autumn 2019. In the meanwhile, he writes, plays basketball and takes care of his home zoo of dogs, cats and turtles with his wife.

**Brian Glenney** is an Assistant Professor of Philosophy at Norwich University publishing on the history and philosophy of sensory perception, from ancient descriptions of blindness

in journals such as the *Journal for the Study of the New Testament*, to contemporary empirical research in journals such as *Biology and Philosophy*. He is currently co-editing a volume of papers on problems of perception from diverse historical periods and regional areas with Routledge.

**Keith Green** is Professor in the Department of Philosophy and Humanities at East Tennessee State University (ETSU). His publications focus on the moral psychology of hatred in Spinoza and Aquinas, and, more broadly, responses to the challenge of naturalism within the context of early modern Christian ethics and philosophy of religion. He is also coordinator of programs in religious studies at ETSU, and a community member of the ethics forum/committee of Mission Health Systems in Asheville, North Carolina.

**Maximilian Kiener** is a graduate student in Philosophy at the University of Oxford (DPhil). Specializing in medical ethics and the philosophy of law, Kiener is also interested in themes in early modern philosophy, particularly those from which modern debates in the aforementioned areas have inherited their methodological tools. His essay on affectivity is part of a larger project on voluntariness and compulsion and hopes to bring to greater attention Descartes' view on a genuinely epistemic kind of affectivity.

**Davide Monaco** is a PhD student at the University of Aberdeen, working under the Normativity Project Award Scheme. He obtained his Master's degree at the University of Naples "Federico II" and published on Ernst Mach's epistemological monism and parallelism. Afterward, he has moved to Spinoza's philosophy and focused on these concepts in a metaphysical perspective. In his dissertation, the relationship between Spinoza's substance monism and his parallelism is investigated from a different angle. The main aim of this research is to demonstrate the compatibility and identity of the two distinct formulations of Spinoza's parallelism theory in the *Ethics*.

**Hans D. Muller** obtained his PhD in Philosophy in 2003 from the University of North Carolina at Chapel Hill. Currently Associate Professor of Philosophy at the American University of Beirut, his work focusses on contemporary philosophy of mind and cognitive science as well as developments in moral psychology during the Scottish Enlightenment. He is a co-editor of and contributor to *Contemporary Philosophical Naturalism and Its Implications* (2014), part of the *Routledge Studies in the Philosophy of Science*. Professor Muller is the author of several recent articles on the role of sympathy in the philosophy of both Adam Smith and David Hume, including "Sympathy for Whom? Smith's Reply to Hume" which appeared in the *Journal of the American Philosophical Association* (2016). His contribution to the current volume is a part of that ongoing research project, which will culminate in a book defending sentimentalist ethics against its rationalist critics.

**Csaba Olay** studied philosophy, mathematics, and physics at Eötvös Loránd University Budapest and obtained his PhD in philosophy at Freiburg University (Germany), with

a scholarship of the German Academic Exchange. He has been teaching at Eötvös Loránd University since 2001 and is currently the Head of Department of Modern and Contemporary Philosophy, where he was appointed as full professor in 2015. His main research areas are 19-20th century continental philosophy, hermeneutics, Heidegger, Gadamer, Hannah Arendt, and Frankfurt School. He has published books and articles, and edited collections of essays in these fields.

**Ádám Smrcz** is currently an assistant research fellow at the Institute of Philosophy of the Hungarian Academy of Sciences. He was a student of the Doctoral School of Literary Studies from 2012, and that of the Doctoral School of Philosophy from 2014 at ELTE University. His main fields of interest are early moderns Stoicism and Platonism.

**Judit Szalai** teaches at Eötvös Loránd University, Budapest. Her main interests are seventeenth-century thought, philosophy of mind and philosophical psychology. Her recent publications are on the emotions, rationality/irrationality, and the non-conscious, in historical and contemporary contexts.

**Olivér István Tóth** is PhD student at Eötvös Loránd University Budapest (ELTE) and Alpen-Adria-Universität Klagenfurt (AAU). He is writing his dissertation, under the supervision of Gábor Boros and Ursula Renz, on the relationship of imagination and intellect in Spinoza's philosophy of mind. He has published articles on Spinoza's philosophy of mind and Hegel's objective spirit. He is member of the interdisciplinary research group on Adorno at ELTE (leading researcher: Csaba Olay), external member of "Spinoza on the Concept of the Human Life Form: Towards a Non-Essentialist and Ontologically Liberal Account" research project (leading researcher: Ursula Renz) and junior researcher of "Consciousness in medieval and early modern philosophy of mind" CENTRAL Kolleg Berlin.

**Hanna Vandebussche** is teaching assistant and doctoral researcher at Institute of Philosophy, KU Leuven (University of Leuven) with a co-tutela at the Université de Bourgogne in Dijon. Her research focus is on early modern French philosophy, with a dissertation on the notion of the self ('le moi') in the (later) works of Descartes and Pascal. Recent publications include: 'Descartes en Pascal over de passies en de (on)mogelijkheid van een moraal' (Tijdschrift voor Filosofie, 2015), 'Levi et Pascal et la question de l'humanité' (Transversales, 2016) and 'Passies sprekend tot en met de verbeelding. Descartes over ijdelheid, liefde en genot' (Algemeen Nederlands Tijdschrift voor wijsbegeerte, 2017).

# Index of Names

## A

Abelard, Peter 143  
Adler, Jacob 160, 167  
Airaksinen, Timo 224, 235  
Akkerman, Fokke 93, 99, 127, 128, 139, 140  
Alain, (Emile Auguste Chartier) 64  
Alanen, Lilli 166, 244  
Alexander of Aphrodisias 167  
Allison, Henry 240  
Alquié, Ferdinand 53  
Althusser, Louis 202  
Andreae, Tobias 83, 85–87  
Ansaldi, Saverio 196  
Ardal, Páll S. 247  
Aristotle 16, 167, 168, 189, 273  
Arnauld, Antoine 25, 27  
Aquinas, Thomas 19  
Augustine 152, 157  
Ayers, Michael 91

## B

Baceski, Tina 251  
Baier, Annette 242, 244, 245

Bal, Mieke 185, 189  
Balibar, Étienne 197, 212  
Beauval, Basagne de 115  
Bennett, Jonathan 175, 177–182  
Bennett, Matthew 111  
Berkeley, George 221–239  
Berman, David 227  
Birch, Thomas 112  
Black, Deborah L. 165  
Blumenberg, Hans 152, 155  
Blyenberg, Willem 132  
Boros Gábor 9, 125, 142, 143, 159,  
161  
Bouillier, Francisque 85  
Bourdin, Pierre 46, 52  
Bouwsma, William J. 13  
Boxel, Hugo 208  
Boyle, Robert 104, 105, 108, 112,  
120  
Bras-Cléro, Gérard – Jean-Pierre 62  
Bricke, John 245  
Brooke, Christopher 11  
Broughton, Janet 26, 44–46, 49–53  
Brown, Charlotte 242, 245  
Burman, Frans 34, 35, 53, 83  
Butler, Ronald 245, 246

## C

Calcidius 90  
Calder, Simon 189  
Calvin, John 13, 14, 16, 20  
Camerer, Theodor 127, 139, 140  
Capaldi, Nicholas 243  
Carey, David P. 176  
Carraud, Vincent 67  
Carriero, John 39, 47  
Carter, Richard B. 76  
Caterus, Johannes 97  
Chanut, Pierre Hector 60, 63, 69, 73  
Cicero 16  
Cimini, Amy 189  
Clark, Stephen 222, 223, 230  
Clauberg, Johannes 83–85  
Clerselier, Claude 43, 45, 48, 51, 52  
Colerus, John 185  
Comenius (Jan Amos Komenský) 114,  
115  
Coplan, Amy 241  
Coster, Salomon 113  
Cottingham, John 26, 38, 39, 47, 48,  
91  
Cronin, Timothy J. 91  
Csíkszentmihályi Mihály 281  
Cunning, David 40  
Curley, Edwin 39, 43, 47, 55, 106, 124,  
125, 127, 128, 132, 133, 138, 139,  
140, 164, 174

## D

D'Alembert, Jean le Rond 276  
Danto, Arthur 189  
Darwall, Stephen 223  
Davidson, Christopher 185, 196  
Davidson, Herbert A. 167, 168  
Da Vinci, Leonardo 113  
Debes, Remy 241, 245

De' Dondi, Giovanni 113  
Desmond M. Clarke 63  
De Witt, Johan and Jan (brothers)  
218  
Deleuze, Gilles 118, 119  
Della Rocca, Michael 26, 39, 42–44,  
47, 78, 55, 129, 160, 162, 167,  
182  
Del Medigo, Elijah 167  
Descartes, René 24–102, 113, 120, 143,  
150, 174, 217, 224  
Des Chene, Dennis 78  
Deugd, Cornelis de 188, 189  
Diderot, Denis 147  
Dilthey, Wilhelm 152  
Doney, Willis 26

## E

Edgley, Roy 31  
Egenhofer, Sebastian 189  
Elbe, Ingo 281  
Evans, Gareth 31

## F

Falkenstein, Lorne 240, 253  
Farr, James 245  
Fetscher, Iring 275  
Figal, Günter 273  
Firth, Roderick 259  
Flage, Daniel 225  
Foucault, Michel 195, 202  
Fraenkel, Carlos 167  
Frankfurt, Harry G. 26, 32, 33, 35, 39,  
43–46, 50–53  
Frege, Gottlob 32  
Freud, Sigmund 275  
Fricker, Elizabeth 240  
Frigo, Alberto 59, 62, 68, 71

## G

Galenus Pergamensis 82  
Galilei, Galileo 114, 116  
Garrett, Don 131, 164, 174, 176,  
196  
Gassendi, Pierre 44, 48, 50, 53  
Gatens, Moira 189  
Gerbino, Giuseppe 242  
Gersonides (Levi ben Gershon) 167  
Gerwartz, Alan 26  
Geulincx, Arnold 115, 116  
Gilson, Étienne 53, 91  
Glazemaker, Jan Hendrik 139, 140  
Glossop, Ronald 259  
Goalenius, Rodolphus 94  
Goldie, Peter 241  
Gosling, Justin 14  
Gouhier, Henri 53  
Grant, Brian 43  
Graver, Margaret R. 12, 16, 17  
Greco, Lorenzo 143, 144, 244, 247  
Green, Keith 124, 149  
Grey, John 166  
Guattari, Felix 189  
Guenancia, Pierre 61, 66, 67, 71, 72  
Gueroult, Martial 53, 117

## H

Harris, James 243, 246, 247  
Harrison, Jonathan 259  
Harvey, Warren Zev 167  
Henning, Christoph 271  
Heraclitus 142, 152, 154, 155  
Herdt, Jennifer 242–246, 251  
Hicks, G.D. 222  
Hieronymi, Pamela 36  
Hirschmann, Nancy 243  
Hobbes, Thomas 91, 106, 151, 213, 215,  
277, 281

Holden, Thomas 261  
Holt, Edwin Bissell 261  
Hübner, Karolina 190  
Hugo, André Malan 13  
Hume, David 135, 240–254, 257–269  
Huygens, Chrisitan 106, 109–122  
Huysmans, J.K. 195

## I

Israel, Jonathan 104, 109, 185

## J

Jackson, Stanley W. 82  
Jaeggi, Rahel 271  
James, Susan 189  
Jaquet, Chantal 105  
Joyce, James 195

## K

Kambouchner, Denis 42, 69  
Kant, Immanuel 189, 257  
Keesing, Elisabeth 109  
Kemp Smith, Norman 243, 246,  
253  
Kenny, Anthony 26, 33, 28, 29, 43, 47,  
59, 60, 65, 69  
Kepler, Johannes 113, 114  
Kersting, Wolfgang 277  
Kircher, Athanasius 112, 113  
Kisner, Matthew J. 165  
Klein, Julie R 159, 160, 167  
Klever, Wim 109  
Klibansky, Raymond 90  
Kolakowski, Leszek 280, 282  
Koyré, Alexandre 96  
Kübler, Lukas 280

## L

Lackey, Jennifer 240  
Laerke, Mogens 96, 159  
La Forge, Louis 83–85  
Land, J. P. N. 127, 139  
Larmore, Charles 28–30, 32  
Lebuffe, Michael 175, 177  
Leibniz, Gottfried Wilhelm 13, 115, 116,  
118, 120, 208  
Lenz, Martin 159, 168, 170, 175, 176,  
179, 182  
Licata, Giovanni 167  
Lipsius, Justus 11–22  
Locke, John 223, 281  
Loeb, Louis 28, 30, 240  
Lohmann, Georg 281  
Lord Beth 94, 189,  
Lucchese, Filippo Del 200  
Lyotard, Jean-François 189

## M

MacArthur, David 44, 56  
Machiavelli, Niccolò di Bernardo dei 200  
Maimonides (Moses ben Maimon) 167,  
169, 170  
Malebranche, Nicolas 61, 113  
Malinowski-Charles, Syliane 159  
Mandeville, Bernard 135  
Marks, Jonathan 278  
Marshall, Eugene 159, 196  
Marx, Karl 271, 278, 280–282  
Matheron, Alexandre 212, 214, 215  
McHugh, Conor 36  
Melamed, Yitzhak Y. 167  
Mersenne, Marine 33, 34, 91  
McIntyre, Jane 258, 260, 261  
McMyler, Benjamin 240  
Mercer, Philip 243, 245  
Mignini, Filippo 189, 190

Miller, Jon 196  
Molina, Luis de 207, 208  
Moran, Richard 31  
Moray, Robert 112  
Moriarty, Michael 67  
Morrison, James 188  
Morrison, John 182  
Moses 198, 200, 201  
Mulier, Eco Haitsma 200  
Muller, Hans 257, 266

## N

Naaman-Zauderer, Noa 42, 47  
Nadler, Steven 109, 160, 166, 167, 170, 185  
Negri, Antonio 209  
Nelson, John 26  
Newman, Lex 38, 39, 42, 43, 47, 48, 49, 114  
Nicole, Pierre 61  
Noonan, Harold 241  
Norton, David F. 253  
Novotný, Daniel D. 95  
Nuchelmans, Gabriel 91

## O

Oldenburg, Henry 104, 105, 108, 112,  
113, 119, 121  
Olscamp, Paul J. 223  
Oresme, Nicole 113  
Oestreich, Gerhard 11–13  
Ovid 185  
Owens, David 35, 36

## P

Pascal, Blaise 59–62, 66–73, 152, 155–157,  
208  
Paul 71, 73, 152, 157

Perin, Casey 47, 48  
Pikovsky, Arkady 110  
Plato 90  
Plug, Cornelis 175  
Pocock, J. G. A. 200  
Postema, Gerald 137, 243, 245, 260, 266  
Pylyshyn, Zenon 176

## R

Ragland, C.P 40  
Ravven, Heidi Morrison 196  
Rawes, Peg 189  
Rawls, John 259  
Regius, Henricus 41  
Reitz, Tilman 280  
Rembrandt, Harmenszoon van Rijn 185  
Renz, Ursula 143, 159, 160, 162, 167, 169, 177  
Rivaud, Albert 107–109  
Roannez, Mlle 155  
Rosenthal, David M. 39, 42, 47, 56  
Ross, Helen 175  
Ross, James F. 94  
Rousseau, Jean-Jacques 135, 271–282  
Rubens, Peter Paul 147  
Rubin, Ronald 28–30, 32  
Ryan, Sharon 36

## S

Sangiaco, Andrea 160, 162, 167, 177  
Sartre, Jean-Paul 68  
Sayre-McCord, Geoffrey 252, 259  
Schleisser, Eric 242  
Schüssler, Rudolf 38–40, 42, 44, 47, 49, 53  
Schmaltz, Tad M. 97  
Schmidt-Biggemann, Wilhelm 152  
Schmidt, Jeremy 82  
Schneider, Daniel 193

Schuller, Georg Hermann 208  
Sellars, John 13, 18, 19  
Seneca 13, 147, 162, 185  
Shaftsbury, Anthony Ashley Cooper 135  
Shakespeare, William 159  
Shah, Nishi 36, 55  
Shapiro, Lisa 39, 42, 75, 77, 79, 82, 175, 176, 178–181  
Shem Tov ben Shem Tov 167  
Shieber, Joseph 240  
Skinner, Quentin 200  
Smith, Adam 135, 136, 138, 242, 243, 246, 266, 281  
Smrcz Ádám 11, 19, 20223, 272, 273  
Socrates 223, 272, 273  
Sommers, Fred 175  
Spinoza, Baruch de 89–102, 104–122, 124–140, 142–157, 160–169, 174–182, 185–204, 207–219  
Sprecht, Rainer 85  
Spruit, Leen 127, 138, 185  
Spyk, Hendrik Van 185  
Steenbakkens, Piet 127, 139, 142, 143, 152  
Steinberg, Diane 177  
Steinberg, Justin 131, 181  
Steno, Nicholas 157  
Steup, Matthias 36  
Struma, Dieter 271, 276, 277  
Suarez, Francisco 19, 40, 90, 94, 95, 98

## T

Taylor, Charles 11  
Taylor, Jacqueline 243, 245, 246–248  
Terence 185  
Themistius 168  
Totaro, Pina 127, 139, 185  
Tóth Olivér István 143, 159, 190  
Traiger, Saul 254  
Trevisani, Francesco 84, 86, 87  
Tydemann, Daniel 185

## U

Uhlmann, Anthony 189

## V

van Bunge, Wiep 89, 109

Van Cleve, James 26, 27

van de Ven, Jeroen 103

Van Gent, Pieter 139, 140

Vardoulakis, Dimitris 185, 189

Vinciguerra, Lorenzo 189

Visi Tamás 168

Vitz, Rico 38, 44, 47, 49, 53, 55, 248, 249

Voltaire (François-Marie Arouet) 278

## W

Waldow, Anik 242, 245

Wallingford, Richard of 113

Walski, Gregory 164

Warnock, G. 222, 223

Wee, Cecilia 39, 40

Welbourne, Michael 90, 97

Wells, Norman J. 90, 97

Wiesenfeld, Kurt 111

Williams, Bernard 35, 39, 42, 44, 49

Wilson, Margaret D. 39, 44, 49, 170,  
174

Wood, William 66, 67, 71, 72

Wolfson, Harry Austin 145

## Y

Yakira, Elhanan 93, 96

## Z

Zehnpfennig, Barbara 271, 276, 280

Zintl, Reinhard 279



There is no need to argue for the relevance of affectivity in early modern philosophy. When doing research and conceptualizing affectivity in this period, we hope to attain a basic interpretive framework for philosophy in general, one that is independent of and cutting across such unfruitful divisions as the time-honored interpretive distinction between “rationalists” and “empiricists”, which we consider untenable when applied to 17th-century thinkers.

Our volume consists of papers based on the contributions to the First Budapest Seminar in Early Modern Philosophy, held on 14–15 October 2016 at Eötvös Loránd University, Budapest. When composing this volume, our aim was not to present a systematic survey of affectivity in early modern philosophy. Rather, our more modest goal was to foster collaboration among researchers working in different countries and different traditions. Many of the papers published here are already in implicit or explicit dialogue with others. We hope that they will generate more of an exchange of ideas in the broader field of early modern scholarship.



ISBN 978-963-284-820-4



9 789632 848204